OCR Core 1 Module Revision Sheet

The C1 exam is 1 hour 30 minutes long. You are **not** allowed **any** calculator[1].

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted. . . always an area prone to error.

J.M.S.

Preliminaries

- Changing the subject of an equation. For example in $y = \sqrt{x+6}$, y is the subject of the equation. To change the subject to x we merely need to re–arrange to $x = y^2 - 6$. A harder example is make to make g the subject of $T = 2\pi\sqrt{\frac{l}{g}}$.

$$2\pi\sqrt{\frac{l}{g}} = T \qquad \Rightarrow \qquad \frac{l}{g} = \left(\frac{T}{2\pi}\right)^2 \qquad \Rightarrow \qquad g = l\left(\frac{2\pi}{T}\right)^2 = \frac{4\pi^2 l}{T^2}.$$

- To solve simultaneous equations, isolate x or y from one of the equations and substitute into the other. For example, solve

$$x + 2y = 1,$$
$$x^2 - 2y^2 = 31.$$

From the first we find $x = 1 - 2y$ and putting into the second we find $(1-2y)^2 - 2y^2 = 31$, which simplifies to $y^2 - 2y - 15 = 0$. This gives $y = -3$ or $y = 5$. To calculate the x values we put the y solutions into either original equation. The solutions are $(-9, 5)$ and $(7, -3)$. [Give your solutions as coordinates to show which x and y values go together.]

- If you ever need to find where two lines or curves cross, then merely view it as a pair of simultaneous equations to be solved.

- You must also know how to handle algebraic fractions and how to write two algebraic fractional expressions as one fraction. The general rules are

$$\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd} \qquad \text{and} \qquad \frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}.$$

Therefore to write $x - \frac{x}{x+1}$ as a single fraction we do the following

$$x - \frac{x}{x+1} = \frac{x}{1} - \frac{x}{x+1} = \frac{x(x+1) - x}{x+1} = \frac{x^2}{x+1}.$$

- *Any* line or curve crosses the x-axis when $y = 0$. Similarly, any line or curve crosses the y-axis when $x = 0$. So to find where $y = x^2 + x - 12$ crosses the x-axis we solve $0 = x^2 + x - 12$ and find $(3, 0)$ and $(-4, 0)$. To find where it crosses the y-axis we put in $x = 0$ to discover $(0, -12)$.

- Always, always, always draw a sketch in any problem that is even vaguely geometric; the sooner you do, the sooner you'll get full marks.

[1] In my own school a few candidates were disqualified recently because they had it with them in the exam despite the fact they were not using it. Don't bring it in at all! Also, don't bring in any paper with formulae on that they give with certain pencil cases.

Coordinates, Points and Lines

- Mid point of (x_1, y_1), (x_2, y_2) is $\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right)$. Average the x-coordinates and average the y-coordinates.
- Distance from (x_1, y_1) to (x_2, y_2) is (by Pythagoras) $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$. Be careful about negatives! Remember $(2 - (-3))^2 = (2 + 3)^2$.
- Gradient is defined to be
$$\frac{\text{difference in } y}{\text{difference in } x} = \frac{y_2 - y_1}{x_2 - x_1}.$$
If you need the gradient between two points you should visualise them first to see if you should be getting a positive or negative answer. This should also give you an idea of whether to expect a big (steep) of small (shallow) gradient.
- Two lines with gradients m_1 and m_2 are at right angles (perpendicular) if $m_1 \times m_2 = -1$. So if a line has gradient -3 then the line perpendicular to it has gradient $\frac{1}{3}$.
- Lines can be written in many forms, the most common being $y = mx + c$ and $ax + by = c$. Any form can be converted to any other. For example write $3x - 2y = 4$ in the form $y = mx + c$.
$$\begin{aligned} 3x - 2y &= 4 \\ 2y &= 3x - 4 \\ y &= \tfrac{3}{2}x - 2. \end{aligned}$$
- Given one point (x_1, y_1) and a gradient m the line is given by $y - y_1 = m(x - x_1)$.

Surds

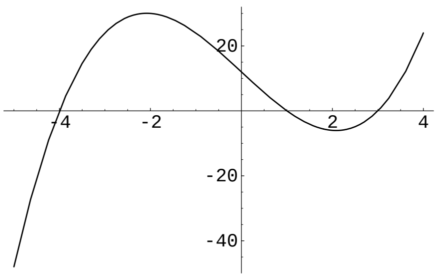
- Know and understand the laws
$$\sqrt{a \times b} = \sqrt{a} \times \sqrt{b} \qquad \text{and} \qquad \sqrt{\frac{a}{b}} = \frac{\sqrt{a}}{\sqrt{b}}.$$
In particular know how to deal with $\frac{\sqrt{44}}{2}$; it is *not* $\sqrt{22}$! It is $\frac{\sqrt{44}}{2} = \frac{\sqrt{44}}{\sqrt{4}} = \sqrt{\frac{44}{4}} = \sqrt{11}$. This comes up in solving quadratics by the formula; check that when you solve $x^2 + 4x - 2 = 0$ by the formula you obtain $x = -2 \pm \sqrt{6}$.
- You also need to be able to rationalise the denominator of certain types of surd expressions. For example to rationalise $\frac{9}{\sqrt{3}}$ is easy; just multiply by $\frac{\sqrt{3}}{\sqrt{3}}$ to obtain $\frac{9\sqrt{3}}{3} = 3\sqrt{3}$. In harder examples you must multiply the top and bottom of the fraction by the denominator with the sign 'flipped'. For example
$$\frac{2 + 2\sqrt{3}}{5 - 2\sqrt{3}} = \frac{2 + 2\sqrt{3}}{5 - 2\sqrt{3}} \times \boxed{\frac{5 + 2\sqrt{3}}{5 + 2\sqrt{3}}} = \frac{10 + 4\sqrt{3} + 10\sqrt{3} + 12}{25 + 10\sqrt{3} - 10\sqrt{3} - 12} = \frac{22 + 14\sqrt{3}}{13}.$$

Some Important Graphs

- Know the shape of the graph $y = x^n$ for $n = \{1, 2, 3, 4 \ldots\}$.
- If the power is even, then the graph will be $\bigcup$-shaped. They all pass through the points $(-1, 1)$, $(0, 0)$ and $(1, 1)$. The bigger the power, the faster it goes to infinity. Slightly more subtle is the point that in the range $-1 < x < 1$ then the higher the power, the *smaller* y-value (because $0.2 \times 0.2 \times 0.2 < 0.2 \times 0.2$). They are all 'even functions' with the y-axis as a line of symmetry.
- If the power is odd then they will (with the exception of $y = x^1 = x$, which is a straight line) be shaped like $y = x^3$. They all pass through $(-1, -1)$, $(0, 0)$ and $(1, 1)$. Similar arguments as for even powers exist here. They are all 'odd functions' with the origin being a point of rotational symmetry.
- The family of curves $y = ax^2 + bx + c$ are parabolas. If a is positive then you get a "happy" $\bigcup$-type curve. If a is negative then you get a "sad" $\bigcap$-type curve. They have a line of symmetry and a vertex (turning point) that you can discover by completing the square (see later).
- If you have a curve that is factorised then you can sketch it easily. For example

  $$y = (x - 1)(x + 4)(x - 3)$$

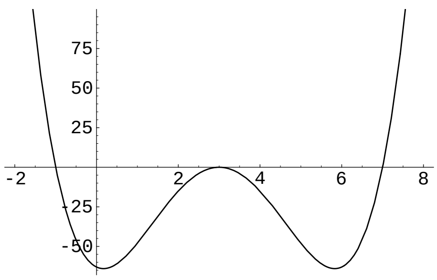
  is a cubic curve that crosses the x-axis at $(1, 0)$, $(-4, 0)$ and $(3, 0)$. It crosses the y-axis when $x = 0$, which gives $(0, 12)$. If x is huge, y is huge and positive and if x is massively negative, then so is y. So



- If a factor is repeated, then it merely touches the x-axis at that point. So

  $$y = (x - 3)^2(x + 1)(x - 7)$$

  is a quartic curve that crosses the x-axis at $(-1, 0)$ and $(7, 0)$, but only touches at $(3, 0)$.



Quadratics

- Factorising quadratics. To check whether a given quadratic factorises calculate the discriminant $b^2 - 4ac$; if it is a perfect square $(4, 49, 81$ etc.) then it factorises.
- When the x^2 coefficient (the number in front of the x^2) is one this is easy. Just spot two numbers which multiply to the constant and add to the x coefficient. For example with $x^2 + 8x + 15$ we need to find two numbers which multiply to 15 and sum to 8; clearly 3 and 5. So $x^2 + 8x + 15 = (x + 3)(x + 5)$.
- If the x^2 coefficient is not one then more work is required. You need to multiply the x^2 coefficient by the constant term and then find 2 numbers which multiply to this and sum to the x coefficient. For example with $6x^2 + x - 12$ we calculate $6 \times -12 = -72$ so the two numbers are clearly 9 and -8. So

$$\begin{aligned} 6x^2 + x - 12 &= 6x^2 + 9x - 8x - 12 & &= 6x^2 - 8x + 9x - 12 \\ &= 3x(2x + 3) - 4(2x + 3) & &= 2x(3x - 4) + 3(3x - 4) \\ &= (3x - 4)(2x + 3) & &= (2x + 3)(3x - 4). \end{aligned}$$

  Notice that it does not matter which way round we write the $9x$ and $-8x$.
- For quadratics that cannot be factorised we need to use the formula. For $ax^2 + bx + c = 0$ the solution is

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

- The $b^2 - 4ac$ part is called the *discriminant*. If it is positive then there are two *distinct* roots. If it is zero then there exists only one root and it is *repeated*. If it is negative then there are no roots. For example: find the values of k such that $x^2 + (k + 3)x + 4k = 0$ has only one root. We need the discriminant to be zero, so

$$\begin{aligned} b^2 - 4ac &= 0 \\ (k + 3)^2 - 16k &= 0 \\ k^2 - 10k + 9 &= 0 \\ k &= 9 \text{ or } k = 1. \end{aligned}$$

- Completing the square. All about halving the x coefficient into the bracket and then correcting the constant term. For example $x^2 - 6x + 10 = (x - 3)^2 - 9 + 10 = (x - 3)^2 + 1$. If the x^2 coefficient isn't one then need to factorise it out. For example

$$\begin{aligned} -2x^2 + 4x - 8 &= -2[x^2 - 2x] - 8 \\ &= -2[(x - 1)^2 - 1] - 8 \\ &= -2(x - 1)^2 - 6. \end{aligned}$$

  From this we can find the maximum or minimum of the quadratic. For $y = -2(x - 1)^2 - 6$ it is when $x = 1$ (to make the bracket 0) and therefore $y = -6$. In this case $(1, -6)$ is a maximum due to negative x^2 coefficient.

  We can also find the vertical line of symmetry by completing the square. For example

$$\begin{aligned} 3x^2 + 5x + 1 &= 3[x^2 + \tfrac{5}{3}x] + 1 \\ &= 3[(x + \tfrac{5}{6})^2 - \tfrac{25}{36}] + 1 \\ &= 3(x + \tfrac{5}{6})^2 - \tfrac{25}{12} + \tfrac{12}{12} \\ &= 3(x + \tfrac{5}{6})^2 - \tfrac{13}{12}. \end{aligned}$$

From this we see that the vertex is at $(-\frac{5}{6}, -\frac{13}{12})$ and consequently the line of symmetry is $x = -\frac{5}{6}$.

- You must be on the lookout for *quadratics in disguise*. You spot these when there are two powers on the variable and one is *twice* the other (or can be manipulated into such an equation[2]). Most students like to solve these by means of a substitution (although some students don't need to do this). For example to solve $x^4 + 2x^2 = 8$ work as follows:

$$\begin{aligned} x^4 + 2x^2 - 8 &= 0 \qquad \text{getting everything to one side} \\ u^2 + 2u - 8 &= 0 \qquad \text{substituting } u = x^2 \\ (u+4)(u-2) &= 0 \\ u = -4 \text{ or } u = 2 \qquad &\Rightarrow \qquad x^2 = -4 \text{ or } x^2 = 2 \end{aligned}$$

  But $x^2 = -4$ has no solutions, so $x = \pm\sqrt{2}$.

- For those who don't like substituting, just factorise and solve:

$$\begin{aligned} 2x^{\frac{2}{3}} &= 5x^{\frac{1}{3}} + 3 \\ 2x^{\frac{2}{3}} - 5x^{\frac{1}{3}} - 3 &= 0 \\ (2x^{\frac{1}{3}} + 1)(x^{\frac{1}{3}} - 3) &= 0 \end{aligned}$$

  So $x^{\frac{1}{3}} = -\frac{1}{2}$ or $x^{\frac{1}{3}} = 3$. Therefore cubing we find $x = -\frac{1}{8}$ or $x = 27$.

- Don't be one of the cretins who sees something like $x^4 + 4x^2 = 9$ and then *thinks* that they are square rooting to obtain $x^2 + 2x = 3$. Remember $\sqrt{x^4 + 4x^2} \neq x^2 + 2x$.

  Likewise $x + \sqrt{x} + 3 = 0$ does not square to $x^2 + x + 9 = 0$.

Differentiation

- We now turn to calculus[3]. Calculus $\equiv$ Differentiation + Integration. You will discover integration in C2.

- Differentiation allows us to calculate the 'gradient function' $\frac{dy}{dx}$. This tells us how the gradient on the original function y changes with x. $\frac{dy}{dx}$ *is* the gradient of a curve. So if you need to find where on a curve the gradient is 7, then you solve $\frac{dy}{dx} = 7$.

- Two alternative notations for derivatives are $\frac{dy}{dx} \equiv f'(x) \equiv y'$.

- The "rules" are;

$$\begin{aligned} y &= \text{constant} &&\Rightarrow \quad \frac{dy}{dx} = 0, \\ y &= ax &&\Rightarrow \quad \frac{dy}{dx} = a, \\ y &= ax^n &&\Rightarrow \quad \frac{dy}{dx} = anx^{n-1}. \end{aligned}$$

  Notice that the first two are merely subsets of the third; the third is the daddy; the big cheese; the head honcho...

[2] For example $3x^3 = 5 + \frac{2}{x^3}$ can be manipulated into $3x^6 - 5x^3 - 2 = 0$ where one power is twice the other.

[3] Isaac Newton. Arguably the greatest physicist ever. [Gottfried Leibniz also came up with it a bit later.]

- For example:

$$y = 4x^4 - 3x^2 + 2x - 5 \qquad \Rightarrow \qquad \frac{dy}{dx} = 16x^3 - 6x + 2,$$
$$y = 4x^{\frac{5}{4}} + 3x^{\frac{4}{5}} \qquad \Rightarrow \qquad \frac{dy}{dx} = 5x^{\frac{1}{4}} + \frac{12}{5}x^{-\frac{1}{5}}.$$

- You must expand brackets or carry out divisions *before* you differentiate[4]. For example:

$$y = x^2(x-3)^2 \qquad \Rightarrow \qquad y = x^4 - 6x^3 + 9x^2 \qquad \Rightarrow \qquad \frac{dy}{dx} = 4x^3 - 18x^2 + 18x,$$
$$y = \frac{x^7 + x}{x^6} \qquad \Rightarrow \qquad y = x + x^{-5} \qquad \Rightarrow \qquad \frac{dy}{dx} = 1 - 5x^{-6} = 1 - \frac{5}{x^6}.$$

- We can use differentiation to find the equation of tangents and normals to curves at specified points. For example find the equation of the normal to the curve $y = x^3 + 2x^2 - 5x - 1$ when $x = 1$.

  Firstly we need the y-coordinate: $x = 1 \quad \Rightarrow \quad y = -3$.

  Secondly $\frac{dy}{dx} = 3x^2 + 4x - 5$. Into this we put $x = 1$, so $\frac{dy}{dx} = 2$. Therefore the *normal* has gradient $-\frac{1}{2}$. So

$$y - y_1 = m(x - x_1)$$
$$y + 3 = -\tfrac{1}{2}(x - 1)$$
$$x + 2y + 5 = 0.$$

- If asked to show that $y + 5x + 17 = 0$ is tangent to the curve $y = x^2 + 3x - 1$, there are two methods to do this:

  1. Find where $y = x^2 + 3x - 1$ and $y + 5x + 17 = 0$ cross. Solving simultaneously we gain the quadratic $-5x - 17 = x^2 + 3x - 1$ which simplifies and factorises to $(x+4)(x+4) = 0$. This gives a *repeated* root, so the line intersects the curve once and we can therefore conclude that the line *must* be a tangent. [I prefer this method.]
  2. The line $y + 5x + 17 = 0$ has gradient -5. Therefore we need to find where on $y = x^2 + 3x - 1$ the gradient is -5. Therefore we differentiate $y = x^2 + 3x - 1$ to get $\frac{dy}{dx} = 2x + 3$ and put $\frac{dy}{dx} = -5$. This gives $x = -4$. On the curve, when $x = -4$, $y = 3$, so to find the equation of the tangent

$$y - y_1 = m(x - x_1)$$
$$y - 3 = -5(x + 4)$$
$$x + 5y + 17 = 0, \text{ as required.}$$

- Another example: Given that the curve $y = ax^3 + 4x^2 + bx + 1$ passes through the point $(-1, 5)$ and (at that point) the tangent is parallel to the line $y + 4x + 1 = 0$. Find a and b.

  There's quite a bit going on here, so take it a bit at a time. Since the curve passes through $(-1, 5)$, then $x = -1$ and $y = 5$ must be a solution to the curve's equation, so $5 = -a + 4 - b + 1$ which simplifies to $a + b = 0$. The line given has gradient -4, so we need to set $\frac{dy}{dx} = -4$ when $x = -1$. So $\frac{dy}{dx} = 3ax^2 + 8x + b$ which gives $-4 = 3a - 8 + b$. These solve to $a = 2$, $b = -2$.

[4] $y = \frac{x^3+x}{x^2}$ does *not* differentiate to $\frac{dy}{dx} = \frac{3x^2+1}{2x}$ and $y = (x^3 + 2)(2x^2 + 5x)$ does *not* differentiate to $\frac{dy}{dx} = 3x^2(4x + 5)$!

Inequalities

- Treat linear inequalities like equations except when multiplying or dividing by a negative number when you reverse the sign. For example

$$\begin{aligned} 2x + 4 &< 3x + 2 \\ -x &< -2 \\ x &> 2. \end{aligned}$$

- To solve quadratic inequalities:

  1. Get all terms over one side so that quadratic > 0 or quadratic < 0 in such a way that the x^2 term is always positive. This will ensure a 'happy' curve.
  2. Solve quadratic $= 0$ to find where it crosses x-axis.
  3. Sketch the graph and read off solution. If it is quadratic > 0 then it is the region(s) above the x-axis, and if quadratic < 0 then it is region below the x-axis.
  4. If one region then express as one triple inequality (e.g. $-2 < x < 5$) and if two regions then two *separate* inequalities (e.g. $x > 5$ or $x < -2$).

- For example solve the inequality $-7x \geqslant 4 - 2x^2$. Firstly get the $2x^2$ on the other side to make it positive to get $2x^2 - 7x - 4 \geqslant 0$. Then solve the equality $2x^2 - 7x - 4 = 0 = (2x+1)(x-4)$, so $x = -\frac{1}{2}$ or $x = 4$. So we have a happy quadratic that crosses the x-axis at $-\frac{1}{2}$ and 4. The inequality is asking for where the curve is bigger than (or equal to) zero, and this is to the right of $x = 4$ and the left of $x = -\frac{1}{2}$. Therefore the solution is $x \leqslant -\frac{1}{2}$ or $x \geqslant 4$.

- Don't fall into the trap of seeing $x^2 < 16$ and saying $x < \pm 4$! Be disciplined and get zero on one side; $x^2 - 16 < 0$ so $(x-4)(x+4) < 0$ so we have happy curve that crosses at 4 and -4. Where is the curve less than zero? Between -4 and 4 so solution is $-4 < x < 4$.

Index Notation

- $(ab)^m = a^m \times b^m$. For example $6^5 = 2^5 \times 3^5$.

- When multiplying a number raised to different powers the powers *add*. Therefore $a^m \times a^n = a^{m+n}$. You can think of this as follows $2^2 \times 2^4 = (2 \times 2) \times (2 \times 2 \times 2 \times 2) = 2^6$.

- Know that $a^{-m} = \frac{1}{a^m}$. Remember this by the standard result that $2^{-1} = \frac{1}{2}$. "When moving something from the bottom line of a fraction to the top (or vice versa), the sign changes."

- From the above two results we can obtain the result $\frac{a^m}{a^n} = a^{m-n}$. This is derived thus; $\frac{a^m}{a^n} = a^m \times a^{-n} = a^{m-n}$ as required.

- We can also derive the important result $a^0 = 1$ for any $a \neq 0$. Derived by considering something like this; $a^0 = a^{1-1} = \frac{a^1}{a^1} = \frac{a}{a} = 1$.

- Know that $(a^m)^n = a^{mn}$. Think about it like this; $(a^3)^4 = a^3 \times a^3 \times a^3 \times a^3 = a^{12}$.

- The n^{th} root of a number can be expressed as a power thus; $\sqrt[n]{a} = a^{\frac{1}{n}}$.

- A few examples:

  1. Write 8 as a power of 4; well $8 = 2^3 = \left(4^{\frac{1}{2}}\right)^3 = 4^{\frac{3}{2}}$.

2. Simplify $\sqrt[4]{16^3} = \left(\left(2^4\right)^3\right)^{\frac{1}{4}} = \left(2^{12}\right)^{\frac{1}{4}} = 2^3 = 8.$

3. Simplify $\dfrac{12x^8y^{\frac{3}{2}}}{6x^6y^{\frac{5}{2}}} = 2x^{8-6}y^{\frac{3}{2}-\frac{5}{2}} = 2x^2y^{-1}.$

4. Simplify $\sqrt{x^6y^4} \times \sqrt[3]{x^3y^{-6}} = \left(x^6y^4\right)^{\frac{1}{2}} \times \left(x^3y^{-6}\right)^{\frac{1}{3}} = x^3y^2x^1y^{-2} = x^4.$

- OCR is particularly 'hot' on linking differentiation, indices and surds. For example, find the equation of the normal to $y = 6x^{\frac{5}{2}} - 4x^{\frac{3}{2}}$ when $x = 2$ in the form $ax + by = c$.

  When $x = 2$, $y = 6 \times 2^{\frac{5}{2}} - 4 \times 2^{\frac{3}{2}} = 24\sqrt{2} - 8\sqrt{2} = 16\sqrt{2}$.

  Differentiating we find $\frac{dy}{dx} = 15x^{\frac{3}{2}} - 6x^{\frac{1}{2}}$.

  So, when $x = 2$, $\frac{dy}{dx} = 15 \times 2\sqrt{2} - 6\sqrt{2} = 24\sqrt{2}$. Therefore the gradient of the normal is $-\frac{1}{24\sqrt{2}} = -\frac{\sqrt{2}}{48}$. Therefore the equation of the normal is

$$\begin{aligned} y - y_1 &= m(x - x_1) \\ y - 16\sqrt{2} &= -\tfrac{\sqrt{2}}{48}(x - 2) \\ \sqrt{2}x + 48y &= 770\sqrt{2}. \end{aligned}$$

Graphs of nth Power Functions

- Be able to sketch $y = \frac{1}{x^n}$ for $n = \{1, 2, 3 \dots\}$.

  If n is even then graphs look like $y = \frac{1}{x^2}$ with the y-axis being a line of symmetry.

  If n is odd then the graph looks like $y = \frac{1}{x}$.

- Shock, horror! The differentiation rule for $y = ax^n$ still works with fractional and negative powers:

$$\frac{dy}{dx} = anx^{n-1}.$$

Polynomials

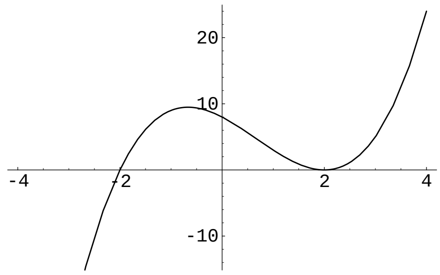
- Must be able to visualise a polynomial curve quickly. This is determined by two things;

  1. Whether the largest power of x is odd or even. x^{1000} has very different shape from x^{1001}.
  2. Whether the coefficient (number in front of) the largest power of x is positive or negative.

- A polynomial of order n has *at most* $n - 1$ stationary points. For example a cubic curve has up to two stationary points, but it might have none or one.

- If $y = x^3 + 2x^2 + 3x + 4$ then the curve crosses the y-axis at 4 (just the constant at the end). To find the x-axis intercept(s) need to solve $0 = x^3 + 2x^2 + 3x + 4$.

- To sketch a polynomial it is best to factorise it. For example given $y = x^3 - 2x^2 - 4x + 8$ we can write $y = (x + 2)(x - 2)^2$. So to sketch we know that it is a cubic with positive x^3 coefficient. y-axis intercept is at 8. It crosses the x-axis at $x = -2$, but only touches it at $x = 2$ due to the repeated root.



Transforming Graphs

- Given $y = f(x)$ then:

Replacement	Graph Shape
None	Normal Graph
x by $x - a$	Graph translated by a vector $\binom{a}{0}$
x by $-x$	Graph reflected in the y-axis
x by $\frac{x}{2}$	Graph stretched by a factor of 2 parallel to the x-axis $\longleftrightarrow$
x by $2x$	Graph stretched by a factor of $\frac{1}{2}$ parallel to the x-axis $\rightarrow\leftarrow$
y by $y - b$	Graph translated by a vector $\binom{0}{b}$
y by $-y$	Graph reflected in the x-axis
y by $\frac{y}{2}$	Graph stretched by a factor of 2 parallel to the y-axis $\updownarrow$
y by $2y$	Graph stretched by a factor of $\frac{1}{2}$ parallel to the y-axis $\downarrow\uparrow$

- For example: Find the equation of $y^2 + 2x^2 = 2x + 1$ after the translation $\binom{1}{-1}$. So we replace x by $x - 1$ and y by $y + 1$. Therefore

$$(y+1)^2 + 2(x-1)^2 = 2(x-1) + 1 \quad \Rightarrow \quad y^2 + 2y + 2x^2 = 6x - 4.$$

- For example: Explain the transformation that maps

$$y = \frac{2}{\sqrt{1+x}} \text{ onto } y = \frac{1}{\sqrt{3+x}}.$$

Rewriting the second equation as $2y = \frac{2}{\sqrt{1+(x+2)}}$ we can see y has been replaced by $2y$ and x has been replaced by $x + 2$. Therefore the curve has been translated by $\binom{-2}{0}$ and also stretched by a factor of $\frac{1}{2}$ parallel to the y-axis; i.e. every y-value has halved.

Investigating Shapes of Graphs

- Stationary points are where the gradient of curve is zero. They are either maxima, minima or points of inflection. To find the turning points of a curve we must find $\frac{dy}{dx}$ and then set $\frac{dy}{dx} = 0$ and solve for x.

- To determine the nature of a turning point we can consider the sign of the gradient either side of the turning point. Present this in a table. In the example of $y = x^2 + 2x + 3$ we find $\frac{dy}{dx} = 2x + 2$ so we solve $0 = 2x + 2$ to give the turning point when $x = -1$:

x	$x < -1$	-1	$x > -1$
dy/dx	negative	0	positive
		minimum	

- We can also use the second derivative to determine the nature of a turning point. This is found by differentiating the function twice;

$$y = 2x^3 + 3x^2 - 2x + 4 \quad \Rightarrow \quad \frac{dy}{dx} = 6x^2 + 6x - 2 \quad \Rightarrow \quad \frac{d^2y}{dx^2} = 12x + 6.$$

  You then evaluate the second derivative with the x value at the turning point and look at its sign. If it is positive it is a minimum, if it is negative it is a maximum. If it is zero then it is *probably* a point of inflection, but you need to do the above analysis either side of the turning point.

- For example, determine the nature of the stationary points on $y = 4x^3 - 21x^2 + 18x + 3$. So

$$y = 4x^3 - 21x^2 + 18x + 3 \quad \Rightarrow \quad \frac{dy}{dx} = 12x^2 - 42x + 18 = 0 \quad \Rightarrow \quad x = 3 \text{ or } x = \tfrac{1}{2}.$$

  Therefore the stationary points are $(3, -24)$ and $(\frac{1}{2}, \frac{29}{4})$. We therefore need the second derivative and evaluate it at 3 and $\frac{1}{2}$.

$$\begin{aligned} \frac{d^2y}{dx^2} &= 24x - 42 \\ \left.\frac{d^2y}{dx^2}\right|_{x=3} &= 24 \times 3 - 42 = 30 > 0 \text{ therefore } (3, -24) \text{ is a minimum.} \\ \left.\frac{d^2y}{dx^2}\right|_{x=\frac{1}{2}} &= 24 \times \tfrac{1}{2} - 42 = -30 < 0 \text{ therefore } (\tfrac{1}{2}, \tfrac{29}{4}) \text{ is a maximum.} \end{aligned}$$

- A function will be *increasing* when $\frac{dy}{dx}$ is positive and *decreasing* when $\frac{dy}{dx}$ is negative[5]. This is obvious if you consider a sketch.

  For example, find the set of values of x for which $y = -2x^3 + 3x^2 + 12x + 1$ is decreasing. First differentiate and recognise we want $\frac{dy}{dx}$ to be negative.

$$\begin{aligned} y &= -2x^3 + 3x^2 + 12x + 1 \\ \frac{dy}{dx} &= -6x^2 + 6x + 12 \\ 0 &> -6x^2 + 6x + 12 \qquad \text{(note that we place } \frac{dy}{dx} < 0\text{)} \\ x^2 - x - 2 &> 0 \\ (x-2)(x+1) &> 0. \end{aligned}$$

  This quadratic inequality solves to $x < -1$ or $x > 2$ which are the values of x for which the curve is decreasing.

Applications of Differentiation

- Differentiation can be used to work out "rates of change". In GCSE Physics you learnt that acceleration is the rate of change of velocity; you will then have learnt that $a = \frac{v-u}{t}$. However, at a higher level rates of change are calculated by differentiating with respect to time. So we now view acceleration as $a = \frac{dv}{dt}$.

[5] Technically a curve is 'increasing' when $\frac{dy}{dx} \geqslant 0$ and 'strictly increasing' when $\frac{dy}{dx} > 0$; likewise for decreasing, but don't get too het up about it; just get the sign the right way round.

- As we have already seen, differentiation allows us to calculate the stationary point of a curve $y = f(x)$. We do this by calculating $\frac{dy}{dx}$ and setting it equal to zero. We can use this to help is in practical problems where we might want to maximise a quantity (e.g. profit) or to minimise a quantity (e.g. cost).

- Worked example: An open topped cuboidal box is to be made from a rectangular piece of metal 10cm by 16cm. Squares are to be cut from each corner and then the four flaps are to be folded up. Find the maximum volume attainable for the box and prove that it is a maximum.

  1. Let x be the side length of the squares cut away, where $0 < x < 5$.
  2. The volume of the box would therefore be
  $$V = x(16-2x)(10-2x) = 160x - 52x^2 + 4x^3.$$
  We imagine a graph of V against x and hope that there is a stationary point in the range $0 < x < 5$.
  3. Differentiate V with respect to x and set equal to zero to find the stationary point;
  $$\frac{dV}{dx} = 160 - 104x + 12x^2 = 0 \quad \Rightarrow \quad x = 2 \text{ or } x = \tfrac{20}{3}.$$
  4. Notice that x can't be $\frac{20}{3}$ because it is outside the range $0 < x < 5$. So we only consider $x = 2$.
  5. If $x = 2$ then $V = 2 \times 12 \times 6 = 144$.
  6. To demonstrate that $x = 2$ is a maximum we need the second derivative and evaluate it at $x = 2$.
  $$\frac{d^2V}{dx^2} = -104 + 24x = -104 + 24 \times 2 = -56 < 0 \text{ therefore a maximum.}$$

Circles

- Circles with centre $(0,0)$ and radius r are expressed by $x^2 + y^2 = r^2$.

- Circles with centre (a,b) and radius r are expressed by $(x-a)^2 + (y-b)^2 = r^2$.

- By 'completing the square' you can convert circles of the form $x^2 + y^2 + \alpha x + \beta y + \gamma = 0$ into the form $(x-a)^2 + (y-b)^2 = r^2$. For example
$$\begin{aligned} x^2 + y^2 + 6x - 4y + 9 &= 0 \\ x^2 + 6x + y^2 - 4y + 9 &= 0 \\ (x+3)^2 - 9 + (y-2)^2 - 4 + 9 &= 0 \\ (x+3)^2 + (y-2)^2 &= 4. \end{aligned}$$

- When finding the intersection of a line and a circle it is easiest to substitute in the value of y from the line into the circle and solve the resulting quadratic. For example; find where the line $y = 2x - 1$ intersects to circle $(x-3)^2 + (y-2)^2 = 25$.
$$\begin{aligned} (x-3)^2 + (y-2)^2 &= 25 \\ (x-3)^2 + (2x-3)^2 &= 25 \\ x^2 - 6x + 9 + 4x^2 - 12x + 9 - 25 &= 0 \\ 5x^2 - 18x - 7 &= 0. \end{aligned}$$

Solve the quadratic (in this case by the formula) and then find the y values by substituting both x values into $y = 2x - 1$ (the original line). There will usually be 2 points of intersection (where the discriminant of the resulting quadratic will be positive) except if the line doesn't intersect the circle at all (discriminant negative) or if the line is a tangent to the circle (discriminant equals zero).

- The gradient of the tangent to a circle is perpendicular to the radius of the circle at that point. For example: The point $B(1, 7)$ lies on the circle $(x - 3)^2 + (y - 4)^2 = 13$. Find the equation of the tangent to the circle at B. The centre of the circle is $(3, 4)$, so the gradient of the radius at B is $-\frac{3}{2}$. Therefore the gradient of the tangent is $\frac{2}{3}$ and will pass through B, so the tangent will be:

$$\begin{aligned} y - y_1 &= m(x - x_1) \\ y - 7 &= \tfrac{2}{3}(x - 1) \\ 0 &= 2x - 3y + 19. \end{aligned}$$

- You must always remember the GCSE theorem that if a triangle is constructed within a circle with one side being a diameter of the circle, then it is a right angled triangle. To demonstrate this one is often required to show that the gradients of certain line segments are perpendicular (i.e. $m_1 \times m_2 = -1$).

  For example; $A(2, 1)$, $B(4, 13)$ and $C(-3, 8)$. The line segment AB is the diameter of a circle and C is a point on its circumference. Find the area of triangle ABC. We know angle $A\hat{C}B$ must be a right angle, so

$$\begin{aligned} \text{Area } ABC &= \tfrac{1}{2} \times \text{base} \times \text{height} \\ &= \tfrac{1}{2} \times (\text{length } AC) \times (\text{length } CB) \\ &= \tfrac{1}{2} \times \sqrt{5^2 + 7^2} \times \sqrt{7^2 + 5^2} \\ &= 37 \text{ units}^2. \end{aligned}$$

- Also, given two points that lie on a circle's circumference, the centre of the circle lies on the perpendicular bisector of the two points.

OCR Core 2 Module Revision Sheet

The C2 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J.M.S.

Trigonometry

- We define
$$\tan\theta \equiv \frac{\sin\theta}{\cos\theta}.$$
This identity is very useful in solving equations like $\sin\theta - 2\cos\theta = 0$ which yields $\tan\theta = 2$. The solutions of this in the range $0° \leqslant \theta \leqslant 360°$ are $\underline{\theta = 63.4°}$ and $\underline{\theta = 243.4°}$ to one decimal place.

- Know the following (or better yet, learn a couple and be able to derive the rest, quickly, from your knowledge of the trigonometric functions):

θ	$\sin\theta$	$\cos\theta$	$\tan\theta$
$0°$	0	1	0
$30°$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{3}}$
$45°$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	1
$60°$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$
$90°$	1	0	undefined
$180°$	0	-1	0

- Be able to sketch $\sin\theta$, $\cos\theta$ and $\tan\theta$ in both degrees and radians.

- By considering a right angled triangle (or a point on the unit circle) we can derive the important result $\sin^2\theta + \cos^2\theta \equiv 1$. This is useful in solving certain trigonometric equations. Worked example; solve $1 = 2\cos^2\theta + \sin\theta$ for $0° \leqslant \theta \leqslant 360°$.
$$\begin{aligned}
1 &= 2\cos^2\theta + \sin\theta \\
1 &= 2(1-\sin^2\theta) + \sin\theta && \text{get rid of } \cos^2\theta, \\
0 &= 1 - 2\sin^2\theta + \sin\theta && \text{quadratic in } \sin\theta, \\
0 &= 2\sin^2\theta - \sin\theta - 1 && \text{factorise as normal,} \\
0 &= (2\sin\theta + 1)(\sin\theta - 1).
\end{aligned}$$
So we just solve $\sin\theta = -\frac{1}{2}$ and $\sin\theta = 1$. Therefore $\underline{\theta = 210°}$ or $\underline{\theta = 330°}$ or $\underline{\theta = 90°}$.

- The above relation is also useful in converting between the different trigonometric functions. For example if $\cos\theta = \frac{6}{7}$ then, to find $\sin\theta$, do **not** use "$\cos^{-1}$" on your calculator and then "sin" the answer. Instead

$$\sin^2\theta + \cos^2\theta = 1,$$
$$\sin^2\theta + \frac{36}{49} = 1,$$
$$\sin\theta = \pm\sqrt{\frac{13}{49}} = \pm\frac{\sqrt{13}}{7}.$$

Without further information you must keep both the positive and negative solution.

- If a question tells you that the angle is 'acute', 'obtuse' or 'reflex' then you must visualise the appropriate graph and interpret. For example given that $\sin\theta = \frac{1}{3}$ and that θ is obtuse, find the value of $\cos\theta$. By the argument above you will find that

$$\cos\theta = \pm\frac{\sqrt{8}}{3} = \pm\frac{2\sqrt{2}}{3}.$$

  However, given an obtuse angle ($90° < \theta < 180°$) the cosine graph is negative, so the final answer should be $\cos\theta = -\frac{2\sqrt{2}}{3}$.

- You must be careful when you see things like $2\tan x \sin x = \tan x$. It is SO tempting to divide both sides by $\tan x$ to yield $2\sin x = 1$. But you must bring everything to one side and factorise;

$$2\tan x \sin x - \tan x = 0 \qquad \Rightarrow \qquad \tan x(2\sin x - 1) = 0.$$

  The full set of solutions can then be by solving $\tan x = 0$ <u>and</u> $2\sin x - 1 = 0$. [It is completely analogous to $x^2 = x$. If we divide by x we find $x = 1$, but we know this has missed the solution $x = 0$. However when we factorise we find $x(x-1) = 0$ and both solutions are found.]

- Given a trigonometric equation it is always best first to isolate the trigonometric function on its own; for example

$$9\cos(\dots) + 2 = 7 \qquad \Rightarrow \qquad \cos(\dots) = \tfrac{5}{9}.$$

- For complicated trigonometric equations where you are not just 'cos'ing, 'sin'ing or 'tan'ing a single variable (x, θ, t or the like), it is often easiest to make a substitution.

  For example to solve $\cos(2x+30) = \frac{1}{4}$ in the range $0° \leqslant x \leqslant 360°$ the desired substitution is clearly $u = 2x + 30$, but you **must** remember to also convert the range also (many students forget this) so:

$$\cos(2x+30) = \tfrac{1}{4} \qquad 0° \leqslant x \leqslant 360°,$$
$$\cos u = \tfrac{1}{4} \qquad 30° \leqslant u \leqslant 750°,$$
$$u = \dots°, \dots°, \dots°, \dots°.$$

  However, we don't want solutions in u, so we need to use $x = \frac{u-30}{2}$ on each u solution to get

$$x = \dots°, \dots°, \dots°, \dots°.$$

Sequences

- You must be comfortable with $\sum$-notation. It works as follows; you put in the number at the bottom of the $\sum$ and then keep summing until you reach the top number[6]. For example:

$$\sum_{i=4}^{8}(2i+3) = 11+13+15+17+19 = 75,$$
$$\sum_{i=1}^{n}(i^2+i) = (1+1)+(4+2)+(9+3)+\cdots+(n^2+n).$$

- A 'sequence' is a list of numbers in a specific order. A 'series' is a sum of the terms of a sequence.

- Sequences are sometimes defined *recursively*. For example the sequence $a_{n+1} = a_n + 3$ with $a_1 = 10$ defines the sequence $10, 13, 16, 19\ldots$ We know that this is an arithmetic sequence which can also be defined *deductively* by $a_n = 10 + 3(n-1)$.

- An arithmetic sequence increases or decreases by a constant amount. The letter a always denotes the first term and d is the difference between the terms (negative for a decreasing sequence!). The nth term is denoted a_n and satisfies the important relationship

$$a_n = a + (n-1)d.$$

For example if told the third term of a sequence is 10 and the seventh term is 34 then we can use the above equation to find the a and d.

$$\begin{aligned} 10 &= a + (3-1)d \\ 34 &= a + (7-1)d \end{aligned} \quad \Rightarrow \quad 4d = 24 \quad \Rightarrow \quad d = 6 \quad \Rightarrow \quad a = -2.$$

- The sum of the n terms of an arithmetic sequence is given by

$$S = \frac{n}{2}(\text{First} + \text{Last}) = \frac{n}{2}(2a + (n-1)d).$$

For example the sum of the first 10 terms of a sequence is 130 and the first term is 4. What is the difference?

$$S = \frac{n}{2}(2a+(n-1)d) \quad \Rightarrow \quad 130 = \frac{10}{2}(8 + (10-1)d) \quad \Rightarrow \quad d = 2.$$

Binomial Theorem

- Binomial expansion allows us to expand $(a+b)^n$ for any integer n. Best explained by means of an example; expand $(2x-y)^5$.

1. Begin by considering 'prototype' expansion of $(a+b)^5$.
2. So $(a+b)^5 = \binom{5}{0}a^5 + \binom{5}{1}a^4b + \binom{5}{2}a^3b^2 + \binom{5}{3}a^2b^3 + \binom{5}{4}ab^4 + \binom{5}{5}b^5$.
3. Calculate binomial coefficients either on calculator or by drawing a mini Pascal's Triangle to give $(a+b)^5 = 1a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + 1b^5$.

---

[6] If you want to multiply instead then use $\prod$-notation. For example $\prod_{i=1}^{n} i = 1 \times 2 \times 3 \times 4 \times \cdots \times n \equiv n!$

4. Next notice that in our case $a = 2x$ and $b = -y$ and substitute in to get $(2x - y)^5 = 1(2x)^5 + 5(2x)^4(-y) + 10(2x)^3(-y)^2 + 10(2x)^2(-y)^3 + 5(2x)(-y)^4 + 1(-y)^5$.
5. Tidying up we get $(2x - y)^5 = 32x^5 - 80x^4y + 80x^3y^2 - 40x^2y^3 + 10xy^4 - y^5$.

- It is worth noting that when the expansion is of the form (something − another thing)^n, then the signs will alternate.

- Also of note is the way each *individual* component is constructed. For example; find the x^5 coefficient in the expansion of $(2 - 3x)^7$. The component with x^5 is given by $\binom{7}{5}(2)^2(-3x)^5 = -20412x^5$, so the coefficient is -20412.

- $\displaystyle\binom{n}{r} = {}^nC_r = \frac{n!}{r!(n-r)!}$. For example $\displaystyle\binom{5}{2} = \frac{5!}{2!\,3!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{(2 \times 1) \times (3 \times 2 \times 1)} = 10$.

- Know that $(1 + x)^n$ expands thus:

$$\begin{aligned}(1+x)^n &= 1 + \binom{n}{1}x + \binom{n}{2}x^2 + \binom{n}{3}x^3 + \cdots + x^n \\ &= 1 + nx + \frac{n(n-1)}{2}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \cdots + x^n.\end{aligned}$$

  (This is particularly useful in Core 4.)

Sine & Cosine Rules

- The sine rule states for any triangle $\dfrac{\sin A}{a} = \dfrac{\sin B}{b} = \dfrac{\sin C}{c}$.

- The cosine rule states that $a^2 = b^2 + c^2 - 2bc\cos A$. Practice both sine and cosine rules on page 293.

- By considering half of a general parallelogram we can show that the area of any triangle is given by $A = \frac{1}{2}ab\sin C$.

- You must be good at bearing problems which result in triangles. Remember to draw lots of North lines and remember also that they are all parallel; therefore you can use Corresponding, Alternate and Allied angle theorems... revise your GCSE notes! Bearings are measured clockwise from North and must contain three digits. For example

$$12.2^\circ \quad \Rightarrow \quad 012.2^\circ.$$

Integration

- *Calculus* is the combined study of differentiation *and* integration (and their relationship). A good description is that calculus is the study of change in the same way that geometry is the study of shapes.

- Integration is the reverse of differentiation. That is if $\frac{dy}{dx} = f(x)$ then $y = \int f(x)\,dx$. For example if $\frac{dy}{dx} = 3x^3$ then $y = \int 3x^3\,dx = \frac{3}{4}x^4 + c$.

- The general rule is therefore $\displaystyle\int ax^n\,dx = \frac{ax^{n+1}}{n+1} + c$.

- $\int y\,dx$ is an *indefinite* integral because there are no limits on the integral sign. When evaluating these integrals *never* forget an *arbitrary constant* added on at the end. For example $\int 6x^2\,dx = 2x^3 + c$.

- $\int_a^b y\,dx$ is a *definite integral* and is the area between the curve and the x-axis from $x = a$ to $x = b$. Areas under the x-axis are negative. (For areas between the curve and the y-axis switch the x and the y and use $\int_p^q x\,dy$ between $y = p$ and $y = q$.)

- To find the area *between* two curves between $x = a$ and $x = b$ evaluate

$$\int_a^b (\text{top} - \text{bottom})\,dx.$$

- For example, given that $y = \sqrt{x} + \frac{1}{\sqrt{x}}$, find the area under the curve from $x = 1$ and $x = 2$.

$$\begin{aligned}\int_1^2 \sqrt{x} + \frac{1}{\sqrt{x}}\,dx &= \int_1^2 x^{\frac{1}{2}} + x^{-\frac{1}{2}}\,dx\\ &= \left[\frac{2}{3}x^{\frac{3}{2}} + 2x^{\frac{1}{2}}\right]_1^2\\ &= \left(\frac{2}{3} \times 2^{\frac{3}{2}} + 2 \times 2^{\frac{1}{2}}\right) - \left(\frac{2}{3} + 2\right)\\ &= \frac{10\sqrt{2}}{3} - \frac{8}{3}.\end{aligned}$$

- To calculate integrals where one of the limits is infinite (∞ or $-\infty$), proceed as normal until you input the ∞ or $-\infty$ into the integral. Then you must **not** write such things as

$$\frac{1}{\infty}, \quad \frac{1}{3^\infty}, \quad \frac{2\infty + 1}{3\infty - 2}, \quad 2^{\frac{1}{\infty}}, \quad \text{and the like.}$$

You must think what these things would equal and just write down the number; in the previous four cases you would get 0, 0, $\frac{2}{3}$ and 1. (If, in the C2 exam, you think that when you put in ∞ you get an infinite answer, chances are you've made a mistake somewhere.) For example:

$$\begin{aligned}\int_2^\infty \frac{8}{5x^3}dx &= \int_2^\infty \frac{8}{5}x^{-3}dx\\ &= \left[-\frac{4}{5}x^{-2}\right]_2^\infty\\ &= (0) - \left(-\frac{4}{5} \times \frac{1}{2^2}\right) = \frac{1}{5}.\end{aligned}$$

Geometric Sequences

- A geometric sequence is one where the terms are multiplied by a constant amount. For example $1, 2, 4, 8, 16, \ldots, [2^{n-1}]$ is a geometric sequence with $a = 1$ and $r = 2$. The n^{th} term is given by

$$a_n = ar^{n-1}.$$

So for the above example the 20th term is $a_{20} = 1 \times 2^{19} = 524288$.

- The sum of n terms of a geometric sequence is given by

$$S = a\left(\frac{r^n - 1}{r - 1}\right) \qquad \text{or (equivalently) by} \qquad S = a\left(\frac{1 - r^n}{1 - r}\right).$$

For example sum the first 20 terms of $4, 2, 1, \frac{1}{2}, \ldots, [4 \times \left(\frac{1}{2}\right)^{n-1}]$. This is given by

$$S = 4\left(\frac{\left(\frac{1}{2}\right)^{20} - 1}{\frac{1}{2} - 1}\right) = 7.999992371\ldots$$

- If the ratio (r) lies between -1 and 1 (i.e. $-1 < r < 1$) then there exists a 'sum to infinity' given by
$$S_\infty = \frac{a}{1-r}.$$
Therefore S_∞ for the above example is $S_\infty = \frac{4}{1-\frac{1}{2}} = 8$. We can see that the sum to 20 terms is very close to S_∞.

Exponentials & Logarithms

- (In these notes if I write $\log x$ I *mean* $\log_{10} x$. If I mean a different base, I will write it *explicitly* as $\log_a x$. When we see $\log_a x$ we say "log to the base a of x".)

- We *define* a logarithm to be the solution (in x) to the equation $a^x = b$. It is written $x = \log_a b$. The fundamental relationship is therefore
$$a^x = b \qquad \Leftrightarrow \qquad x = \log_a b. \qquad\qquad \dagger$$

- From † we see that $\log_a b$ *means*: "The number a has to be raised to, to make b". Therefore some simple logarithms can be calculated without a calculator:
$$\begin{aligned} \log_2 8 &= \text{"the number 2 has to be raised to, to make 8"} &&= 3, \\ \log_{10} 10000 &= \text{"the number 10 has to be raised to, to make 10000"} &&= 4, \\ \log_9 3 &= \text{"the number 9 has to be raised to, to make 3"} &&= \tfrac{1}{2}, \quad (\because 3 = \sqrt{9} = 9^{\frac{1}{2}}) \\ \log_a a &= \text{"the number } a \text{ has to be raised to, to make } a\text{"} &&= 1. \end{aligned}$$

- We see from † that logarithms 'pluck out powers' from equations. Therefore if you ever see an equation with the unknown in the power, then that is the clue that you will need to use logarithms. For example to solve $7^{2x-1} = 22$ we discover
$$\begin{aligned} 7^{2x-1} &= 22, \\ 2x - 1 &= \log_7 22, \\ x &= \tfrac{1}{2}\log_7 22 + \tfrac{1}{2}. \end{aligned}$$
However we need to build on † because not all equations are this simple (e.g. $3 \times 2^{2x-1} = 5 \times 7^{x+1}$) and not all calculators can calculate $\log_7 22$.

- You can also use † to eliminate logarithms from an equation. Given an equation of the form $\log_a(\cdots) = b$, you can eliminate the logarithm instantly to get $(\cdots) = a^b$. A good way to remember this[7] is 'Girvan's Bullying Base'. So if we have $\log_3 x = 8$, then the bullying base '3' knocks the log out of the way and moves to the other side and squeezes up the 8 to put it in its place; therefore $x = 3^8$.

- Logarithms and exponentials (powers) are the inverse functions of each other (as can be seen from † if one puts one into the other). Therefore
$$\log_{10} 10^x = x \quad \text{and} \quad 10^{\log_{10} x} = x.$$
So if $\log a = 5.4$ then $a = 10^{5.4}$.

[7] From a colleague I respect; Mr Girvan

- There are some rules that can be derived from † that *must* be learnt. They are (for all bases):

$$\begin{aligned} \log(ab) &= \log a + \log b & \log 1 &= 0 \\ \log(\frac{a}{b}) &= \log a - \log b & \log_a a &= 1 \\ \log(a^n) &= n \log a & \log_a b &= \frac{\log_c b}{\log_c a}. \\ \log(\frac{1}{a}) &= -\log a & & \end{aligned}$$

- When we need to solve an equation where the unknown is in the exponent such as $5^{2x-1} = 8$ take $\log_{10}$ of both sides and simplify:

$$\begin{aligned} 5^{2x-1} &= 8 \\ \log_{10}(5^{2x-1}) &= \log_{10} 8 \\ (2x-1)\log_{10} 5 &= \log_{10} 8 \\ 2x - 1 &= \frac{\log_{10} 8}{\log_{10} 5} \\ x &= \frac{1}{2} \times \left(\frac{\log_{10} 8}{\log_{10} 5} + 1\right) \\ x &= 1.15 \text{ (3sf)}. \end{aligned}$$

Factors and Remainders

- Need to know how to divide any polynomial by a linear factor of the form $ax - b$. For example divide $x^3 + 2x^2 + 3x - 6$ by $x - 2$. (*Always* devote a column to each power of x.)

$$\begin{array}{r|rrrr} & & x^2 & +4x & +11 \\ \hline x-2 & +x^3 & +2x^2 & +3x & -6 \\ & +x^3 & -2x^2 & & \\ \hline & & +4x^2 & +3x & \\ & & +4x^2 & -8x & \\ \hline & & & +11x & -6 \\ & & & +11x & -22 \\ \hline & & & & +16 \end{array}$$

So the remainder is $+16$. Therefore $x^3 + 2x^2 + 3x - 6 = (x-2)(x^2 + 4x + 11) + 16$.

- If the remainder is zero, then the divisor is said to be a *factor* of the original polynomial.

- The Factor Theorem states:

$$(x - a) \text{ is a factor of } f(x) \quad \Leftrightarrow \quad f(a) = 0.$$

[More generally (but used less often in exams) is:

$$(ax - b) \text{ is a factor of } f(x) \quad \Leftrightarrow \quad f(\tfrac{b}{a}) = 0.]$$

- For example if $x^3 + ax^2 + 8x - 4$ has $(x - 2)$ as a factor, find a. From factor theorem we know $f(2) = 0$, so we discover $2^3 + a \times 2^2 + 8 \times 2 - 4 = 0$, and therefore $a = -5$.

- The Remainder Theorem states:

When $f(x)$ is divided by $(x-a)$ the remainder is $f(a)$.

[More generally (but used less often in exams) is:

When $f(x)$ is divided by $(ax-b)$ the remainder is $f(\frac{b}{a})$.]

Notice that the factor theorem is a subset of the remainder theorem. In the factor theorem all remainders are zero, by definition.

- For example if told that when $f(x) = x^3 + 2x^2 - 3x - 7$ is divided by $x - 2$ the remainder is 3, we know $f(2) = 3$.

- Worked example: $f(x) = 2x^3 + 3x^2 + kx - 2$. The remainder when $f(x)$ is divided by $(x-2)$ is four times the remainder when $f(x)$ is divided by $(x+1)$. Find k. We know

$$\begin{aligned} f(2) &= 4 \times f(-1) \\ 2 \times 2^3 + 3 \times 2^2 + 2k - 2 &= 4[2 \times (-1)^3 + 3 \times (-1)^2 - k - 2] \\ k &= -5. \end{aligned}$$

Radians

- There are (by definition) 2π radians in a circle. So $360° = 2\pi$. To convert from degrees to radians we use the conversion factor of $\frac{\pi}{180}$. For example to convert 45° to radians we calculate $45 \times \frac{\pi}{180} = \frac{\pi}{4}$ rad. From radians to degrees we use its reciprocal $\frac{180}{\pi}$.

- *When using radians* the formulae for arc length and area of a sector of a circle become simpler. They are $s = r\theta$ and $A = \frac{1}{2}r^2\theta$.

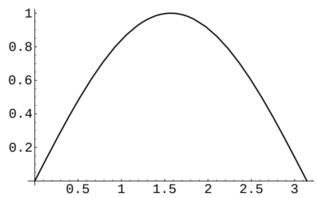
The Trapezium Rule

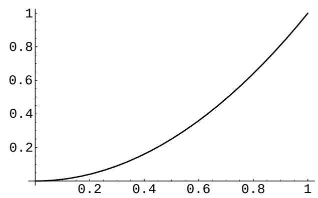
- The area under *any* curve can be *approximated* by the Trapezium Rule. The governing formula is given by (and contained in the formula booklet you will have in the exam)

$$\int_a^b y\,dx \approx \frac{1}{2}h\left[y_0 + y_n + 2(y_1 + y_2 + \cdots + y_{n-1})\right],$$

where h is the width of each trapezium, y_0 and y_n are the 'end' heights and $y_1 + y_2 + \cdots + y_{n-1}$ are the 'internal' heights.

- By considering the shape of the graph in the interval over which you are approximating it should be clear whether your estimate of the area is an over or under-estimate of the *true* area.





For example if you were to estimate $\int_0^\pi \sin x\,dx$ (above, left) using the trapezium rule, due to the shape of the curve, the trapezia would all fall below the curve, so we would obtain an *under*-estimate. However, with $\int_0^1 x^2\,dx$ (above, right) we would obtain an *over*-estimate.

OCR Core 3 Module Revision Sheet

The C3 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J.M.S.

Functions

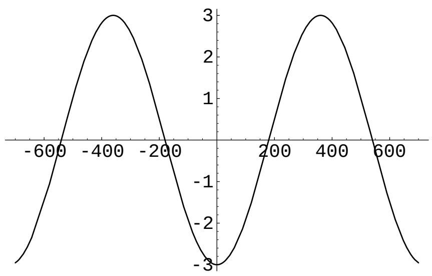
- A *function* is a *one-to-one* or a *many-to-one mapping*. There are also *many-to-many* and *one-to-many* mappings, but these are ***not*** functions. In a function, for every value you feed into the function you obtain one (and only one) value out.
- The *domain* of a function $y = f(x)$ is all the possible values of x the function can take. For example the domain of $y = \sqrt{x-4}$ is $x \geqslant 4$. In other words all the *inputs* the function can take.
- The *range* of a function is all the possible *outputs*. That is all the possible values of $f(x)$. So for $f(x) = -x^2 + 5$ the range is $f(x) \leqslant 5$.
- Functions are transformed as follows

Function	Graph Shape
$f(x)$	Normal Graph
$2f(x)$	Graph stretched by a factor of 2 parallel to the y-axis i.e. every value of $f(x)$ in the original graph is multiplied by 2
$f(2x)$	Graph stretched by factor of $\frac{1}{2}$ parallel to the x-axis
$3f(\frac{x}{4})$	Graph stretched by factor of 4 parallel to the x-axis and a stretch by a factor of 3 parallel to the y-axis
$f(x)+6$	Graph translated vertically *up* 6 units
$f(x)-6$	Graph translated vertically *down* 6 units
$f(x+4)$	Graph translated 4 units to the *left*
$f(x-6)$	Graph translated 6 units to the *right*
$f(x-6)+9$	Graph translated 6 units to the *right* and 9 units *up*. This is a translation and can be expressed as $\binom{6}{9}$ where $\binom{\text{change in } x}{\text{change in } y}$
$-f(x)$	Graph reflected in the x-axis
$f(-x)$	Graph reflected in the y-axis

- When faced with more than one of the above transformations it sometimes matters which order you carry out the transformations. In the example of $2f(x-3)$ it doesn't matter because you end up with the same result both ways, regardless of whether you do the translation right, or the stretch parallel to the y-axis first (think about it). However with $f(2x+10)$ you get a different result depending on the order you carry out the translation 10 left and then stretch factor $\frac{1}{2}$ parallel to the x-axis. If the conflict occurs within the bracket you should do the *opposite* of what you expect. So here you do the translation first and then the stretch.

  For $2f(x)+6$ the transformations are outside the bracket, so here you would do the stretch *then* the translation.

- So for example if you were asked to sketch $y = 3\sin(\frac{x}{2} - 90)$ you would translate '$y = \sin x$' $90°$ to the right, *then* stretch factor 2 parallel to the x-axis and stretch factor 3 parallel to the y-axis.



- If $f(x) = f(-x)$ then the function is called an *even* function. An even function is one where the y-axis is a line of symmetry. Examples are

$$f(x) = \cos x \qquad \text{since} \qquad f(-x) = \cos(-x) = \cos x = f(x),$$
$$g(x) = x^2 + 1 \qquad \text{since} \qquad g(-x) = (-x)^2 + 1 = x^2 + 1 = g(x).$$

- If $-f(x) = f(-x)$ then the function is called an *odd* function. An odd function is one where the function is unchanged if you rotate it $180°$ around the point $(0, 0)$. Examples are

$$f(x) = \sin x \qquad \text{since} \qquad f(-x) = \sin(-x) = -\sin x = -f(x),$$
$$g(x) = x^3 \qquad \text{since} \qquad g(-x) = (-x)^3 = -x^3 = -g(x).$$

- You must be able to construct compositions of functions. Note that $f(g(x))$ is not usually the same as $g(f(x))$. For example if $f(x) = x^2$ and $g(x) = x+1$ then $f(g(x)) = f(x+1) = (x+1)^2 = x^2 + 2x + 1$. Contrast this with $g(f(x)) = g(x^2) = x^2 + 1$.

- Sometimes you will be asked to describe a quadratic of the form $ax^2 + bx + c$ in terms of $f(x) = x^2$. It is often useful to *complete the square*. Very quickly I will go through a couple of examples of how to do this:

$x^2 + 10$	$\Rightarrow$	Clearly just $f(x) + 10$.
$x^2 + 6x + 10$	$\Rightarrow$	Complete square to get $(x+3)^2 - 9 + 10 = (x+3)^2 + 1$ so it is $f(x+3) + 1$, which is the translation $\binom{-3}{1}$ of x^2.
$2x^2 + 16x + 1$	$\Rightarrow$	Complete square to get $2(x+4)^2 - 31$ so it is $2f(x+4) - 31$, which is a stretch of factor 2 away from the x-axis, followed by a translation $\binom{-4}{-31}$ of x^2.

- The inverse of a function $f(x)$ is denoted $f^{-1}(x)$. To find the inverse of a function you swap round the x and the y and make y the subject again. This will be the inverse of the original function. For example find the inverse of $f(x) = \sqrt{x^3 + 2}$ gives

$$\begin{aligned} f(x) &= \sqrt{x^3 + 2}, \\ \Rightarrow \qquad y &= \sqrt{x^3 + 2}, \\ \Rightarrow \qquad x &= \sqrt{y^3 + 2}, \\ \Rightarrow \qquad y &= \sqrt[3]{x^2 - 2}, \\ \Rightarrow \qquad f^{-1}(x) &= \sqrt[3]{x^2 - 2}. \end{aligned}$$

- A function only has an inverse if it is a one-to-one mapping. If the original function is a many-to-one function (e.g. $y = x^2$ or any of the trig functions) you must restrict its domain to make it a one-to-one mapping (e.g. for $y = x^2$ restrict domain to $x \geqslant 0$). The domain and range of a function are switched in its inverse. For example if $f(x)$ has domain $x > 8$ and range $f(x) \leqslant -10$, then its inverse $f^{-1}(x)$ has domain $x \leqslant -10$ and range $f^{-1}(x) > 8$.

- Geometrically the relationship between a function and its inverse is a reflection in the line $y = x$. A useful spin-off from this result is that if you are asked to find where a function equals its inverse (i.e. $f(x) = f^{-1}(x)$) all you need to do is solve $f(x) = x$ or $f^{-1}(x) = x$; take your pick.

- Given a point on a function ($(3,4)$, say) then the equivalent point on its inverse is $(4,3)$ because it has been reflected in $y = x$. If the gradient at $(3,4)$ was 7, then the gradient on the inverse will be its reciprocal $\frac{1}{7}$.

Modulus

- The modulus function makes everything you put into it positive. For example $|4| = 4$ and $|-6| = 6$. If something negative is 'fed in' to the mod function then it multiplies it by -1 to turn it positive; otherwise it leaves it alone.

- If you have an expression such as $|x-4|$, then the critical value for x is $x = 4$; if $x > 4$ then the expression is just $x - 4$ and if $x < 4$ then the expression becomes $-x + 4$ because the mod function multiplies it by -1 to turn it positive. This idea helps us solve modulus equations; for example to solve $|2x - 1| = 6$ we first look for the critical values of x; here clearly $x = \frac{1}{2}$. We therefore set up two equations depending on whether $x > \frac{1}{2}$ or $x < \frac{1}{2}$:

$$\begin{aligned} \text{If } x &< \tfrac{1}{2} \\ \text{then } -2x + 1 &= 6 \\ x &= -\tfrac{5}{2}, \end{aligned} \qquad\qquad \begin{aligned} \text{If } x &> \tfrac{1}{2} \\ \text{then } 2x - 1 &= 6 \\ x &= \tfrac{7}{2}. \end{aligned}$$

  We perform a little check at the end to check that the solutions found actually satisfy the conditions on x are met; the left hand equation is valid if $x < \frac{1}{2}$ and the solution we have found *is* less that $\frac{1}{2}$; the right hand equation is valid if $x > \frac{1}{2}$ and the solution *is* greater than $\frac{1}{2}$. Therefore both solutions found are valid.

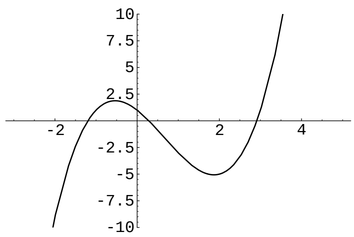
- If you merely have an equation such as |something| = |something else| then just get rid of the mods and square both sides to get (something)^2 = (something else)^2. *Check* your answers back in the original mod equation to check they work!

- Consider the intimidating looking $|2x-1| - 1 < |x+2|$. As with most inequalities a good first step is to solve the *equality*; i.e. solve $|2x - 1| - 1 = |x + 2|$. The critical x values are $x = -2$ and $x = \frac{1}{2}$ so we need to set up three different equations depending whether x is $x < -2$, $-2 < x < \frac{1}{2}$ or $x > \frac{1}{2}$ and solve:

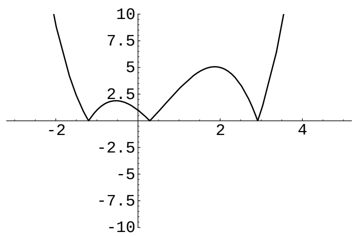
$$\begin{aligned} \text{If } x &< -2 \\ \text{then } -2x + 1 - 1 &= -x - 2 \\ x &= 2, \end{aligned} \qquad \begin{aligned} \text{If } -2 < x &< \tfrac{1}{2} \\ \text{then } -2x + 1 - 1 &= x + 2 \\ x &= -\tfrac{2}{3}, \end{aligned} \qquad \begin{aligned} \text{If } x &> \tfrac{1}{2} \\ \text{then } 2x - 1 - 1 &= x + 2 \\ x &= 4. \end{aligned}$$

  Performing our check again we see that two solutions are fine, but $x = 2$ is *not* a solution because the equation was only valid if $x < -2$. Therefore the solution of the equation is $x = -\frac{2}{3}$ or $x = 4$. To solve the inequality we need to see if a number less than $-\frac{2}{3}$ works in the inequality (it doesn't), to see if a number between $-\frac{2}{3}$ and 4 works (it does) and to see if a number greater than 4 works (it doesn't). Therefore $-\frac{2}{3} < x < 4$ is the solution to the question.

- Given a graph of $y = f(x)$ you must be able to draw the graph of $y = |f(x)|$; this is done by leaving any parts of the curve above the x-axis where they are and reflecting parts of the curve under the x-axis so that the are above the x-axis. In the reflected parts, the equation of the curve would be $y = -f(x)$. For example:



goes to



.

Trigonometry

- By definition $\sec\theta \equiv \dfrac{1}{\cos\theta}$, $\quad \operatorname{cosec}\theta \equiv \dfrac{1}{\sin\theta}$, $\quad \cot\theta \equiv \dfrac{1}{\tan\theta}$.
- If you get an equation where one of the new trig functions equals a constant, then just take the reciprocal of each side and solve *à la* C2. For example

$$\sec\theta = 5 \qquad \Rightarrow \qquad \frac{1}{\cos\theta} = 5 \qquad \Rightarrow \qquad \cos\theta = \frac{1}{5}.$$

- Know the graphs of $y = \sec x$, $y = \operatorname{cosec} x$ and $y = \cot x$. Page 91/2 of your textbook.
- By dividing $\sin^2 x + \cos^2 x \equiv 1$ by $\sin^2 x$ and $\cos^2 x$ we can derive

$$1 + \cot^2 x \equiv \operatorname{cosec}^2 x \quad \text{and} \quad \tan^2 x + 1 \equiv \sec^2 x \quad \text{respectively.}$$

These create a whole new family of equations that reduce to a quadratic in disguise. For example solve $3\cot^2\theta + 5\operatorname{cosec}\theta + 1 = 0$ in the range $0 \leqslant \theta \leqslant 2\pi$. Firstly note we will need to replace the $\cot^2\theta$ by $\operatorname{cosec}^2\theta - 1$ to reduce the equation to one trig function only.

$$\begin{aligned} 3\cot^2\theta + 5\operatorname{cosec}\theta + 1 &= 0 \\ 3(\operatorname{cosec}^2\theta - 1) + 5\operatorname{cosec}\theta + 1 &= 0 \\ 3\operatorname{cosec}^2\theta + 5\operatorname{cosec}\theta - 2 &= 0 \\ (3\operatorname{cosec}\theta - 1)(\operatorname{cosec}\theta + 2) &= 0 \\ \operatorname{cosec}\theta = \tfrac{1}{3} \quad \text{or} \quad \operatorname{cosec}\theta &= -2. \end{aligned}$$

Therefore $\sin\theta = 3$ which has no solutions, or $\sin\theta = -\frac{1}{2}$ which gives the solutions $\theta = \frac{7\pi}{6}, \frac{11\pi}{6}$.

- You must know, and be able to apply, the compound angle formulae:

$$\begin{aligned} \sin(A \pm B) &= \sin A\cos B \pm \cos A\sin B, \\ \cos(A \pm B) &= \cos A\cos B \mp \sin A\sin B, \\ \tan(A \pm B) &= \frac{\tan A \pm \tan B}{1 \mp \tan A\tan B}. \end{aligned}$$

- You must know, and be able to apply, the double angle formulae (derived by setting $A = B$ in the compound angle formulae above):

$$\begin{aligned}\sin 2A &= 2\sin A\cos A,\\ \cos 2A &= \cos^2 A - \sin^2 A,\\ &= 2\cos^2 A - 1,\\ &= 1 - 2\sin^2 A,\\ \tan 2A &= \frac{2\tan A}{1-\tan^2 A}.\end{aligned}$$

  Notice there are three versions of the double angle formula for $\cos 2A$; you need to *think hard* about which form you will need for the question you are solving. You will hardly ever need the fist of the three $(\cos^2\theta - \sin^2\theta)$ because it involves two different trig functions; the aim is, usually, to get only one.

- You must be able to convert form the form $a\cos\theta \pm b\sin\theta$ into either $R\cos(\theta\pm\alpha)$ or $R\sin(\theta\pm\alpha)$; the question will specify which. This then enables us to solve equations of the form

$$a\cos\theta \pm b\sin\theta = \text{constant}.$$

  For example express $3\cos\theta - 5\sin\theta$ in the form $R\cos(\theta+\alpha)$. Always start by looking at the coefficients of $\cos\theta$ and $\sin\theta$ in the original expression; here they are 3 and 5 (ignore the sign). Sum their squares and square root (like Pythagoras) and factorise out:

$$3\cos\theta - 5\sin\theta \equiv \sqrt{34}\left[\frac{3}{\sqrt{34}}\cos\theta - \frac{5}{\sqrt{34}}\sin\theta\right].$$

  Next consider the form of the answer we are aiming for; here "$R\cos(\theta+\alpha)$". The expansion of "$R\cos(\theta+\alpha)$" is "$R(\cos\theta\cos\alpha - \sin\theta\sin\alpha)$". Comparing

$$\sqrt{34}\left[\frac{3}{\sqrt{34}}\cos\theta - \frac{5}{\sqrt{34}}\sin\theta\right] \quad \text{with} \quad R(\cos\theta\cos\alpha - \sin\theta\sin\alpha)$$

  we see instantly $R = \sqrt{34}$. We also require $\frac{3}{\sqrt{34}} = \cos\alpha$ and $\frac{5}{\sqrt{34}} = \sin\alpha$; solving either of those two we find $\alpha = 59.0^\circ$ (to 1 d.p). Therefore

$$3\cos\theta - 5\sin\theta \equiv \sqrt{34}\cos(\theta + 59.0^\circ).$$

- The trig functions all have inverses if we restrict the domain. The conventional restrictions to allow inversion are

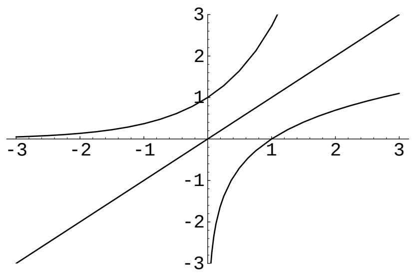
FUNCTION	DOMAIN	DOMAIN
$y = \sin x$	$-90^\circ \leqslant x \leqslant 90^\circ$	$-\frac{\pi}{2} \leqslant x \leqslant \frac{\pi}{2}$
$y = \cos x$	$0^\circ \leqslant x \leqslant 180^\circ$	$0 \leqslant x \leqslant \pi$
$y = \tan x$	$-90^\circ < x < 90^\circ$	$-\frac{\pi}{2} < x < \frac{\pi}{2}$

  Know what the graphs of $y = \sin^{-1} x$, $y = \cos^{-1} x$ and $y = \tan^{-1} x$ look like.

Exponentials & Logarithms

- Know that e is a special number in mathematics. It is approximately 2.7182818284... and it is irrational (i.e. it can't be expressed as a fraction; similar to π).

- If the base of a logarithm is e then we call it a 'natural logarithm'. Written $\log_e x \equiv \ln x$.
- We already know that logarithms and exponentials are inverses of each other with the relationships
$$\log_{10}(10^x) \equiv x \qquad \text{and} \qquad 10^{\log_{10} x} \equiv x.$$
The same is true for natural logarithms and exponents of e;
$$\ln(e^x) \equiv x \qquad \text{and} \qquad e^{\ln x} \equiv x.$$
- Below is a graph of $y = e^x$ and $y = \ln x$ showing the inverse relationship between the two (reflecting in $y = x$):



This also shows that you can't 'ln' a negative number and that $\ln 1 = 0$.
- All the laws of logarithms from C2 are true for natural logarithms (e.g. $\ln ab = \ln a + \ln b$). For example make a the subject of the following equation (a few steps missed out):
$$\begin{aligned}\ln(a-1) - \ln(a+1) &= b\\ \ln\left(\frac{a-1}{a+1}\right) &= b\\ \frac{a-1}{a+1} &= e^b\\ a(1-e^b) &= 1+e^b\\ a &= \frac{1+e^b}{1-e^b}.\end{aligned}$$
- You must understand that many physical systems can be modelled by either exponential growth or exponential decay. The most general form is $y = a \times b^x$. If $b > 1$ then the curve represents *exponential growth.* If $b < 1$ then the curve represents *exponential decay.* For example if the number of swine flu sufferers is modelled by $N = 5 \times 7^t$, where t is time measured in days, then find the amount of time for 2 billion people to have caught the disease. We need to solve $2 \times 10^9 = 5 \times 7^t$. So
$$\frac{2 \times 10^9}{5} = 7^t \quad \Rightarrow \quad \log\left(\frac{2 \times 10^9}{5}\right) = t \log 7 \quad \Rightarrow \quad t = 10.2 \text{ days! (to 3 s.f.)}.$$
(Cue dramatic music...)
- Any exponential relationship $y = a \times b^x$ can be converted to an exponential form using e. This is useful because to differentiate exponential relationships they have to be of the form $y = a \times e^{kt}$. This is done using the powerful statement that (something $\equiv e^{\ln \text{something}}$), so
$$\begin{aligned}y &= a \times b^x\\ y &= a \times e^{\ln(b^x)}\\ y &= a \times e^{x \ln b} && \text{(by 'log law' } \log(a^n) = n \log a)\\ y &= a \times e^{kx}, && \text{(where } k = \ln b).\end{aligned}$$

- An exponential can never equal zero (see graph above). Therefore if you have an equation with lots of exponential 'bits' that you can factorise out, then you are allowed to divide through (in a way that is forbidden with trig functions). For example if $2x^2e^{2x} + 3xe^{2x} - 2e^{2x} = 0$, factorise out the e^{2x} to get $e^{2x}(2x^2 + 3x - 2) = 0$. Divide by e^{2x} to get $2x^2 + 3x - 2 = 0$ which solves to $x = \frac{1}{2}$ or $x = -2$.

- To differentiate an exponential the basic building block is

$$y = e^x \qquad \Rightarrow \qquad \frac{dy}{dx} = e^x.$$

  That is *why* 'e' is so important; it gives us the exponential that differentiates to itself. Combined with the chain/product/quotient rule (below) we can build on this starting point. (Some students think that if $y = e^x$, then $\frac{dy}{dx} = xe^{x-1}$. Do not be one of them! Exponentials are fundamentally different to polynomials.)

- To differentiate a natural logarithm the basic building block is

$$y = \ln x \qquad \Rightarrow \qquad \frac{dy}{dx} = \frac{1}{x}.$$

  Again combined with the chain/product/quotient rule (below) we can build on this starting point.

Differentiation

- The basic building blocks for differentiation (that we know at present) are:

$$\begin{array}{ccc} y = ax^n & y = e^x & y = \ln x \\ \dfrac{dy}{dx} = anx^{n-1}, & \dfrac{dy}{dx} = e^x, & \dfrac{dy}{dx} = \dfrac{1}{x}. \end{array}$$

  Also we know the idea that (for $f(x) \equiv f$, $g(x) \equiv g$ and $k =$ constant)

$$\frac{d}{dx}(f+g) = \frac{d}{dx}(f) + \frac{d}{dx}(g) \quad \text{and} \quad \frac{d}{dx}(kf) = k\frac{d}{dx}(f).$$

  (In big-boy speak we say that $\frac{d}{dx}$ is a linear operator.)

- The *chain rule* is incredibly important! It states that

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}.$$

  This seems obvious from the way that differentials are written, but remember that they should not be thought of as fractions.

- If a bit of a $y = \ldots$ is making the differentiation difficult, then ask yourself the question "would making the complicated bit u make it easier for me to deal with?" For example with $y = (2x-5)^{20}$ the function would be considerably easier if $u = 2x - 5$ because y becomes $y = u^{20}$. Similarly with $y = e^{x^2+1}$ my life would be easier if $u = x^2 + 1$ because y would become $y = e^u$.

- It can be applied as follows to the example $y = (x^4 + x)^{10}$. Let $u = x^4 + x$, so

$$\begin{array}{cc} y = u^{10} & u = x^4 + x \\ \dfrac{dy}{du} = 10u^9 & \dfrac{du}{dx} = 4x^3 + 1. \end{array}$$

  Therefore $\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx} = 10u^9 \times (4x^3 + 1) = 10(4x^3 + 1)(x^4 + x)^9$.

- The above method works all the time but it is a little slow. You will notice the general result that if $y = [f(x)]^n$ then $\frac{dy}{dx} = n[f(x)]^{n-1} \times f'(x)$. So we can just write down the answer to similar problems. For example if $y = (3x^2 + 1)^5$ then $\frac{dy}{dx} = 30x(3x^2 + 1)^4$.

- We can also combine the chain rule with exponentials and logarithms to gain the following important results:

$$\begin{aligned}
\frac{d}{dx}(e^{ax}) &= ae^{ax} && \text{using } u = ax \\
\frac{d}{dx}(e^{f(x)}) &= f'(x)e^{f(x)} && \text{using } u = f(x) \\
\frac{d}{dx}(\ln ax) &= \frac{a}{ax} = \frac{1}{x} && \text{using } u = ax \\
\frac{d}{dx}(\ln f(x)) &= \frac{f'(x)}{f(x)} && \text{using } u = f(x).
\end{aligned}$$

- The *product rule* states that when $y = u \times v$ (where u and v are functions of x) we can differentiate it using the product rule. It states that

$$\frac{dy}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}.$$

For example if $y = x^2(x^3 - 1)^3$ then

$$\begin{aligned}
\frac{dy}{dx} &= [2x \times (x^3 - 1)^3] + [x^2 \times 3(x^3 - 1)^2 \times 3x^2] \\
&= 2x(x^3 - 1)^3 + 9x^4(x^3 - 1)^2 \\
&= x(x^3 - 1)^2[2(x^3 - 1) + 9x^3] \\
&= x(x^3 - 1)^2(11x^3 - 2).
\end{aligned}$$

- With the product rule you often end up with expressions such as

$$\frac{dy}{dx} = 2x^3(2x + 1)^{-4} - x^4(2x + 1)^{-5}.$$

When tidying these things up you must pull out (as always) *the lowest power* of any common elements even if they are negative or fractional; here we have x^3 and $(2x + 1)^{-5}$:

$$\begin{aligned}
\frac{dy}{dx} &= 2x^3(2x + 1)^{-4} - x^4(2x + 1)^{-5} \\
&= x^3(2x + 1)^{-5}[2(2x + 1) - x] \\
&= \frac{x^3(3x + 2)}{(2x + 1)^5}.
\end{aligned}$$

- Very similar to the product rule is the *quotient rule*. It is used for functions of the form $y = \frac{u}{v}$. It states

$$\frac{dy}{dx} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}.$$

For example differentiating $y = \frac{x^3}{x^2+1}$ gives

$$\frac{dy}{dx} = \frac{(x^2 + 1) \times 3x^2 - x^3 \times 2x}{(x^2 + 1)^2} = \frac{x^2(x^2 + 3)}{(x^2 + 1)^2}.$$

- Once again, although $\frac{dy}{dx}$ is not a fraction, it can be treated as such when taking its reciprocal, so

$$\frac{dy}{dx} = \frac{1}{\left(\frac{dx}{dy}\right)}.$$

For example if you have $V = \frac{4}{3}\pi r^3$ then $\frac{dV}{dr} = 4\pi r^2$ and also $\frac{dr}{dV} = \frac{1}{4\pi r^2}$. This idea most useful in the topic of...

- ... *connected rates of change*. Here you need to use the chain rule to 'connect' differentials you know to get one you need. Questions mostly ask you for $\frac{dy}{dx}$ (say) and you need to find a third variable to construct $\frac{dy}{dx} = \frac{dy}{d...} \times \frac{d...}{dx}$ by the chain rule. For example: The area A of a circle is increasing a rate of $3\text{cm}^2/\text{s}$, find the rate at which the radius r is increasing when $r = 20\text{cm}$. We want to find $\frac{dr}{dt}$ so

$$\begin{aligned}\frac{dr}{dt} &= \frac{dr}{dA} \times \frac{dA}{dt} \qquad \text{but} \qquad A = \pi r^2, \text{ so } \frac{dA}{dr} = 2\pi r.\\ &= \frac{1}{2\pi r} \times 3\\ &= \frac{3}{40\pi}.\end{aligned}$$

Integration

- The central idea of calculus is that integration and differentiation are the inverse operations of each other in the same way that plus is the inverse operation of subtraction. In C3 a favourite type of question is to differentiate something using the above rules and then integrate something similar later in the question. View the question as a whole!

- Our basic building blocks for integration are therefore

$$\int ax^n\,dx = \frac{ax^{n+1}}{n+1} + c, \qquad \int e^x\,dx = e^x + c, \qquad \int \frac{1}{x}\,dx = \ln x + c.$$

- A big result is gained by inspection below, but worth stating alone:

$$\int e^{ax}\,dx = \frac{1}{a}e^{ax} + c.$$

- Integration by *inspection* is effectively "spotting the answer" by an intermediate guess. The intermediate guess is then differentiated mentally and the final answer should then only be a constant factor out. Things to note are that the power on $e^{\text{something}}$ never changes and be on the lookout for integrals where the top line is *almost* the derivative of the bottom line. Here are a few examples:

INTEGRAL	GUESS	ANSWER
$\int (2x+3)^{15}\,dx$	$(2x+3)^{16}+c$	$\frac{1}{32}(2x+3)^{16}+c$,
$\int (1-3x)^{-4}\,dx$	$(1-3x)^{-3}+c$	$\frac{1}{9}(1-3x)^{-3}+c$,
$\int 2\sqrt{4x+1}\,dx$	$(4x+1)^{\frac{3}{2}}+c$	$\frac{1}{3}(4x+1)^{\frac{3}{2}}+c$,
$\int 2e^{3x-5}\,dx$	$e^{3x-5}+c$	$\frac{2}{3}e^{3x-5}+c$,
$\int 7xe^{x^2+1}\,dx$	$e^{x^2+1}+c$	$\frac{7}{2}e^{x^2+1}+c$,
$\int \frac{7}{1-4x}\,dx$	$\ln(1-4x)+c$	$-\frac{7}{4}\ln(1-4x)+c$,
$\int \frac{e^{2x}}{1-e^{2x}}\,dx$	$\ln(1-e^{2x})+c$	$-\frac{1}{2}\ln(1-e^{2x})+c$.

You must practice this a lot. . . it only comes easily after a while. [Since most students also take C3 and C4 at the same time it is worth noting that all the above can be done by the C4 technique of *integration by substitution*.]

- $\int_a^b \pi y^2\,dx$ is the volume of revolution of the curve y rotated about the x-axis between $x=a$ and $x=b$. All that is needed for you to do is calculate y^2 in terms of x from y. For example find the volume of revolution of the solid formed by rotating the curve $y=\sqrt{2x+3}$ about the x-axis between $x=10$ and $x=14$. We need to evaluate $\int_a^b \pi y^2\,dx = \int_{10}^{14} \pi y^2\,dx$. Now the curve is $y=\sqrt{2x+3}$ so to find y^2 in terms of x we need only square the equation $\Rightarrow y^2 = 2x+3$. We therefore evaluate

$$\int_{10}^{14} \pi y^2\,dx = \pi\int_{10}^{14}(2x+3)\,dx = \pi\left[x^2+3x\right]_{10}^{14} = 108\pi.$$

- For volumes of revolution around the y-axis switch the x and the y and use $\int_p^q \pi x^2\,dy$ between $y=p$ and $y=q$. For example find the volume of revolution of the solid formed by rotating the line $y=3x-2$ about the y-axis between $y=0$ and $y=5$. We need to evaluate $\int_p^q \pi x^2\,dy = \int_0^5 \pi x^2\,dy$. Now the line is $y=3x-2$ so to find x^2 in terms of y, we make x the subject and square;

$$y=3x-2 \quad\Rightarrow\quad x=\frac{y+2}{3} \quad\Rightarrow\quad x^2=\frac{y^2+4y+4}{9}.$$

We therefore evaluate

$$\int_0^5 \pi x^2\,dy = \pi\int_0^5\left(\frac{y^2+4y+4}{9}\right)dy = \frac{\pi}{9}\left[\frac{y^3}{3}+2y^2+4y\right]_0^5 = \frac{335\pi}{27}.$$

Numerical Methods

- Given an equation $f(x)=g(x)$ it is often not possible to solve them *analytically* (by algebraic manipulation) and we are forced to use numerical methods that 'home in' on the solution. You need to know two for C3: "search for a change of sign" and "fixed point iteration".

- *Search for a change of sign* 'homes in' on a solution to an equation by sandwiching the solution between two numbers. Those two numbers can gradually be brought together to improve knowledge of where the solution is. Given an equation ($e^x = 15x+3$, say) it is best to get one side equal to zero ($0 = e^x - 15x - 3$). Then *define* $f(x) = e^x - 15x - 3$. Then put values into $f(x)$ and look for a change of sign.

$$\begin{aligned}
f(-1) &= 12.36787944\ldots & +\text{ve}\\
f(0) &= -2 & -\text{ve}\\
f(1) &= -15.28171817\ldots & -\text{ve}\\
f(2) &= -25.6109439\ldots & -\text{ve}\\
f(3) &= -27.91446308\ldots & -\text{ve}\\
f(4) &= -8.401849967\ldots & -\text{ve}\\
f(5) &= 70.4131591\ldots & +\text{ve}
\end{aligned}$$

From this we can see that there are two solutions (α and β) such that

$$-1 < \alpha < 0 \qquad \text{and} \qquad 4 < \beta < 5.$$

If you were interested in finding β to 2 decimal places (say) then you would next evaluate

$$f(4.1), f(4.2), \ldots, f(4.9)$$

and you should discover $4.1 < \beta < 4.2$. Next

$$f(4.11), f(4.12), \ldots, f(4.19)$$

and you should discover $4.18 < \beta < 4.19$. You should resist the temptation (however strong) to state $\beta = 4.18$ (to 2 d.p.) as your final answer. It is still possible that the answer could still be $\beta = 4.19$ (to 2 d.p.). You must check 4.185 and then think! *Hard*!

We find $f(4.185) < 0$, so the change of sign exists between 4.185 and 4.19 so final stated answer should be $\beta = 4.19$ (to 2 d.p.)

- *Fixed point iteration* works by taking an equation and rearranging to isolate an x in the form $x = g(x)$. From this rearrangement we form an iterative formula

  $$x_{n+1} = g(x_n).$$

  It is important to note that there exist many possible rearrangements of an equation; for the equation $x^3 - 3x + 4 = 0$ here are a few:

  $$x = \sqrt[3]{3x-4} \qquad x = \frac{x^3+4}{3} \qquad x = \frac{3x-4}{x^2}.$$

  However, the exam will usually specify which one they want[8]. In the above example let's use the first one and create $x_{n+1} = \sqrt[3]{3x_n - 4}$. The starting value for the iteration is denoted x_0 (or x_1) and you should either use the value specified in the question or choose a value close to where you know the solution exists[9]. Here let's use $x_0 = -1$.

  To save time, you can use your calculator to speed up the process a lot. First type "$-1 =$" to enter -1 as the "Ans" on your calculator. Then type "$\sqrt[3]{(3 \times \text{Ans} - 4)}$". Press "="

[8] *rant* It is worth noting just how unrealistic this situation is; in practice you will discover that some of these rearrangements work a treat and some of them fail miserably. This should be part of a coursework (*à la* MEI) and not part of an exam! *rant*

[9] If the first part of a question gets you to show the solution exists between 2.1 and 2.2 then start with $x_0 = 2.1$

repeatedly to see the results of the iteration. You should find:

$$
\begin{aligned}
x_0 &= -1 \\
x_1 &= -1.912931183 \\
x_2 &= -2.134410543 \\
x_3 &= -2.18324263 \\
x_4 &= -2.19321102 \\
x_5 &= -2.19528142 \\
&\vdots \qquad \cdots \text{keep pressing “=” lots and eventually} \cdots \\
x &= -2.19582 \text{ to (5 d.p.)}
\end{aligned}
$$

Always state the accuracy to which you give your answer (sig figs or d.p.s). If when you keep pressing “=” it settles to one number we say the iteration *converges*; otherwise it *diverges*.

- Sometimes a question gives you an iteration and asks for the equation which has been solved (or to show that the number the iteration converges to represents a solution of another given equation). All you do is remove the $n+1$ and n subscripts and rearrange: For example

$$
\begin{aligned}
x_{n+1} &= \ln(\sqrt[3]{1-2x_n}) \\
x &= \ln(\sqrt[3]{1-2x}) \\
e^x &= \sqrt[3]{1-2x} \\
e^{3x} + 2x - 1 &= 0.
\end{aligned}
$$

Simpson's Rule

- Similar to the trapezium rule is *Simpson's Rule*. It can be used to approximate integrals. It uses a quadratic curve to approximate the curve rather than a straight line, and is therefore rather more accurate. Unlike the trapezium rule it is hard to say whether the approximation will be an over or under-estimate. Therefore you don't get questions on it.

- In class I refer to “Simpson Chunks”; this is a Stone-ism you will not hear elsewhere. One “Simpson Chunk” contains two intervals/strips and three ordinates.

  In general

  n “Simpson Chunks” $\Leftrightarrow$ $2n$ Intervals/Strips $\Leftrightarrow$ $2n+1$ Ordinates.

  The heights on the ordinates are the y-values of the curve. They are labelled y_0, y_1, …, y_{2n}. **Never** forget that the first height on the left is denoted y_0 and **not** y_1; if you do the whole question will go wrong because your ‘odds’ and ‘evens’ will be wrong!

- Simpson's Rule states (where h is the distance between each ordinate/height):

$$\int_a^b y\,dx \approx \frac{h}{3}\left[y_0 + y_{2n} + 4(y_1 + y_3 + \cdots + y_{2n-1}) + 2(y_2 + y_4 + \cdots + y_{2n-2})\right].$$

$$\int_a^b y\,dx \approx \frac{h}{3}\left[\text{‘sum ends’} + 4(\text{‘sum internal odds’}) + 2(\text{‘sum internal evens’})\right].$$

- For example use 8 intervals to approximate $\int_{-4}^{4} \frac{1}{1+x^2}\,dx$. Each interval must have width 1 since the total width is 8. There must be nine ordinates. A table for the ordinates:

$$\begin{array}{ccccccccc} y_0 & y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_7 & y_8 \\ \frac{1}{17} & \frac{1}{10} & \frac{1}{5} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{5} & \frac{1}{10} & \frac{1}{17} \end{array}.$$

  Therefore

$$\begin{aligned} \int_{-4}^{4} \frac{1}{1+x^2}\,dx &\approx \frac{1}{3}\left[\frac{1}{17}+\frac{1}{17}+4\left(\frac{1}{10}+\frac{1}{2}+\frac{1}{2}+\frac{1}{10}\right)+2\left(\frac{1}{5}+1+\frac{1}{5}\right)\right] \\ &\approx 2.573 \text{ (to 3 d.p.)}. \end{aligned}$$

OCR Core 4 Module Revision Sheet

The C4 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J.M.S.

Algebra

- Review binomial expansion from C2 for $(x+y)^n$ for positive integer n. Notice that it is valid for *any* x and y and that the expansion has $n+1$ terms.
- The general binomial expansion is given by

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \cdots$$

and is valid for any n (fractional or negative) but $-1 < x < 1$ (i.e. $|x| < 1$). Notice also it must start with a 1 in the brackets. For example expand $(4-x)^{-1/2}$.

$$\begin{aligned}(4-x)^{-1/2} &= \left(4\left(1-\frac{x}{4}\right)\right)^{-1/2}\\ &= \frac{1}{2}\left(1-\frac{x}{4}\right)^{-1/2}\\ &= \frac{1}{2}\left[1+\left(-\frac{1}{2}\right)\left(-\frac{x}{4}\right)+\frac{(-\frac{1}{2})(-\frac{3}{2})}{2!}\left(-\frac{x}{4}\right)^2+\frac{(-\frac{1}{2})(-\frac{3}{2})(-\frac{5}{2})}{3!}\left(-\frac{x}{4}\right)^3+\cdots\right]\\ &= \frac{1}{2}\left[1+\frac{x}{8}+\frac{3x^2}{128}+\frac{15x^3}{3072}+\cdots\right] = \frac{1}{2}+\frac{x}{16}+\frac{3x^2}{256}+\frac{15x^3}{6144}+\cdots.\end{aligned}$$

It is only valid for $|x/4| < 1 \quad \Rightarrow \quad |x| < 4$.

- Another example: Find first 3 terms in the expansion for $\dfrac{(3+x)^2}{1+\frac{x}{2}}$.

$$(3+x)^2\left(1+\frac{x}{2}\right)^{-1} = (9+6x+x^2)\left(1-\frac{x}{2}+\frac{x^2}{4}+\ldots\right) = 9+\frac{3}{2}x+\frac{1}{4}x^2+\ldots.$$

- You must be able to simplify algebraic fractions; best tactic is always to factorise and cancel:

$$\frac{4x^5-10x^4-6x^3}{12x^6-18x^5-12x^4} = \frac{2x^3(2x^2-5x-3)}{6x^4(2x^2-3x-2)} = \frac{2x^3(2x+1)(x-3)}{6x^4(x-2)(2x+1)} = \frac{x-3}{3x(x-2)}.$$

- You must be able to divide a polynomial ($p(x)$) by a divisor ($a(x)$), finding the quotient ($q(x)$) and remainder ($r(x)$). If there is no remainder then $a(x)$ is a factor of $p(x)$. It is always such that

$$\frac{p(x)}{a(x)} = q(x) + \frac{r(x)}{a(x)} \qquad \Rightarrow \qquad p(x) = a(x)q(x) + r(x).$$

The order of $q(x)$ is the order of $p(x)$ subtract the order of $a(x)$. The order of $r(x)$ is *at most* one less than $a(x)$. For example if you have a quintic (power 5 polynomial) divided by a quadratic you would expect

$$\frac{\text{quintic}}{\text{quadratic}} = \text{cubic} + \frac{\text{linear}}{\text{quadratic}},$$
$$\frac{\text{quintic}}{\text{quadratic}} = Ax^3 + Bx^2 + Cx + D + \frac{Ex+F}{\text{quadratic}}.$$

Of course it *may* turn out that $Ex + F$ is just a constant or zero (if the cubic divides the quintic).

- Division is most easily done step-by-step working *down* the powers of $p(x)$. For example divide $2x^4 - x^3 + 3x^2 - 7x + 1$ by $x^2 + 2x + 3$:

$$\begin{aligned} 2x^4 - x^3 + 3x^2 - 7x + 1 &= (x^2+2x+3)(\text{quadratic}) + (\text{remainder}) \\ &= (x^2+2x+3)(2x^2 + \dots) + (\text{remainder}) && x^4\checkmark \\ &= (x^2+2x+3)(2x^2 - 5x + \dots) + (\text{remainder}) && x^3\checkmark \\ &= (x^2+2x+3)(2x^2 - 5x + 7) + (\text{remainder}) && x^2\checkmark \\ &= (x^2+2x+3)(2x^2 - 5x + 7) - 6x + \text{const.} && x\checkmark \\ &= (x^2+2x+3)(2x^2 - 5x + 7) - 6x - 20 && \text{const.}\checkmark \end{aligned}$$

  Therefore $q(x) = 2x^2 - 5x + 7$ and $r(x) = -6x - 20$.

- Partial fractions is effectively the reverse of combining together two algebraic fractions. For example

$$\frac{1}{x+1} + \frac{1}{x+2} \quad \begin{array}{c} \longrightarrow \\ \longleftarrow \end{array} \quad \begin{array}{c} \text{Algebraic Fractions} \\ \text{Partial Fractions} \end{array} \quad \begin{array}{c} \longrightarrow \\ \longleftarrow \end{array} \quad \frac{2x+3}{(x+1)(x+2)}.$$

  You can use partial fractions provided the order of the top line is less than the order of the bottom line.

- If you simply have a product of linear factors on the bottom line then you split out into that many terms with constants placed on top (usually denoted by A, B, C, etc.). Place this equal (in an identity "$\equiv$") to the original expression and then multiply through to get rid of the denominators. For example:

$$\frac{7x-1}{(2x+1)(x-1)} \equiv \frac{A}{2x+1} + \frac{B}{x-1}$$
$$\Rightarrow \qquad 7x - 1 \equiv (x-1)A + (2x+1)B.$$

  Because this is an identity we can choose any value of x we fancy to help us discover A and B. In this case letting $x = 1$ is a good choice because one of the brackets become zero. Similarly $x = -\frac{1}{2}$ is another good choice[10]; we put these into the identity.

$$x = 1 \quad \Rightarrow \quad 7 - 1 \equiv 3B \quad \Rightarrow \quad \underline{B = 2},$$
$$x = -\tfrac{1}{2} \quad \Rightarrow \quad -\tfrac{7}{2} - 1 \equiv -\tfrac{3}{2}A \quad \Rightarrow \quad \underline{A = 3}.$$

  Therefore $\dfrac{7x-1}{(2x+1)(x-1)} \equiv \dfrac{3}{2x+1} + \dfrac{2}{x-1}$.

[10]If you don't spot (or run out of) clever values to use, fret not! Just put in some other numbers (usually something 'nice' like $x = 0$, $x = 1$ or $x = -1$) and solve the resulting equations. If your girlfriend's lucky number is 53 and your mistress's lucky number is 178 then feel free to use those if you wish, although I wouldn't advise it; choose simple numbers!

- If you have a repeated factor in the denominator then you deal with it as follows (notice the top line is a quadratic and the bottom a cubic, so partial fractions are fine):

$$\frac{5x^2-10x+1}{(x-3)(x-1)^2} \equiv \frac{A}{x-3}+\frac{B}{x-1}+\frac{C}{(x-1)^2}$$
$$\Rightarrow \quad 5x^2-10x+1 \equiv (x-1)^2A+(x-3)(x-1)B+(x-3)C.$$

  Similarly, good values of x to choose are $x=1$ and $x=3$:

$$x=1 \quad \Rightarrow \quad 5-10+1 \equiv -2C \quad \Rightarrow \quad \underline{C=2},$$
$$x=3 \quad \Rightarrow \quad 45-30+1 \equiv 4A \quad \Rightarrow \quad \underline{A=4}.$$

  This tells us that

$$5x^2-10x+1 \equiv 4(x-1)^2+(x-3)(x-1)B+2(x-3),$$

  but it hasn't told us B. I would consider $x=0$, here, and sub in to discover

$$0-0+1 \equiv 4+3B-6 \Rightarrow \underline{B=1}.$$

  Therefore $\dfrac{5x^2-10x+1}{(x-3)(x-1)^2} \equiv \dfrac{4}{x-3}+\dfrac{1}{x-1}+\dfrac{2}{(x-1)^2}$.

- Partial fractions are often very useful in evaluating integrals. If you see a quadratic in the bottom line of an integral, then one option your brain should turn to is "can this be split into partial fractions?". For example in this integral, it can be split and then integrated:

$$\begin{aligned}\int \frac{x+10}{x^2+5x+4}\,dx &= \int \frac{x+10}{(x+1)(x+4)}\,dx \\ &= \int \frac{3}{x+1}-\frac{2}{x+4}\,dx \\ &= 3\ln(x+1)-2\ln(x+4)+c.\end{aligned}$$

Differentiation & Integration

- Know the contents of the formula booklet well. Very well! Lots of problems can be solved simply by looking at the table of differentials and integrals and knowing that integration 'undoes' a differentiation. Some questions get you to differentiate something and *then* get you to integrate something similar. *Always view the question as a whole!*

- When using radians we can differentiate the trigonometric functions. The results are as follows:

$$\begin{array}{lll} y=\sin x & y=\cos x & y=\tan x \\ \dfrac{dy}{dx}=\cos x, & \dfrac{dy}{dx}=-\sin x, & \dfrac{dy}{dx}=\sec^2 x. \end{array}$$

  One can derive the third result from the other two using the quotient rule and that $\tan x \equiv \frac{\sin x}{\cos x}$.

- You can also use these results along with the chain rule to differentiate functions like the following; $y=\sin(x^2+1)$ by letting $u=x^2+1$ and $y=(\tan x)^{10}$ by letting $u=\tan x$.

$$\begin{array}{ll} y=\sin(x^2+1) & y=(\tan x)^{10} \\ \dfrac{dy}{dx}=2x\cos(x^2+1), & \dfrac{dy}{dx}=10\sec^2 x(\tan x)^9. \end{array}$$

- Integration by substitution is a way of integrating by replacing the variable given to you (usually x) and replacing it by another (usually u). These days the substitution you are to use is given to you in the exam, but practice will get you better at spotting what to substitute (usually the most complicated term in the integration or the denominator of a fraction). For example $\int x^3(x^4+1)^7\,dx$ we should use $u = x^4+1$.

$$\begin{aligned}
&\int x^3(x^4+1)^7\,dx & u &= x^4+1\\
=&\int x^3u^7\,dx & \frac{du}{dx} &= 4x^3\\
=&\int x^3u^7\,\frac{du}{4x^3} & \frac{du}{4x^3} &= dx\\
=&\frac{1}{4}\int u^7\,du & &\\
=&\frac{u^8}{32}+c = \frac{(x^4+1)^8}{32}+c. & &
\end{aligned}$$

We have effectively *"used and abused"* u to help us to get the answer. (NOTE: I have been *very* sloppy in the above integration because I have mixed my x and u variables; you shouldn't really do this, but it makes the process of conversion clearer.)

- When dealing with definite integrals we need to also convert the limits of the integration and there is no need to convert back to x at the end since all definite integrals are merely numbers. For example

$$\begin{aligned}
&\int_3^4 2x\sqrt{x^2-4}\,dx & u &= x^2-4 & x=3 &\Rightarrow u=5\\
=&\int_5^{12} 2xu^{1/2}\,\frac{du}{2x} & \frac{du}{dx} &= 2x & x=4 &\Rightarrow u=12\\
=&\int_5^{12} u^{1/2}\,du & \frac{du}{2x} &= dx & &\\
=&\left[\frac{2}{3}u^{3/2}\right]_5^{12} & & & &\\
=&\,20.3 \text{ (3sf)}. & & & &
\end{aligned}$$

- Know the result $\int e^{ax}\,dx = \frac{1}{a}e^{ax}+c$.

- We know that if $y = e^{f(x)}$ then $\frac{dy}{dx} = f'(x)e^{f(x)}$. Therefore by reversal we find

$$\int f'(x)e^{f(x)}\,dx = e^{f(x)}+c.$$

  For example[11]

$$\int x^3e^{x^4}\,dx = \frac{1}{4}\int 4x^3e^{x^4}\,dx = \frac{1}{4}e^{x^4}+c.$$

- Know that $\int \frac{1}{x}\,dx = \ln x + c$.

- We know (by the chain rule) that if $y = \ln(f(x))$ then $\frac{dy}{dx} = \frac{f'(x)}{f(x)}$. Therefore by reversal we find

$$\int \frac{f'(x)}{f(x)}\,dx = \ln|f(x)|+c.$$

[11] This could also have been evaluated (more slowly) by a substitution of $u = x^4$ which would then have reduced to $\int x^3e^{x^4}\,dx = \frac{1}{4}\int e^u\,du = \frac{1}{4}e^{x^4}+c$.

Be on the lookout for expressions where the top line is almost the derivative of the bottom line. For example[12]

$$\int \frac{x^3}{x^4+1}\,dx = \frac{1}{4}\int \frac{4x^3}{x^4+1}\,dx = \frac{1}{4}\ln|x^4+1| + c.$$

- Know the results

$$\int \cos ax\,dx = \frac{1}{a}\sin ax + c \qquad \text{and} \qquad \int \sin ax\,dx = -\frac{1}{a}\cos ax + c.$$

- Always be on the look out for integrals involving a mixture of trigonometric functions. These are usually handled by means of a substitution. For example $\int \cos x(\sin x)^7\,dx$ is best handled by $u = \sin x$ to give $\frac{1}{8}(\sin x)^8 + c$.

- Also know the useful results (all derived from reverse chain rule)

$$\int f'(x)\cos f(x)\,dx = \sin f(x) + c \qquad \text{and} \qquad \int f'(x)\sin f(x)\,dx = -\cos f(x) + c.$$

  For example $\int x^3\cos(x^4)\,dx = \frac{1}{4}\sin(x^4) + c$.

- When an integral is made up of two 'bits' then we can sometimes use *integration by parts.* It states

$$\int u\frac{dv}{dx}\,dx = uv - \int v\frac{du}{dx}\,dx.$$

  So you will need to decide which 'bit' of the integral you will need to differentiate and which 'part' to integrate. For example in $\int x\sin x\,dx$ it is quite clear that we will need to differentiate the x 'part' and integrate the $\sin x$ 'part'.

$$\begin{aligned}\int x\sin x\,dx &= -x\cos x - \int -\cos x\,dx\\ &= -x\cos x + \sin x + c.\end{aligned}$$

- Another example (this time a definite integral)

$$\begin{aligned}\int_0^2 xe^{2x}\,dx &= \left[\tfrac{1}{2}xe^{2x}\right]_0^2 - \int_0^2 \tfrac{1}{2}e^{2x}\,dx\\ &= \left[\tfrac{1}{2}xe^{2x}\right]_0^2 - \left[\tfrac{1}{4}e^{2x}\right]_0^2\\ &= \left(e^4 - 0\right) - \left(\tfrac{e^4}{4} - \tfrac{1}{4}\right) = \tfrac{3e^4}{4} + \tfrac{1}{4}.\end{aligned}$$

- Initially $\int \ln x\,dx$ looks nothing like it has anything to do with integration by parts because it only has one 'part'. However if we write $\ln x$ as $1 \times \ln x$ we can integrate the 1 and differentiate the $\ln x$:

$$\int \ln x\,dx = \int 1\ln x\,dx = x\ln x - \int x.\tfrac{1}{x}\,dx = x\ln x - x + c.$$

  This principle can be extended to integrals of the type $\int x^n \ln x\,dx$:

$$\int x^n\ln x\,dx = \tfrac{x^{n+1}}{n+1}\ln x - \int \tfrac{x^{n+1}}{n+1}.\tfrac{1}{x}\,dx = \tfrac{x^{n+1}}{n+1}\ln x - \tfrac{x^{n+1}}{(n+1)^2} + c.$$

[12]Again, this could also have been evaluated by the substitution $u = x^4 + 1$.

- Very occasionally you will need to integrate by parts *twice* to get the final answer. This will almost always be of the form $\int kx^2(\text{something})\,dx$. For example find $\int x^2e^{2x}\,dx$:

$$\begin{aligned}\int x^2e^{2x}\,dx &= \frac{x^2}{2}e^{2x} - \left(\int xe^{2x}\,dx\right)\\ &= \frac{x^2}{2}e^{2x} - \left(\frac{x}{2}e^{2x} - \int \frac{1}{2}e^{2x}\,dx\right)\\ &= \frac{x^2}{2}e^{2x} - \left(\frac{x}{2}e^{2x} - \frac{1}{4}e^{2x}\right) + c\\ &= \frac{x^2}{2}e^{2x} - \frac{x}{2}e^{2x} + \frac{1}{4}e^{2x} + c.\end{aligned}$$

- For the cases of $\int \sin^2 x\,dx$ and $\int \cos^2 x\,dx$ we need to recall two forms of the double angle formula for $\cos 2x$: Namely $\cos 2x = 1 - 2\sin^2 x$ (for $\int \sin^2 x\,dx$) and $\cos 2x = 2\cos^2 x - 1$ (for $\int \cos^2 x\,dx$). Re-arranging them both we find:

$$\begin{aligned}\int \sin^2 x\,dx &= \int (\tfrac{1}{2} - \tfrac{1}{2}\cos 2x)\,dx = \tfrac{x}{2} - \tfrac{1}{4}\sin 2x + c,\\ \int \cos^2 x\,dx &= \int (\tfrac{1}{2} + \tfrac{1}{2}\cos 2x)\,dx = \tfrac{x}{2} + \tfrac{1}{4}\sin 2x + c.\end{aligned}$$

Learn the technique rather than the result!

Implicit Functions

- Given a function in the form $y = f(x)$ we can differentiate it. Implicit differentiation allows us to differentiate a function without making y the subject first. It uses the chain rule that

$$\frac{d\,f(y)}{dx} = \frac{d\,f(y)}{dy} \times \frac{dy}{dx}.$$

So all you do is differentiate the y bits with respect to y and then multiply by $\frac{dy}{dx}$. For example differentiate $y^4 + x^4 = \sin y$ with respect to x. This gives

$$4y^3\frac{dy}{dx} + 4x^3 = \cos y\frac{dy}{dx} \qquad \Rightarrow \qquad \frac{dy}{dx} = \frac{4x^3}{\cos y - 4y^3}.$$

You must be on the lookout for products in terms of x and y; for example $2xy = e^{2y}$ would differentiate to

$$2x\frac{dy}{dx} + 2y = 2e^{2y}\frac{dy}{dx} \quad \text{so} \quad \frac{dy}{dx} = \frac{2y}{2e^{2y} - 2x} = \frac{y}{e^{2y} - x}.$$

- Another example; find all the stationary points on the curve $x^2 + y^2 + xy = 3$. Differentiating w.r.t. x we find

$$2x + 2y\frac{dy}{dx} + y + x\frac{dy}{dx} = 0 \qquad \Rightarrow \qquad \frac{dy}{dx} = -\frac{2x + y}{2y + x}.$$

Stationary points are where $\frac{dy}{dx} = 0$ so solve

$$0 = -\frac{2x + y}{2y + x} \qquad \Rightarrow \qquad y = -2x.$$

Substituting this *back into the original equation* we find

$$x^2 + (-2x)^2 + x(-2x) = 3 \qquad \Rightarrow \qquad x = \pm 1 \qquad \Rightarrow \qquad \text{Points are } (1, -2) \text{ and } (-1, 2).$$

- If you discover $\dfrac{dy}{dx} = \dfrac{f(x,y)}{g(x,y)}$ and are asked to find where the tangents to a curve are parallel to the y-axis (i.e. vertical) then you need to solve where the bottom line is zero, i.e. solve $g(x,y) = 0$.

Parametric Equations

- A parametric equation is one where

$$x = f(\text{some parameter}) \quad \text{and} \quad y = g(\text{some parameter}).$$

  The parameter in a set of parametric equations can be any letter, but usually either t or θ. As the parameter varies it sketches out a curve. If no restriction is given, assume the parameter varies $-\infty < t < \infty$. However the parameter can be restricted in any way, defined by an inequality on the parameter. Standard examples: $0 \leqslant \theta < 2\pi$ or $-\pi < \theta \leqslant \pi$.

- You must be able to convert a parametric curve to Cartesian form. Sometimes this is just obvious; isolate t from one of the equations and put into the other. For example

$$\begin{array}{c} x = 2t \\ y = \frac{t}{t+1} \end{array} \quad \Rightarrow \quad t = \frac{x}{2} \quad \Rightarrow \quad y = \frac{\frac{x}{2}}{\frac{x}{2}+1} = \frac{x}{x+2}.$$

  If one of x or y involves a "sin" and the other involves a "cos" then use $\sin^2 x + \cos^2 x = 1$:

$$\begin{array}{c} x = 3\cos\theta \\ y = \sin\theta + 4 \end{array} \quad \Rightarrow \quad \begin{array}{c} \cos^2\theta = \left(\frac{x}{3}\right)^2 \\ \sin^2\theta = (y-4)^2 \end{array} \quad \Rightarrow \quad \frac{x^2}{9} + (y-4)^2 = 1.$$

- To find where a line intersects a parametric curve, place the parameters (in terms of t) into the line and solve for t. For example find the points of intersection of

$$\begin{array}{c} x = 2t^2 + 1 \\ y = \frac{1}{t} \end{array} \quad \text{and the line} \quad x + 4y = 7.$$

  Replace the x and y in the line by $2t^2+1$ and $\frac{1}{t}$ respectively. Therefore

$$x + 4y = 7, \quad \Rightarrow \quad (2t^2+1) + 4\left(\frac{1}{t}\right) = 7, \quad \Rightarrow \quad t^3 - 3t + 2 = 0.$$

  This cubic factorises to $(t-1)^2(t+2) = 0$ which gives $t = 1$ or $t = -2$ as solutions. Plugging these back into the original parametric equation we discover the two points $(3,1)$ and $(9, -\frac{1}{2})$. [It is worth noting that the squared factor $(t-1)^2$ in the cubic implies the the line is a *tangent* to the curve at the point $(3,1)$.]

- To differentiate a parametric curve

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt}.$$

  For stationary points you still equate $\frac{dy}{dx} = 0$ and solve. All other properties you are used to for normals and tangents still work.

For example find the equation of the normal to $x = 2t^3$, $y = \frac{1}{t}$ at the point $(16, \frac{1}{2})$. Firstly we need to discover the value of the parameter at the stated point: $y = \frac{1}{t} = \frac{1}{2}$ implies $t = 2$. Next differentiate and put in $t = 2$:

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{-t^{-2}}{6t^2} = -\frac{1}{6t^4}.$$

$$\text{When } t = 2, \ \frac{dy}{dx} = -\frac{1}{96}.$$

Therefore the gradient of the normal is 96. Thus, $y - \frac{1}{2} = 96(x - 16)$ which 'simplifies' to $192x - 2y - 3071 = 0$.

- In harder examples questions will leave the parameter unevaluated; either leaving it as t or setting $t = p$. For example, find the equation of the tangent to the curve $x = 2t$, $y = \frac{1}{t^2}$ where $t = p$. When $t = p$, the point becomes $(2p, \frac{1}{p^2})$. Differentiating we find

$$\frac{dy}{dx} = \tfrac{-2t^{-3}}{2} = -\tfrac{1}{t^3}.$$

Therefore the gradient of the tangent when $t = p$ is $-\frac{1}{p^3}$. Therefore the tangent is

$$y - \tfrac{1}{p^2} = -\tfrac{1}{p^3}(x - 2p) \qquad \Rightarrow \qquad x + p^3y = 3p.$$

The question could further be extended to find the area of the triangle formed by the points where the tangent crosses the x-axis and y-axis and the origin. The tangent $(x+p^3y = 3p)$ crosses the x-axis when $y = 0$ which gives $x = 3p$. The tangent crosses the y-axis when $x = 0$ which gives $y = \frac{3}{p^2}$. So the three vertices of the triangle are at $(0, 0)$, $(0, \frac{3}{p^2})$ and $(3p, 0)$. The area of the triangle is therefore

$$\text{Area} = \tfrac{1}{2} \times \text{base} \times \text{height} = \tfrac{1}{2} \times 3p \times \tfrac{3}{p^2} = \tfrac{3}{2p}.$$

Differential Equations

- If you are told that ("something") is proportional to ("something else") then we write ("something") $\propto$ ("something else"). This implies that

$$(\text{“something”}) = \pm k(\text{“something else”})$$

for some *constant* k. k can then be determined by putting in one pair of values (x, y) into the equation. If you read that something is decreasing then use $-k$, if it is increasing then use $+k$. The expression "varies as" also implies proportionality between two quantities.

- The words "rate of change of (something)" $\Rightarrow \dfrac{d(\text{something})}{dt}$. Also the word "initially" $\Rightarrow$ $t = 0$. Also be on the lookout for phrases such as "where t is measured from now" implying $t = 0$ now.

- In simple cases you need to be able to construct a differential equation of a situation. For example: The number of people infected with bird flu (N) is growing at a rate proportional to the square of the number of people infected:

$$\frac{dN}{dt} \propto N^2 \qquad \Rightarrow \qquad \frac{dN}{dt} = +kN^2.$$

- The notation dy/dx lets us believe the it is a normal fraction. Although this is not the case we can manipulate it like a fraction in a differential equation. You must move the variables to different sides of the equation and integrate (separation of variables). Only add the ever-present "$+c$" to one side. For example solve

$$\frac{dy}{dx} = y^2 \cos x \quad \Rightarrow \quad \int \frac{1}{y^2}\, dy = \int \cos x\, dx \quad \Rightarrow \quad y = -\frac{1}{\sin x + c}.$$

$$\frac{dN}{dt} = +kN^2 \quad \Rightarrow \quad \int \frac{1}{N^2}\, dN = \int k\, dt \quad \Rightarrow \quad N = \frac{-1}{kt + c}.$$

In the second example above you will notice that there are two constants; the constant of proportionality and the arbitrary integration constant. This means you will need to be given two pieces of data (t_1, N_1) and (t_2, N_2) to figure them both out.

- A final example involving partial fractions:

$$\begin{aligned}
(3P+1)\frac{dP}{dt} &= kt(P-1)(P+3) \\
\int \tfrac{3P+1}{(P-1)(P+3)}\, dP &= \int kt\, dt \\
\int \tfrac{1}{P-1} + \tfrac{2}{P+3}\, dP &= \tfrac{kt^2}{2} + c \\
\ln(P-1) + 2\ln(P+3) &= \tfrac{kt^2}{2} + c \\
\ln(P-1)(P+3)^2 &= \tfrac{kt^2}{2} + c.
\end{aligned}$$

- If the arbitrary constant is left unevaluated, then your solution represents the *general solution* of the differential equation. If you put a value in to work out its value then your solution is called the *particular solution*.

Vectors

- The vector $\begin{pmatrix}3\\4\end{pmatrix}$ can be written $3\mathbf{i} + 4\mathbf{j}$ and represents a vector going 3 right and 4 up. By Pythagoras' Theorem it can be shown that the magnitude of this vector is $\sqrt{3^2 + 4^2} = 5$ and by trigonometry the direction is $\tan^{-1}\frac{4}{3}$ above the horizontal.
- Two vectors are equal if their magnitudes and directions are the same. Two vectors are parallel if one is a scalar multiple of the other. For example show that $\begin{pmatrix}2\\3\end{pmatrix}$ is parallel to $\begin{pmatrix}3\\4.5\end{pmatrix}$; so show that $1.5 \times \begin{pmatrix}2\\3\end{pmatrix} = \begin{pmatrix}3\\4.5\end{pmatrix}$.
- When multiplying a vector by a positive scalar it changes the length of the vector but not the direction. If the scalar is negative then it also reverses the direction of the vector. For example $3\begin{pmatrix}2\\-3\end{pmatrix} = \begin{pmatrix}6\\-9\end{pmatrix}$.
- When adding vectors, you just add the x components and add the y components. For example $\begin{pmatrix}2\\-3\end{pmatrix} + \begin{pmatrix}2\\5\end{pmatrix} = \begin{pmatrix}4\\2\end{pmatrix}$.
- A unit vector is a vector with a magnitude 1. A unit vector in a given direction can be constructed by dividing a vector by its magnitude. For example the unit vector from $\begin{pmatrix}2\\3\end{pmatrix}$ is $\frac{1}{\sqrt{2^2+3^2}} \times \begin{pmatrix}2\\3\end{pmatrix} = \frac{1}{\sqrt{13}} \times \begin{pmatrix}2\\3\end{pmatrix}$.

- You must know the geometric interpretation of $\mathbf{a}+\mathbf{b}$ and $\mathbf{a}-\mathbf{b}$. Also know that in general if you have position vectors $\overrightarrow{OA} = \mathbf{a}$ and $\overrightarrow{OB} = \mathbf{b}$ then $\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$.

- It cannot be stressed enough that <u>*subtraction*</u> is the most important operation with vectors. If you wish to travel *from* one point ($\mathbf{a}$) *to* another ($\mathbf{b}$) then we use subtraction: $\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$.

  I will repeat that! If you wish to travel *from* $\mathbf{a}$ *to* $\mathbf{b}$ then use subtraction: $\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$.

- If you wish to calculate a length in 3D space then you merely need to calculate the magnitude of the vector that travels between the two points (i.e. $|\mathbf{b} - \mathbf{a}|$).

- A line can be written in vector form. If you know a line goes through a point (a, b) and has the gradient m then its vector form is $\mathbf{r} = \begin{pmatrix} a \\ b \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ m \end{pmatrix}$ where λ is a scalar that takes different values on different points on the line. The vector $\begin{pmatrix} 1 \\ m \end{pmatrix}$ can be re-written to make the components 'nicer'. For example $\begin{pmatrix} 1 \\ 2/3 \end{pmatrix} \Rightarrow \begin{pmatrix} 3 \\ 2 \end{pmatrix}$. (These vectors are not equal, but they have the same direction.) The most general form is

$$\mathbf{r} = \mathbf{a} + \lambda\mathbf{d}$$

  where $\mathbf{a}$ is the point it passes through and $\mathbf{d}$ is the direction vector (i.e. the vector that points *along* the line).

- We can therefore show that the equation of the line through $\mathbf{a}$ and $\mathbf{b}$ is given by $\mathbf{r} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a})$, because $\mathbf{b} - \mathbf{a}$ is the vector that travels *from* $\mathbf{a}$ *to* $\mathbf{b}$ <u>along the line</u>. For example find the line that passes through $(2, 3, 1)$ and $(3, 6, -1)$. This gives

$$\mathbf{r} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} \qquad \text{or} \qquad \mathbf{r} = \begin{pmatrix} 3 \\ 6 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ -3 \\ 2 \end{pmatrix}.$$

- To find the angle between two vectors we use the scalar product result

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$$

  where $|\mathbf{u}|$ represents the magnitude of vector $\mathbf{u}$. From this we can see that two vectors are perpendicular if their scalar product is zero.

- The scalar product is most easily calculated as follows; $\begin{pmatrix} a_x \\ a_y \end{pmatrix} \cdot \begin{pmatrix} b_x \\ b_y \end{pmatrix} = a_x b_x + a_y b_y$. (It is just a number, *not* a vector!)

- The following table sums up the 3D equivalents of the 2D results we have already found:

2D	3D
$\mathbf{i}, \mathbf{j}$	$\mathbf{i}, \mathbf{j}, \mathbf{k}$
$\|\mathbf{a}\| = \sqrt{a_x + a_y}$	$\|\mathbf{a}\| = \sqrt{a_x + a_y + a_z}$
$\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y$	$\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z$

  Most of the results from the 2D section (above) still hold true for 3D vectors.

- Obviously in 2D provided lines have different gradient then they *must* intercept somewhere. However in 3D it is possible for two lines to have different direction vectors (i.e. not be parallel) and still not cross: these lines are called *skew*. This example shows how to discover if lines in 3D intercept or are skew.

$$\mathbf{r} = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} \qquad \text{and} \qquad \mathbf{r} = \begin{pmatrix} 0 \\ -6 \\ 5 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

  Firstly we note the different direction vectors, so they cannot be parallel. Equate the x and y components[13] of both lines and solve for λ and μ:

$$\begin{aligned} (x): \quad & 4 + \lambda = 2\mu, \\ (y): \quad & -1 - \lambda = -6 + \mu. \end{aligned}$$

  These solve to $\lambda = 2$ and $\mu = 3$. Put these values back into the original lines and compare z-coordinates: if they are the same then they intercept, if different then skew.

$$\mathbf{r} = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 6 \\ -3 \\ 8 \end{pmatrix} \qquad \text{and} \qquad \mathbf{r} = \begin{pmatrix} 0 \\ -6 \\ 5 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ -3 \\ 8 \end{pmatrix}.$$

  Therefore the lines cross at $(6, -3, 8)$. (You should find that the x and y-coordinates are *always* the same, it is only the z-coordinate that might be different; a nice little check!)

- To find the angle between two lines then *dot their direction vectors*. Using the two lines in the above example we find:

$$\begin{aligned} \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} &= \sqrt{1^2 + 1^2 + 4^2}\sqrt{2^2 + 1^2 + 1^2} \cos\theta \\ 2 - 1 + 4 &= \sqrt{18}\sqrt{6} \cos\theta \\ \theta &= 61.2^\circ \text{ (to 3s.f.)}. \end{aligned}$$

  If you get an answer $90 < \theta \leqslant 180$ then give $180^\circ - \theta$ as your answer (Between any two lines there are two possible angles between them; think about it. The acute angle tends to be 'nicer').

- When working out angles in 3D you must be very careful that you are 'dotting' the right vectors! For example if $A = (1, 2, -2)$, $B = (3, 1, -4)$ and $C = (7, 5, 1)$ find the angle $A\hat{B}C$. Draw a sketch! We require the angle at B so we need to dot $\overrightarrow{BA}$ and $\overrightarrow{BC}$.

  Now $\overrightarrow{BA} = \mathbf{a} - \mathbf{b} = \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}$ and $\overrightarrow{BC} = \mathbf{c} - \mathbf{b} = \begin{pmatrix} 4 \\ 4 \\ 5 \end{pmatrix}$. Therefore dotting we find:

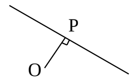
$$\begin{aligned} \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 4 \\ 5 \end{pmatrix} &= \sqrt{9}\sqrt{57} \cos\theta \\ -8 + 4 + 10 &= 3\sqrt{57} \cos\theta \\ \frac{2}{\sqrt{57}} &= \cos\theta \qquad \Rightarrow \qquad \theta = 74.6^\circ \text{(to 3s.f.)}. \end{aligned}$$

[13] You can take any pair of components you like here (x & y, x & z, or y & z) but most students just take x and y.

- Some tough problems involve the use of

$$\text{“two vectors are at perpendicular”} \quad \Leftrightarrow \quad \text{“the dot product is zero”}.$$

For example find the point (P) on the line l: $\mathbf{r} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ closest to the origin.

Firstly draw a sketch of a line running some distance past an origin. 

At the point P the vector $\overrightarrow{OP}$ must be perpendicular to the line. The direction vector is the vector *along* the line l, so we need

$$(\overrightarrow{OP}).\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = 0.$$

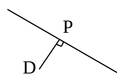
The point P is some point on the line, so $P = \begin{pmatrix} 1+\lambda \\ 2-\lambda \\ 2\lambda \end{pmatrix}$ for some λ. So $\overrightarrow{OP} = \begin{pmatrix} 1+\lambda \\ 2-\lambda \\ 2\lambda \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1+\lambda \\ 2-\lambda \\ 2\lambda \end{pmatrix}$. Therefore

$$\begin{pmatrix} 1+\lambda \\ 2-\lambda \\ 2\lambda \end{pmatrix} . \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = 0.$$

This gives $1+\lambda-2+\lambda+4\lambda = 0$ which solves to $\lambda = \frac{1}{6}$. Putting this λ back into l we find

$$P = (\tfrac{7}{6}, \tfrac{11}{6}, \tfrac{1}{3}).$$

- Another tough example done in two ways: Find the shortest distance from point $D = (2,-1,3)$ to the line $\mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$.

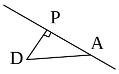
  - **Method I**: First method similar to above. Draw a sketch! 

Let the point on the line closest to D be P. So $P = \begin{pmatrix} 1+2\lambda \\ -\lambda \\ 1+\lambda \end{pmatrix}$. We require $\overrightarrow{DP}$ to be perpendicular to the line if it is the closest point. Thus

$$\overrightarrow{DP} \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = 0 \quad \Rightarrow \quad \left(\begin{pmatrix} 1+2\lambda \\ -\lambda \\ 1+\lambda \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} \right) \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = 0.$$

This solves to $\lambda = \frac{5}{6}$. Therefore $P = (\frac{8}{3}, -\frac{5}{6}, \frac{11}{6})$. Therefore the distance is

$$\text{dist.} = |\overrightarrow{DP}| = |\mathbf{p} - \mathbf{d}| = \sqrt{(2-\tfrac{8}{3})^2 + (\tfrac{5}{6}-1)^2 + (3-\tfrac{11}{6})^2} = \frac{\sqrt{66}}{6}.$$

– **Method II**: Again, draw a sketch. 

Let P be the point closest to D. This time also include the point that we know the line passes through $A = (1, 0, 1)$. We have therefore created a right angled triangle APD with a right angle at P. Length AD is just the magnitude of $\mathbf{d} - \mathbf{a}$; $|\mathbf{d} - \mathbf{a}| = \sqrt{(2-1)^2 + (-1-0)^2 + (3-1)^2} = \sqrt{6}$. Angle $D\hat{A}P$ can be worked out by $\overrightarrow{AD} \cdot$ (direction vector). So

$$(\mathbf{d} - \mathbf{a}) \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = |\mathbf{d} - \mathbf{a}| \left| \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \right| \cos\theta$$

$$\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \sqrt{6}\sqrt{6} \cos D\hat{A}P.$$

So $\cos D\hat{A}P = \frac{5}{6}$.
By right angled trigonometry $\sin D\hat{A}P = \frac{DP}{\sqrt{6}}$. To convert a sin into a cos we use $\sin^2\theta + \cos^2\theta = 1$ which gives $\sin D\hat{A}P = \frac{\sqrt{11}}{6}$. Therefore $DP = \sqrt{6} \times \frac{\sqrt{11}}{6} = \frac{\sqrt{66}}{6}$, just as before.

OCR Further Pure 1 Module Revision Sheet

The FP1 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J.M.S.

Summing Series

- You must know the important results:

$$\sum_{r=1}^{n} 1 = 1 + 1 + 1 + \cdots + 1 = n,$$
$$\sum_{r=1}^{n} r = 1 + 2 + 3 + \cdots + n = \tfrac{1}{2}n(n+1),$$
$$\sum_{r=1}^{n} r^2 = 1^2 + 2^2 + 3^2 + \cdots + n^2 = \tfrac{1}{6}n(n+1)(2n+1),$$
$$\sum_{r=1}^{n} r^3 = 1^3 + 2^3 + 3^3 + \cdots + n^3 = \tfrac{1}{4}n^2(n+1)^2.$$

- Also know the important properties (for constant λ and μ)

$$\sum_{r=1}^{n}(\lambda f(r) \pm \mu g(r)) = \lambda \sum_{r=1}^{n} f(r) \pm \mu \sum_{r=1}^{n} g(r).$$

  However beware of these!

$$\sum_{r=1}^{n}(f(r) \times g(r)) \neq \sum_{r=1}^{n} f(r) \times \sum_{r=1}^{n} g(r) \qquad \text{and} \qquad \sum_{r=1}^{n}\left(\frac{f(r)}{g(r)}\right) \neq \frac{\sum_{r=1}^{n} f(r)}{\sum_{r=1}^{n} g(r)}.$$

  To apply the first of these is equivalent to the heinous crime of $(a+b+c)^2 = a^2+b^2+c^2$!!!

- These must be applied in cases such as:

$$\begin{aligned}\sum_{r=1}^{n}(4r^2 - 2r + 3) &= 4\sum_{r=1}^{n} r^2 - 2\sum_{r=1}^{n} r + 3\sum_{r=1}^{n} 1\\ &= \tfrac{2}{3}n(n+1)(2n+1) - n(n+1) + 3n\\ &= \tfrac{n}{3}[2(n+1)(2n+1) - 3(n+1) + 9]\\ &= \tfrac{n}{3}(4n^2 + 3n + 8).\end{aligned}$$

- If the sum starts from a number other than 1 then you can use the trick (which should be obvious)

$$\sum_{r=a}^{n}(\text{something}) = \sum_{r=1}^{n}(\text{something}) - \sum_{r=1}^{a-1}(\text{something}).$$

- The *method of differences* can be used to sum certain expressions where cancellation occurs when the sum is written out. For example find $\sum_{r=1}^{n}\left(\frac{1}{2r+1}-\frac{1}{2r+3}\right)$. Write the sum out, starting a new line for each value of r and you should see that some nice cancelling occurs;

$$\begin{aligned}\sum_{r=1}^{n}\left(\frac{1}{2r+1}-\frac{1}{2r+3}\right) &= \frac{1}{3}-\frac{1}{5}\\ &+\frac{1}{5}-\frac{1}{7}\\ &+\frac{1}{7}-\frac{1}{9}\\ &\quad\vdots\\ &+\frac{1}{2n+1}-\frac{1}{2n+3}.\end{aligned}$$

You can see that everything cancels except the $\frac{1}{3}$ and the $\frac{1}{2n+3}$ so

$$\sum_{r=1}^{n}\left(\frac{1}{2r+1}-\frac{1}{2r+3}\right)=\frac{1}{3}-\frac{1}{2n+3}.$$

It is usually best *not* to combine these terms together into one fraction in order to make it easier to see if there is a sum to infinity.

- A sum to infinity exists if the expression for the sum to n has a finite limit as $n\to\infty$. In the above example it does, so

$$\sum_{r=1}^{\infty}\left(\frac{1}{2r+1}-\frac{1}{2r+3}\right)=\lim_{n\to\infty}\left(\frac{1}{3}-\frac{1}{2n+3}\right)=\frac{1}{3}.$$

- Questions of this sort invariably start "Show that $f(r)-g(r)=h(r)$", and then ask you to sum $h(r)$; this, clearly, is the same as summing $f(r)-g(r)\Rightarrow$ 'method of differences'.

Matrices

- Capital letters tend to be used to denote matrices and you should underline them, just as you do with vectors. An $n\times m$ matrix has n rows and m columns. So $\begin{pmatrix}1&2&-3\\2&-2&7\end{pmatrix}$ is a 2×3 matrix. You must be able to add, subtract and multiply matrices. To add or subtract matrices they must be the same size and it works as you would expect. To multiply matrices ($\mathbf{A}\times\mathbf{B}$, say) the number of columns of $\mathbf{A}$ must be the same as the number of rows of $\mathbf{B}$. Your teacher will have explained this better than I ever can here, but a few examples: test for yourself!

$$\begin{pmatrix}1&5\\-2&1\end{pmatrix}\begin{pmatrix}2&7&-1\\0&1&4\end{pmatrix}=\begin{pmatrix}2&12&19\\-4&-13&6\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}2&3&-1\end{pmatrix}\begin{pmatrix}1\\0\\4\end{pmatrix}=\begin{pmatrix}-2\end{pmatrix},$$

$$\begin{pmatrix}1\\0\\4\end{pmatrix}\begin{pmatrix}2&3&-1\end{pmatrix}=\begin{pmatrix}2&3&-1\\0&0&0\\8&12&-4\end{pmatrix}.$$

- The determinant of an $n\times n$ ('*square*') matrix can be denoted by the letter Δ. A matrix with $\Delta=0$ is called a '*singular*' matrix; otherwise it is '*non-singular*'. For a 2×2 matrix $\det\begin{pmatrix}a&b\\c&d\end{pmatrix}\equiv\Delta=ad-bc.$

- The inverse of a 2×2 matrix is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

- The inverse of a matrix (if it exists) is such that $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ where $\mathbf{I}$ is the identity matrix $\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

- Matrix multiplication is not, in general, commutative; i.e. $\mathbf{AB} \neq \mathbf{BA}$. Matrix multiplication is, however, associative; i.e. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.

- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. An extension of this is $(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3 \ldots \mathbf{A}_n)^{-1} = \mathbf{A}_n^{-1} \ldots \mathbf{A}_3^{-1}\mathbf{A}_2^{-1}\mathbf{A}_1^{-1}$; prove it by induction yourself if you fancy...

- You must be very careful when manipulating matrix equations because of this non-commutativity. With normal numbers we are happy with $ax = b$ giving $x = ba^{-1}$, but this is wrong in matrix-world.

$$\begin{aligned} \mathbf{AX} &= \mathbf{B} \\ \mathbf{A}^{-1}\mathbf{AX} &= \mathbf{A}^{-1}\mathbf{B} \quad \underline{\text{pre}}\text{-multiply both sides by } \mathbf{A}^{-1} \\ \mathbf{IX} &= \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{X} &= \mathbf{A}^{-1}\mathbf{B}. \end{aligned}$$

  Or

$$\begin{aligned} \mathbf{XA} &= \mathbf{B} \\ \mathbf{XAA}^{-1} &= \mathbf{BA}^{-1} \quad \underline{\text{post}}\text{-multiply both sides by } \mathbf{A}^{-1} \\ \mathbf{XI} &= \mathbf{BA}^{-1} \\ \mathbf{X} &= \mathbf{BA}^{-1}. \end{aligned}$$

- Know that linear simultaneous equations can be expressed by matrices:

$$\begin{aligned} ax + by &= c \\ dx + ey &= f \end{aligned} \quad \Rightarrow \quad \begin{pmatrix} a & b \\ d & e \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c \\ f \end{pmatrix}.$$

  The system of equations has a unique solution provided $\Delta \neq 0$. If $\Delta = 0$ then there are no unique solutions: there are either an infinite set of solutions or no solutions at all depending on whether $ax + by = c$ and $dx + ey = f$ represent parallel lines (no solutions) or the same line (infinite set of solutions).

  The unique solution (if it exists) is given by $\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{ae-bd}\begin{pmatrix} e & -b \\ -d & a \end{pmatrix}\begin{pmatrix} c \\ f \end{pmatrix}$.

Matrix Transformations

- Matrices can be thought of as transformations. To discover what a matrix does, consider what it does to the arbitrary point (x, y) and *think*! For example

  1. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ so $(x, y) \to (x, y)$. Therefore matrix does nothing.
  2. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$ so $(x, y) \to (y, x)$. Therefore the x and y-coordinates get flipped, so matrix reflected in the line $y = x$.

3. $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}$ so $(x, y) \to (-x, y)$. Therefore matrix changes the sign of the x-coordinate, so it represents a reflection in the y-axis.
4. $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix}$ so $(x, y) \to (y, -x)$. Draw a few sample points and we see it represents a rotation 90° clockwise about the origin.
5. $\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4x \\ 4y \end{pmatrix}$ so $(x, y) \to (4x, 4y)$. So the x and y-coordinates get multiplied by 4. Therefore an enlargement scale factor 4, centre the origin.

- You need to know the family of matrices that represent shears.

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} = \text{"Shear with } x\text{-axis invariant with shear factor } k\text{".} \quad \overrightarrow{\leftarrow}$$

$$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix} = \text{"Shear with } y\text{-axis invariant with shear factor } k\text{".} \quad \downarrow\uparrow$$

- For combined transformations you write the matrices in the opposite order to which the transformations occur[14]. For example if we apply transformation **A** followed by transformation **B**, then the matrix for this combined transformation would be **BA**.

- If a 2×2 matrix **M** represents a transformation, then $|\det(\mathbf{M})|$ represents the *area scale factor* of the transformation.

  For example $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ represents an enlargement with length scale factor 2. We see the determinant is 4, so areas get multiplied by 4 in the transformation, which is consistent.

- If we consider an arbitrary matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ acting on the point $(1, 0)$ we find

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}$$

  (i.e. the first column of the matrix). Similarly if we act on the point $(0, 1)$ we find

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}$$

  (i.e. the second column of the matrix). This immensely powerful pair of statements tells us that *if* a transformation can be expressed by a matrix, then all we need to do to find the matrix that does what we want is to find where $(1, 0)$ maps to under the transformation and write this image point as the first column of our matrix and find where $(0, 1)$ maps to under the transformation and write this as the second column.

- For example find the matrix that:
  - reflects in the x-axis. $(1, 0) \to (1, 0)$ and $(0, 1) \to (0, -1)$ $\Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
  - reflects in the $y = x$. $(1, 0) \to (0, 1)$ and $(0, 1) \to (1, 0)$ $\Rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
  - rotates 90° clockwise. $(1, 0) \to (0, -1)$ and $(0, 1) \to (1, 0)$ $\Rightarrow \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

[14] Just like functions: If you apply f then g, we do $gf(x)$.

– enlarges scale factor 3. $(1,0) \to (3,0)$ and $(0,1) \to (0,3)$ $\Rightarrow \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$.

– stretch factor 2 parallel to y-axis. $(1,0) \to (1,0)$ and $(0,1) \to (0,2)$ $\Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.

– rotates $\theta°$ anticlockwise. $(1,0) \to (\cos\theta, \sin\theta)$ and $(0,1) \to (-\sin\theta, \cos\theta)$

$$\Rightarrow \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

– rotates 90° CW and then reflects in y-axis. $(1,0) \to (0,-1)$ and $(0,1) \to (1,0)$

$$\Rightarrow \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

3 × 3 Matrices

- To calculate the determinant of a 3×3 matrix, you pick a column or a row (most students choose the first column, but it works with any row or column) and you work down/across it using the plus/minus checkerboard approach and multiplying by the determinant of the 2×2 matrix left when the column and row of the number you have chosen is crossed out.

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \Delta = a(ei - fh) - d(bi - ch) + g(bf - ce).$$

- To invert a 3×3 matrix you do:

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}^{-1} = \frac{1}{\Delta} \begin{pmatrix} +(ei-hf) & -(di-gf) & +(dh-eg) \\ -(bi-ch) & +(ai-cg) & -(ah-bg) \\ +(bf-ce) & -(af-cd) & +(ae-bd) \end{pmatrix}^{\mathrm{T}},$$
$$= \frac{1}{\Delta} \begin{pmatrix} +(ei-hf) & -(bi-ch) & +(bf-ce) \\ -(di-gf) & +(ai-cg) & -(af-cd) \\ +(dh-eg) & -(ah-bg) & +(ae-bd) \end{pmatrix}.$$

  Don't forget to transpose at the end! There is an elegant pattern to all of the above; it's easy to do once you get into the swing of it. (Mr Stone has a spreadsheet where you can practice this to your heart's content.)

- As before, a system of linear simultaneous equations can be written with a matrix.

$$\begin{aligned} ax + by + cz &= d \\ ex + fy + gz &= h \\ ix + jy + kz &= l \end{aligned} \quad \Rightarrow \quad \begin{pmatrix} a & b & c \\ e & f & g \\ i & j & k \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d \\ h \\ l \end{pmatrix}.$$

  If the matrix in non-singular then the equations have a unique solution. If the matrix is singular then the system either has an infinite set of solutions or no solutions at all. If the equations generate an inconsistency (e.g. 4=19) then there are no solutions at all.

- If the matrix is non-singular, then the unique solution is given by:

$$\mathbf{M} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{M}^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Complex Numbers

- Complex numbers start with one idea only; that we can find a number that squares to -1; we call it i. Therefore $i^2 = -1$. It is not a number that exists on the number line so it is refered to as *complex* or *imaginary*. Therefore the square root of any negative number can now be calculated; $\sqrt{-36} = \sqrt{36}\sqrt{-1} = 6i$.

- In general a complex number can consist of a real part and an imaginary/complex part i.e. $a + ib$, where a is the real part and b is the complex part. We write $\text{Re}(a+ib) = a$ and $\text{Im}(a + ib) = b$. It is important to note that a and b themselves *must* be real numbers.

- We can use complex numbers to solve *any* quadratic equation. For example solve $3x^2 + 2x + 7 = 0$ by the formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-2 \pm \sqrt{2^2 - 4 \times 3 \times 7}}{2 \times 3} = -\frac{1}{3} \pm i\frac{2\sqrt{5}}{3}.$$

- A complex (or real) number can be represented as a point in an *Argand* diagram. So the complex number $6 + 2i$ would be the point 6 across and 2 up, at the equivalent point where $(6, 2)$ would be in a Cartesian coordinate system.

- The complex conjugate $(z*)$ of a complex number (z) is where the complex part has the sign changed. For example if $z = 3 - 7i$, then $z* = 3 + 7i$. Real numbers are, therefore, their own conjugates. Any number with an imaginary component is reflected in the real axis in the Argand diagram.

- If two complex numbers are equal, then the real parts must be equal and the complex parts must be equal: i.e.

$$\begin{aligned} z_1 = z_2 \quad &\Rightarrow \quad \text{Re}(z_1) = \text{Re}(z_2) \quad \text{and} \quad \text{Im}(z_1) = \text{Im}(z_2), \\ a + ib = c + id \quad &\Rightarrow \quad a = c \quad \text{and} \quad b = d. \end{aligned}$$

- To add, subtract or multiply complex numbers the results are pretty obvious:

$$\begin{aligned} (a + bi) + (c + id) &= (a + c) + i(b + d) \\ (a + bi) - (c + id) &= (a - c) + i(b - d) \\ (a + bi)(c + id) &= ac + adi + bci + bdi^2 = (ac - bd) + i(ad + bc) \end{aligned}$$

- To divide by a complex number use a trick taken from surds; in C1 if the bottom line was $a \pm b\sqrt{k}$ then you multiplied top and bottom by $a \mp b\sqrt{k}$. In FP1 if you want to divide by $a \pm ib$, then you multiply top and bottom by the complex conjugate $a \mp ib$. For example:

$$\frac{3 - 2i}{2 + 5i} = \frac{3 - 2i}{2 + 5i} \times \frac{2 - 5i}{2 - 5i} = \frac{6 + 10i^2 - 4i - 15i}{4 - 25i^2 + 10i - 10i} = \frac{-4 - 19i}{29}.$$

- Complex numbers exhibit the elegant property of *closure*[15]. This means that any operation on complex numbers involving $+, -, \times, \div, \sqrt{\ldots}, \sqrt[n]{\ldots}$ etc. will produce an answer that is also complex[16]. This allows us to state that the answer to a given problem *must* be $a + ib$ for some a and b and then proceed to calculate a and b by equating the real part and, separately, the imaginary part.

[15] "And that, my friend, is what they call closure" - *Rachel Green, Friends*
[16] Note this does not happen with the real numbers; you cannot always square root a number.

- In the above problem to find $\frac{3-2i}{2+5i}$ we could also have approached it by stating that the answer is $a+ib$ and manipulating:

$$\begin{aligned}\frac{3-2i}{2+5i} &= a+ib\\ 3-2i &= (a+ib)(2+5i)\\ 3-2i &= (2a-5b)+i(5a+2b).\end{aligned}$$

This yields the simultaneous equations $3 = 2a-5b$ and $-2 = 5a+2b$. These solve to $a = -\frac{4}{29}$ and $b = -\frac{19}{29} \Rightarrow \frac{-4-19i}{29}$, just as before. I wouldn't use this method in this case but I would certainly use it...

- ...to find square roots. The square roots of 16 are (obviously) ± 4. With the exception of zero, we should expect two roots and the same is true of complex numbers. For example find the square roots of $8-6i$: We know that the answers must be of the form $a+ib$ such that

$$\begin{aligned}8-6i &= (a+ib)^2\\ 8-6i &= (a^2-b^2)+(2ab)i\\ \text{Therefore, } 8 = a^2-b^2 &\text{ and } -6 = 2ab.\end{aligned}$$

From the second we find $b = -\frac{3}{a}$. Putting this in the first we find $0 = a^4-8a^2-9 = (a^2-9)(a^2+1)$. The first bracket yields $a = \pm 3$. (The second bracket yields $a = \pm i$, but we can discard this because a must be real.) Therefore this yields the square roots $3-i$ and $-3+i$. In the Argand diagram you should find that square roots come out in opposite directions from the origin.

- If a polynomial has *real coefficients* then its roots are either real, or exist in complex conjugate pairs. Therefore if $z = a \pm ib$ is a root, then so is $z = a \mp ib$.

For example, given that $z^4 - z^3 + 2z^2 + 7z - 5 = 0$ has one root $1-2i$, solve the equation fully. Since the coefficients are real we know that the conjugate $1+2i$ must also be a root. Therefore $(z-(1-2i))$ and $(z-(1+2i))$ must be factors by the factor theorem. Multiplying out the two factors we find $(z-(1-2i))(z-(1+2i)) = (z^2-2z+5)$ which must also be a factor. By polynomial division we find

$$z^4 - z^3 + 2z^2 + 7z - 5 = (z^2-2z+5)(z^2+z-1) = 0.$$

The second quadratic solves to $z = -\frac{1}{2} \pm \frac{\sqrt{5}}{2}$. Therefore the solutions are

$$z = 1-2i, \quad z = 1+2i, \quad z = -\tfrac{1}{2}+\tfrac{\sqrt{5}}{2}, \quad z = -\tfrac{1}{2}-\tfrac{\sqrt{5}}{2}.$$

- The modulus of a complex number ($z = x+iy$) is defined $|z| = \sqrt{x^2+y^2}$. It represents the distance of a complex number from the origin. For example the modulus of $z = 2-2\sqrt{3}i$ would be $|z| = \sqrt{2^2+(-2\sqrt{3})^2} = 4$.

- The argument of a complex number is defined as the angle a line from the origin to a complex number makes with the positive real axis. By convention $-\pi < \arg(z) \leqslant \pi$. For example

$$\begin{aligned}
\arg(4) &= 0,\\
\arg(i) &= \frac{\pi}{2},\\
\arg(-3) &= \pi,\\
\arg(1+i) &= \frac{\pi}{4},\\
\arg(-1-i) &= -\frac{3\pi}{4}.
\end{aligned}$$

Arguments are best calculated by drawing a suitable right angled triangle in an Argand diagram and then calculating the desired angle (not always an angle in the triangle you've drawn, but π minus it, etc.)

- You must be able to sketch loci of points obeying a rule defined by a modulus or an argument. The most important fact here is often that the operation *subtraction* takes you *from* one complex number *to* another[17]; i.e. $z - w$ takes you *from* w *to* z. An addition can be converted into a subtraction by $z + w = z - (-w)$; this therefore represents the movement from $-w$ to z.
- This idea allows us to draw certain loci very easily indeed.

  For example: $|z| = 4$ means the length of z from the origin is 4; i.e. a circle of radius 4, centre the origin.

  For example: $|z - 2| = 5$ means the length travelling from 2 to z is 5, so a circle radius 5, centre 2(+0i).

  For example: $|z + i| < 2$ is the same as $|z - (-i)| < 2$ which means the length travelling from $-i$ to z is less than 2, so the inside of a circle radius 2, centre $(0) - i$.

  For example: $|z| = |z + 1 - i|$ is the same as $|z| = |z - (-1 + i)|$ which means the length travelling from 0 to z must be the same as the distance travelling from $-1 + i$ to z so it must be the perpendicular bisector of 0 and $-1 + i$, i.e. $y = x + 1$.
- The above type of questions can also be done using a method in your textbook (see top of P144, "Method 2"), but I prefer the 'intuitive' way demonstrated above.
- Argument loci also come up and we can use the same principles.

  For example: $\arg(z) = \frac{\pi}{2}$ means the argument z makes is $\frac{\pi}{2}$ so it is a vertical line going up from the origin (with a hollow circle drawn at the origin to indicate that it is not included in the half-line).

  For example: $\arg(z - i) = \frac{\pi}{6}$ means the argument going from i to z is $\frac{\pi}{6}$ so it is a half-line from i (hollow circle) at angle $\frac{\pi}{6}$ with the positive real axis.

Roots Of Equations

- By considering the general quadratic equation $ax^2 + bx + c = 0$ we re-write it as $x^2 + \frac{b}{a}x + \frac{c}{a} = 0$. Quadratics can be factorised into two linear factors $(x - \alpha)(x - \beta)$. By equating the two we find

$$\begin{aligned}
(x-\alpha)(x-\beta) &= x^2 + \frac{b}{a}x + \frac{c}{a}\\
x^2 - (\alpha+\beta)x + \alpha\beta &= x^2 + \frac{b}{a}x + \frac{c}{a}.
\end{aligned}$$

[17] In precisely the same way that with vectors $\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$.

So we see that the sum of the roots of a quadratic is $-\frac{b}{a}$ and the product of the roots is $\frac{c}{a}$.

- By two tedious derivations (that you *should* do for yourself) similar to the one above we find that for the cubic ($ax^3+bx^2+cx+d=0$) and the quartic ($ax^4+bx^3+cx^2+dx+e=0$) the following:

QUADRATICS	CUBICS	QUARTICS
$\alpha+\beta=-\frac{b}{a}$	$\alpha+\beta+\gamma=-\frac{b}{a}$	$\alpha+\beta+\gamma+\delta=-\frac{b}{a}$
$\alpha\beta=\frac{c}{a},$	$\alpha\beta+\alpha\gamma+\beta\gamma=\frac{c}{a}$	$\alpha\beta+\alpha\gamma+\alpha\delta+\beta\gamma+\beta\delta+\gamma\delta=\frac{c}{a}$
	$\alpha\beta\gamma=-\frac{d}{a},$	$\alpha\beta\gamma+\alpha\beta\delta+\alpha\gamma\delta+\beta\gamma\delta=-\frac{d}{a}$
		$\alpha\beta\gamma\delta=\frac{e}{a}.$

For your exam you only need quadratics and cubics, but the pattern continues fairly easily to quartics, quintics and beyond.

- For speed of writing we use the following shorthand:

$$\alpha+\beta+\gamma\equiv\sum\alpha \qquad \text{and} \qquad \alpha\beta+\alpha\gamma+\beta\gamma\equiv\sum\alpha\beta.$$

- We can therefore find properties of roots from equations without having to solve the equations themselves.

  For example from $2x^2-3x-6=0$ I can say that $\alpha\beta=\frac{-6}{2}=-3$ and $\alpha+\beta=\frac{3}{2}$.

  For example from $2x^3-4x^2-3x+6=0$ I can say that $\alpha\beta\gamma=\frac{-6}{2}=-3$, $\sum\alpha\beta=-\frac{3}{2}$ and $\sum\alpha=\frac{4}{2}=2$. Watch those signs!

- You are often asked to construct new equations with roots related to the original equation's roots. There are two basic methods for this:

- Method I (the "Lo-Tech" approach) is to find the new 'sum'/'sum of prod'/'prod' of roots etc., from knowledge of the old 'sum'/'sum of prod'/'prod'. Two examples:

  1. The equation $2x^2+5x+7=0$ has roots α and β. Find an equation with roots $2\alpha-1$ and $2\beta-1$. We can see that $\alpha\beta=\frac{7}{2}$ and $\alpha+\beta=-\frac{5}{2}$. Therefore for the new equation must have the following:
     New sum of roots $=(2\alpha-1)+(2\beta-1)=2(\alpha+\beta)-2=2\times(-\frac{5}{2})-2=-7$.
     New prod of roots $=(2\alpha-1)(2\beta-1)=4\alpha\beta-2(\alpha+\beta)+1=4\times\frac{7}{2}-2\times(-\frac{5}{2})+1=20$.
     Therefore the new equation is $u^2+7u+20=0$.
  2. The equation $2x^3-x^2+4x+2=0$ has roots α, β and γ. Find an equation with roots $\alpha+1$, $\beta+1$ and $\gamma+1$. We can see that $\alpha\beta\gamma=-1$, $\sum\alpha\beta=2$ and $\sum\alpha=\frac{1}{2}$. Therefore for the new equation we must have the following:
     New sum of roots $=(\alpha+1)+(\beta+1)+(\gamma+1)=\sum\alpha+3=\frac{7}{2}$.
     New sum of prods $=(\alpha+1)(\beta+1)+(\alpha+1)(\gamma+1)+(\beta+1)(\gamma+1)=\sum\alpha\beta+2\sum\alpha+3=2+2\times\frac{1}{2}+3=6$.
     New prod of roots $=(\alpha+1)(\beta+1)(\gamma+1)=\alpha\beta\gamma+\sum\alpha\beta+\sum\alpha+1=-1+2+\frac{1}{2}+1=\frac{5}{2}$.
     Therefore the equations becomes $u^3-\frac{7}{2}u^2+6u-\frac{5}{2}=0$ which we double to make it 'nice':

$$2u^3-7u^2+12u-5=0.$$

- Method II (the "Hi-Tech" approach) is to make a substitution into the original equation to construct a second. We will do the same two examples as above.

  1. The equation $2x^2+5x+7=0$ has roots α and β. Find an equation with roots $2\alpha-1$ and $2\beta-1$. Let u be one of the new roots; $u=2\alpha-1$. Rearrange to make α the subject; $\alpha=\frac{u+1}{2}$. We know that α satisfies the original equation because it is a root, so if we substitute $\alpha=\frac{u+1}{2}$ into the original equation we will have an equation in u which has the desired roots.
  $$2\left(\frac{u+1}{2}\right)^2+5\left(\frac{u+1}{2}\right)+7=0 \quad\Rightarrow\quad u^2+7u+20=0.$$
  2. The equation $2x^3-x^2+4x+2=0$ has roots α, β and γ. Find an equation with roots $\alpha+1$, $\beta+1$ and $\gamma+1$. So, let $u=\alpha+1$. Therefore $\alpha=u-1$. Sub in we find:
  $$2(u-1)^3-(u-1)^2+4(u-1)+2=0 \quad\Rightarrow\quad 2u^3-7u^2+12u-5=0.$$

Proof by Induction

- Let $P(n)$ be a proposition which depends on some integer value n. The principle of induction works as follows: Start by demonstrating the truth of $P(1)$ (say). Then we show that <u>*if*</u> $P(k)$ is true for *some value* k <u>*then*</u> it implies the truth of $P(k+1)$, then $P(n)$ must be true for all integer $n \geqslant 1$. This is because we have shown
$$P(1)\Rightarrow P(2),\text{ and } P(2)\Rightarrow P(3),\text{ and } P(3)\Rightarrow P(4) \text{ etc. etc. etc.}$$
- Your answer should always follow this template:
  - "Let $P(n)$ be the proposition that $f(n)=g(n)$ for all $n \geqslant 1$."
  - "Basis Case: If $n=1$, $f(1)=\dots$ and $g(1)=\dots$. We see $f(1)=g(1)$ so $P(1)$ is true."
  - "Let us suppose that $P(n)$ is true for some $n=k$:
  $$f(k)=g(k)."$$
  [Then manipulate $f(k)=g(k)$ using algebra to obtain the next line]
  $$"f(k+1)=g(k+1)."$$
  - "This is the statement of $P(k+1)$."
  - "Therefore we have shown that *if* $P(k)$ is true *then* $P(k+1)$ is also true and since $P(1)$ is also true we can conclude by the principle of mathematical induction that $P(n)$ is true for all $n \geqslant 1$."
- If an induction question includes a "$\sum$", can I suggest you get rid of it by writing out the sum term-by-term; students tend to get muddled on when to use r, n, k and $k+1$ in my experience (although this might be my teaching). Also leave initial numerical values unevaluated; 1×2^2 is preferable to 4. For example
$$\sum_{r=1}^{n} r(r+2) \Rightarrow 1\times 3+2\times 4+\cdots+n(n+2).$$

- For example use induction to prove that for $n \geqslant 2$, $\displaystyle\sum_{r=2}^{n}(r-1)r = \tfrac{1}{3}n(n-1)(n+1)$.

  - "The question is the same as proving $1\times2+2\times3+\cdots+(n-1)n = \tfrac{1}{3}n(n-1)(n+1)$."
  - "Let $P(n)$ be the proposition $1\times2+2\times3+\cdots+(n-1)n = \tfrac{1}{3}n(n-1)(n+1)$."
  - "Basis case: If $n=2$, $1\times2+2\times3+\cdots+(n-1)n = 2$ and $\tfrac{1}{3}n(n-1)(n+1) = 2$. We see that LHS = RHS = 2 so $P(2)$ is true."
  - "Let us suppose that $P(n)$ is true for some $n=k$:

$$1\times2+2\times3+\cdots+(k-1)k = \frac{1}{3}k(k-1)(k+1)."$$

  - (Add the next term to the LHS to both sides:)

$$\text{“}1\times2+2\times3+\cdots+(k-1)k+\underline{k(k+1)} = \frac{1}{3}k(k-1)(k+1)+\underline{k(k+1)}$$
$$= \frac{1}{3}k(k+1)[(k-1)+3]$$
$$= \frac{1}{3}k(k+1)(k+2)."$$

  - "This is the statement of $P(k+1)$."
  - "Therefore we have shown that if $P(k)$ is true then $P(k+1)$ is also true and since $P(2)$ is also true we can conclude by the principle of mathematical induction that $P(n)$ is true for all $n \geqslant 2$."

- In a recent official mark scheme, the use of the words 'mathematical induction' in your conclusion was needed for full marks.

OCR Further Pure 2 Module Revision Sheet

The FP2 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J.M.S.

Rational Functions

- Review all partial fractions and polynomial division work from C4 before starting this section.
- A nice 'trick' that can be used from time-to-time is to make the top line of an algebraic fraction look like a multiple of the bottom. Then you can split it up. For example

$$\frac{x-1}{x+3} = \frac{x+3-4}{x+3} = 1 - \frac{4}{x+3},$$
$$\frac{2x^2-1}{x-2} = \frac{2x^2-4x+4x-1}{x-2} = \frac{2x(x-2)+4x-1}{x-2} = 2x + \frac{4x-8+7}{x-2} = 2x+4+\frac{7}{x-2}.$$

  Some students like this & others don't; it's up to you if you use it. You could just use polynomial division.

- For some reason best known to the examiners at OCR, C4 only contains two of the three partial fraction types[18]. In C4 you dealt with $\frac{ax+b}{(cx+d)(ex+f)}$ and $\frac{ax^2+bx+c}{(dx+e)(fx+g)^2}$. In FP2 you also need to know how do deal with $\frac{ax^2+bx+c}{(dx+e)(fx^2+g)}$. The general technique is

$$\frac{ax^2+bx+c}{(dx+e)(fx^2+g)} \equiv \frac{A}{dx+e} + \frac{Bx+C}{fx^2+g}.$$

  Remember that to use 'pure' partial fractions the numerator has to have order less than the denominator.

- For example to express $\frac{5x^2-7x+14}{(x-3)(2x^2+1)}$ in partial fractions we start:

$$\frac{5x^2-7x+14}{(x-3)(2x^2+1)} \equiv \frac{A}{x-3} + \frac{Bx+C}{2x^2+1},$$
$$\Rightarrow \qquad 5x^2-7x+14 \equiv (2x^2+1)A + (x-3)(Bx+C).$$

  Clearly a good x-value to use is $x=3$, so

$$x=3 \qquad \Rightarrow \qquad 45-21+14 \equiv 19A \qquad \Rightarrow \qquad \underline{A=2}.$$

  We've no more cunning values so just use $x=0$ to find $14 = 2-3C$, which gives $\underline{C=-4}$.

[18]Note that the wonderful MEI has all three in C4 which is much more coherent...

Next use $x = 1$ to give $12 = 6 + (-2)(B - 4)$, which solves to $\underline{B = 1}$. Therefore

$$\frac{5x^2 - 7x + 14}{(x-3)(2x^2+1)} \equiv \frac{2}{x-3} + \frac{x-4}{2x^2+1}.$$

- You can always make the leap from polynomial division to partial fractions in one go if you like. For example to divide

$$\frac{3x^4 + 13x^3 + 27x^2 + 56x + 59}{x^3 + 3x^2 + 4x + 12}$$

we have '$\frac{\text{quartic}}{\text{cubic}}$'. We are therefore expecting 'linear + $\frac{\text{quadratic}}{\text{cubic}}$'. But, because the denominator can be factorised to $(x + 3)(x^2 + 4)$, we could split the '$\frac{\text{quadratic}}{\text{cubic}}$' term into partial fractions too. So

$$\frac{3x^4 + 13x^3 + 27x^2 + 56x + 59}{(x+3)(x^2+4)} \equiv Ax + B + \frac{C}{x+3} + \frac{Dx+E}{x^2+4},$$
$$3x^4 + 13x^3 + 27x^2 + 56x + 59 \equiv (Ax + B)(x + 3)(x^2 + 4) + C(x^2 + 4) + (Dx + E)(x + 3),$$

Clearly $A = 3$ by considering the x^4 coefficient. Running through the rest of the calculations in the usual way (you should do this yourself) we find

$$\frac{3x^4 + 13x^3 + 27x^2 + 56x + 59}{(x+3)(x^2+4)} \equiv 3x + 4 + \frac{2}{x+3} + \frac{x+1}{x^2+4}.$$

Graphs

- To sketch a graph of $y = \frac{f(x)}{g(x)}$ there are a series of steps to follow. If one step contradicts another, chances are you've made a mistake. Firstly to find where a curve crosses the x-axis, set $y = 0$ and solve. Similarly to find where a curve crosses the y-axis, set $x = 0$ and solve.

- To find stationary points just solve $\frac{dy}{dx} = 0$ as usual. To discover their nature you can use the second derivative as normal; or the lo-tech approach. Review your C1 notes.

- To find the vertical asymptotes of the curve $y = \frac{f(x)}{g(x)}$ you need to find where $g(x) = 0$. So $y = \frac{x+3}{(2x-1)(x+2)}$ will have vertical asymptotes $x = -2$ and $x = \frac{1}{2}$.

- To find a horizontal asymptotes of $y = \frac{f(x)}{g(x)}$ you must look at the order of $f(x)$ and the order of $g(x)$.

   1. If "order of $f(x)$" $<$ "order of $g(x)$" then, as $x \to \pm\infty$, $g(x)$ is much, *much* larger than $f(x)$, so $y = 0$ is the horizontal asymptote.
   2. If "order of $f(x)$" $=$ "order of $g(x)$" then, as $x \to \pm\infty$, the dominant term of $f(x)$ and $g(x)$ becomes the largest power of x. Therefore the horizontal asymptote becomes the ratio of the coefficients of the largest power of x. For example $y = \frac{7x^2+2x-1}{5x^2-x-2}$ has $y = \frac{7}{5}$ as its horizontal asymptote.

      (Another, possibly better, way of thinking about this to divide the improper fraction into quotient and remainder by polynomial division. Since the order of the numerator is the same as the order of the denominator, then the quotient is a constant. This constant is the value of the horizontal asymptote.)

3. If "order of $f(x)$" $>$ "order of $g(x)$" then there is *no* horizontal asymptote because as $x \to \pm\infty$, $f(x)$ is much, *much* larger than $g(x)$, so y is unbounded, heading up to $+\infty$ or down to $-\infty$ (just think about what happens when x gets really big).

- If the numerator has order one more than the denominator, there will exist an *oblique asymptote*; a line of the form $y = mx + c$ that the curve approaches when $x \to \pm\infty$. To find this line, you must carry out the polynomial division and find the quotient and remainder. For example $y = \frac{3x^2+4x+5}{x+1}$. We know $y = \frac{3x^2+4x+5}{x+1} = Ax + B + \frac{C}{x+1}$ and, carrying out the calculation (do it yourself!), we find $y = 3x + 1 + \frac{4}{x+1}$. Therefore the oblique asymptote is $y = 3x + 1$ because the $\frac{4}{x+1} \to 0$ as $x \to \pm\infty$.

  Similarly you should find (again, do it yourself!) $y = \frac{-x^3+x+2}{x^2+x+1} = -x + 1 + \frac{x+1}{x^2+x+1}$, so the oblique asymptote is $y = -x + 1$.

- You must be able to discover the range of y-values for which the curve exists, and, equivalently, the values for which it does not. This can be done by finding the stationary points on the curve and considering a sketch. However, there is quite a neat algebraic method. For example: Find the values of y for which the curve

$$y = \frac{x^2 + x + 1}{x^2 + 1}$$

  exists. Multiplying by the denominator we discover $y(x^2 + 1) = x^2 + x + 1$. This can be rearranged as a quadratic in x: $(y - 1)x^2 - x + (y - 1) = 0$. For the curve to exist, we need the quadratic to have at least one solution, so $b^2 - 4ac \geqslant 0$. So

$$\begin{aligned} (-1)^2 - 4(y-1)(y-1) &\geqslant 0, \\ 4y^2 - 8y + 3 &\leqslant 0, \\ (2y-3)(2y-1) &\leqslant 0. \end{aligned}$$

  This quadratic inequality solves to $\frac{1}{2} \leqslant y \leqslant \frac{3}{2}$. So the curve only exists between the horizontal lines $y = \frac{1}{2}$ and $y = \frac{3}{2}$.

- Given a graph of $y = f(x)$, you must also be able to sketch the graph of $y^2 = f(x)$. Most students (including myself) mentally re-cast the problem as drawing $y = \pm\sqrt{f(x)}$. Things to look for include

  1. Anything below the x-axis on the original graph (y) is negative and therefore *cannot* be square rooted. Therefore these x-values represent a forbidden region where y^2 doesn't exist.
  2. All y-values above the x-axis get square rooted. Then these new points also get reflected in the x-axis. *Any* graph $y^2 = f(x)$, *must* have a line of symmetry in the x-axis.
  3. All positive y-values on the original graph get square rooted; therefore points on the line $y = 1$ are invariant. If $y > 1$, then they get 'pulled down' towards the x-axis ($\sqrt{100} = 10$). If $y < 1$ then the points get pushed further away from the x-axis ($\sqrt{\frac{1}{4}} = \frac{1}{2}$).
  4. Vertical asymptotes on y remain vertical asymptotes on y^2.
  5. Horizontal asymptotes above the x-axis ($y = k$, say) become horizontal asymptotes $y = \pm\sqrt{k}$.
  6. Any points where the original curve hits the x-axis are also invariant ($\sqrt{0} = 0$). Also, the gradient of any points where y hits the x-axis become vertical on the y^2 graph.
  7. Any stationary point above the x-axis on y ($(2, 16)$, say) remain stationary points at $(2, \pm 4)$, say.

Polar Coordinates

- Polar coordinates are given as points with (r, θ) with the constraints $r \geqslant 0$ and (usually) either $0 \leqslant \theta < 2\pi$ or $-\pi < \theta \leqslant \pi$. The distance from the origin is r and θ is the angle made with the initial line (i.e. the positive x-axis) measured anti-clockwise. The angle constraints are used so that each point in space has a unique angle[19]. The 'pole' is sometimes used to describe the origin of your xy-grid in the context of a polar graph.

- Circles are described by '$r =$ constant'. Lines running out from the pole are described by '$\theta =$ constant'.

- Polar curves are a new way of describing curves by showing a relationship between r and θ. They are usually given in the form $r = f(\theta)$.

- To sketch a polar curve $r = f(\theta)$ there are various tools to help you (in most cases completely analogous to the tools have help you draw $y = f(x)$).

  1. If in doubt, throw θ values into your $r = f(\theta)$ and work out r-values for given θ-values and plot them.
  2. If you are trying to draw $r = f(\theta)$, some students find it helpful to draw $y = f(x)$ to discover the general behavior of f.
  3. Understand that solving $\frac{dr}{d\theta} = 0$ gives you points where the curve is locally closest or furthest from the pole. If $\frac{d^2r}{d\theta^2} > 0$ then it is a point closest to the pole. If $\frac{d^2r}{d\theta^2} < 0$ then it is a point furthest from the pole.
  4. Solving $r = 0$ will give you the θ values (θ_1, say) that represent where the curve drops into the pole. Therefore the line $\theta = \theta_1$ will represent a tangent to the curve.
  5. Look for symmetries in your function. For example $r = \sin\theta$. We know the sine wave has symmetry about $\frac{\pi}{2}$; i.e. $\sin(\pi - \theta) \equiv \sin\theta$. Therefore the line $\theta = \frac{\pi}{2}$ must represent a line of symmetry on the polar curve.

     More generally if you can show that $f(2\alpha - \theta) \equiv f(\theta)$ then $\theta = \alpha$ represents a line of symmetry on the polar curve $r = f(\theta)$.

- To find areas on a polar graph we use the formula

$$\text{Area} = \int_{\theta_1}^{\theta_2} \tfrac{1}{2} r^2 \, d\theta.$$

- To convert from cartesian form to polar form use the relationships

$$x^2 + y^2 = r^2, \qquad x = r\cos\theta, \qquad y = r\sin\theta, \qquad \tan\theta = \tfrac{y}{x}.$$

  These can be derived easily from the point (x, y) drawn with a right angled triangle to the origin.

Hyperbolic Functions

- Know that the hyperbolic trigonometric functions[20] are defined

$$\cosh x \equiv \frac{e^x + e^{-x}}{2}, \qquad \sinh x \equiv \frac{e^x - e^{-x}}{2}, \qquad \tanh x \equiv \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

[19] Otherwise $(3, 0)$, $(3, 2\pi)$, $(3, 4\pi)$, ... would all be the same point.

[20] Compare with normal trig functions $\cos x \equiv \frac{e^{ix}+e^{-ix}}{2}$, $\quad \sin x \equiv \frac{e^{ix}-e^{-ix}}{2i}$, $\quad \tan x \equiv \frac{\sin x}{\cos x}$.

Similarly (as you would expect) we define

$$\operatorname{sech} x \equiv \frac{1}{\cosh x}, \qquad \operatorname{cosech} x \equiv \frac{1}{\sinh x}, \qquad \coth x \equiv \frac{1}{\tanh x}.$$

- Two important relationships that drop out instantly are

$$\cosh x + \sinh x = e^x \qquad \text{and} \qquad \cosh x - \sinh x = e^{-x}.$$

- You must know (or, better yet be able to work out from the definitions) the sketches for all 6 hyperbolic curves. Also $\sinh x$ is 'one-to-one' so can be inverted without restricting the domain. However, $\cosh x$ is 'many-to-one' so a domain restriction is required ($x \geqslant 0$) to invert it.

- Differentiating the above definitions we quickly find

$$\frac{d}{dx}\cosh x = \sinh x \qquad \frac{d}{dx}\sinh x = \cosh x.$$

- We also find $\cosh^2 x - \sinh^2 x = 1$ and $\sinh 2x = 2\sinh x\cosh x$. To derive results like these, run back to the exponential definitions and work from one side to the other. For example to prove the latter of the two results stated, start with $2\sinh x\cosh x$:

$$\begin{aligned} 2\sinh x\cosh x &= 2\left(\frac{e^x - e^{-x}}{2}\right)\left(\frac{e^x + e^{-x}}{2}\right) \\ &= \frac{(e^x - e^{-x})(e^x + e^{-x})}{2} \\ &= \frac{e^{2x} + 1 - 1 - e^{-2x}}{2} = \sinh 2x. \end{aligned}$$

- You need to know the logarithmic forms[21] for the inverse hyperbolic functions:

$$\begin{aligned} \sinh^{-1} x &= \ln\left(x + \sqrt{x^2+1}\right) && \text{for all } x, \\ \cosh^{-1} x &= \ln\left(x + \sqrt{x^2-1}\right) && \text{for } x \geqslant 1, \\ \tanh^{-1} x &= \tfrac{1}{2}\ln\left(\frac{1+x}{1-x}\right) && \text{for } -1 < x < 1. \end{aligned}$$

- For example: Solve $24\cosh x + 16\sinh x = 21$. Re-write the $\sinh x$ and $\cosh x$ in terms of e^x and then solve the resulting 'quadratic in disguise'.

$$\begin{aligned} 12(e^x + e^{-x}) + 8(e^x - e^{-x}) &= 21, \\ 20e^x - 21 + 4e^{-x} &= 0, \\ 20(e^x)^2 - 21(e^x) + 4 &= 0, \\ (5e^x - 4)(4e^x - 1) &= 0. \end{aligned}$$

This then solves to $x = \ln\frac{4}{5}$ or $x = \ln\frac{1}{4}$.

- A particularly useful identity which helps in some tougher problems is $(x - \sqrt{x^2-1})(x + \sqrt{x^2-1}) \equiv 1$. So

$$\frac{1}{x - \sqrt{x^2-1}} \equiv x + \sqrt{x^2-1} \qquad \text{and} \qquad \frac{1}{x + \sqrt{x^2-1}} \equiv x - \sqrt{x^2-1}.$$

[21] Derived by considering $y = \sinh^{-1} x$, $\Rightarrow$ $\sinh y = x$, $\Rightarrow$ $e^y - \frac{1}{e^y} = 2x$, $\Rightarrow$ $(e^y)^2 - 2x(e^y) - 1 = 0$. Then solve the resulting 'quadratic in disguise' for e^y.

Differentiation & Integration

- In C3 you will have seen the wonderful trick to find the derivative of $\ln x$:

$$\begin{aligned} y &= \ln x \\ e^y &= x \\ e^y &= \frac{dx}{dy} \\ \frac{1}{e^y} &= \frac{dy}{dx} \\ \frac{dy}{dx} &= \frac{1}{x}. \end{aligned}$$

We can use the same trick for inverse trig functions:

$$\begin{aligned} y &= \sin^{-1} x \\ \sin y &= x \\ \cos y &= \frac{dx}{dy} \\ \frac{1}{\cos y} &= \frac{dy}{dx} \\ \frac{dy}{dx} &= \frac{1}{\sqrt{1-\sin^2 y}} = \frac{1}{\sqrt{1-x^2}}. \end{aligned}$$

Similarly we can derive the following important results (you should do so for yourself!):

$$\frac{d}{dx}\sin^{-1} x = \frac{1}{\sqrt{1-x^2}}, \qquad \frac{d}{dx}\cos^{-1} x = -\frac{1}{\sqrt{1-x^2}}, \qquad \frac{d}{dx}\tan^{-1} x = \frac{1}{1+x^2},$$
$$\frac{d}{dx}\sinh^{-1} x = \frac{1}{\sqrt{1+x^2}}, \qquad \frac{d}{dx}\cosh^{-1} x = \frac{1}{\sqrt{x^2-1}}, \qquad \frac{d}{dx}\tanh^{-1} x = \frac{1}{1-x^2}.$$

- A glance at the formula book[22] shows that the above six derivations yield results such as $\int \frac{1}{a^2+x^2}\,dx = \frac{1}{a}\tan^{-1}\frac{x}{a} + c$. However, you should not allow yourself to get tied down to the formula book. I am a firm believer that the formula booklet should act as a guide only; showing you what substitution to use. For example if we needed to find $\int \frac{5}{9+4x^2}\,dx$ we would be lost with only the formula book, because it is not in precisely the same form. However it 'looks like' the '$\tan^{-1}$' answer in the formula book, so this is the hint to use a 'tan' substitution. Here we want $4x^2 = 9\tan^2\theta$; or, more simply, $2x = 3\tan\theta$. So (deep breath!)

$$\begin{aligned} &\int \frac{5}{9+4x^2}\,dx && 2x = 3\tan\theta \\ =&\int \frac{5}{9+9\tan^2\theta}\frac{3}{2}\sec^2\theta\,d\theta && 2\,dx = 3\sec^2\theta\,d\theta \\ =&\frac{15}{2}\int \frac{\sec^2\theta}{9\sec^2\theta}\,d\theta \\ =&\frac{5}{6}\theta + c = \frac{5}{6}\tan^{-1}\left(\frac{2x}{3}\right) + c. \end{aligned}$$

[22] If your school has not provided you with a copy, you should ask for (demand) one. It is very useful to know what's in it. However, if you're going to an interview at a top university and you say to your interviewer "I would have to look at a formula book to answer that" then you can expect a rejection letter soon after.

- Completing the square is another useful thing to look for. Here I don't necessarily mean the strict C1 method where $9x^2 + 6x - 15$ becomes $9(x + \frac{1}{3})^2 - 16$. A much more useful form for the former is $(3x + 1)^2 - 16$, keeping everything in integers. So if asked to work out $\int \frac{7}{\sqrt{9x^2+6x-15}}\,dx$ we can re-write it as $\int \frac{7}{\sqrt{(3x+1)^2-16}}\,dx$. This looks very similar to the '$\cosh^{-1}$' differential above. Therefore we need a 'cosh' substitution. Here we want $(3x + 1)^2 = 16\cosh^2 u$; or, more simply, $3x + 1 = 4\cosh u$. So (here we go!)

$$
\begin{aligned}
&\int \frac{7}{\sqrt{(3x+1)^2 - 16}}\,dx && 3x + 1 = 4\cosh u \\
=&\int \frac{7}{\sqrt{16\cosh^2 u - 16}}\frac{4}{3}\sinh u\,du && 3\,dx = 4\sinh u\,du \\
=&\frac{28}{3}\int \frac{\sinh u}{4\sinh u}\,du \\
=&\frac{7}{3}u + c = \frac{7}{3}\cosh^{-1}\left(\frac{3x+1}{4}\right) + c.
\end{aligned}
$$

- Another useful trick is to split the numerator of a fraction in an integral into two bits of more use. For example, if faced with $\int \frac{2x+3}{x^2+4x+1}\,dx$ you can split it into $\int \frac{2x+4}{x^2+4x+1} - \frac{1}{x^2+4x+1}\,dx$, each bit of which is now more easily handled.

- A useful substitution for integrals that involve trigonometric functions is $t = \tan\left(\frac{x}{2}\right)$. This is a boon because it changes horrible integrals with trig functions into new integrals with no trig at all[23]. Given this substitution it can be shown that

$$
\tan x = \frac{2t}{1-t^2}, \qquad \sin x = \frac{2t}{1+t^2}, \qquad \cos x = \frac{1-t^2}{1+t^2}.
$$

These must be learnt! When applying this, you must also use the fact that $\frac{dx}{dt} = \frac{2}{1+t^2}$, to replace the '$dx$' at the end of the integral by '$\frac{2}{1+t^2}\,dt$'. For example to evaluate $\int \frac{\sin x}{1+\cos x}\,dx$ we find

$$
\begin{aligned}
\int \frac{\sin x}{1+\cos x}\,dx &= \int \frac{\frac{2t}{1+t^2}}{1+\frac{1-t^2}{1+t^2}}\frac{2}{1+t^2}\,dt && \text{using } t = \tan\left(\frac{x}{2}\right) \\
&= \int \frac{2t}{1+t^2}\,dt \\
&= \ln(1+t^2) + c \\
&= \ln(1+\tan^2\left(\frac{x}{2}\right)) + c.
\end{aligned}
$$

The only thing to add is that if you were faced with $\int \frac{\sin 10x}{1+\cos 10x}\,dx$ the substitution would be $t = \tan 5x$. This would change the 'dx' replacement (by the chain rule) to '$\frac{1}{5(1+t^2)}\,dt$'. Therefore

$$
\int \frac{\sin 10x}{1+\cos 10x}\,dx = \int \frac{\frac{2t}{1+t^2}}{1+\frac{1-t^2}{1+t^2}}\frac{1}{5(1+t^2)}\,dt = \frac{1}{10}\int \frac{2t}{1+t^2}\,dt.
$$

Make sure you 'get' this; it is a little subtle.

[23] I don't know about you, but I *hate* trig integrals and the sooner I can get rid of the trig bits the better.

Reduction Formulae

- Reduction formulae involve integrals which do not only involve x, but also n. In general we write
$$I_n = \int (\text{something to do with } x \text{ and } n)\,dx$$
to indicate that the integral depends on n. The aim (usually) is to find a relationship between I_n and I_{n-1} or a relationship between I_n and I_{n-2} and then use this relationship to evaluate a specific integral (I_6, say). Integration by parts tends to be the method needed to find such a relationship since the integration by parts formula[24] contains an integral on each side of the equation which may be manipulated into the desired relationship. You do occasionally need to be quite cunning![25]

- For example find $\int_0^1 x^6 e^{2x}\,dx$. Clearly[26] the hint is to let $I_n = \int_0^1 x^n e^{2x}\,dx$. By parts
$$\begin{aligned} I_n &= \int_0^1 x^n e^{2x}\,dx = \left[\frac{x^n e^{2x}}{2}\right]_0^1 - \frac{n}{2}\int_0^1 x^{n-1} e^{2x}\,dx, \\ I_n &= \frac{e^2}{2} - \frac{n}{2} I_{n-1}. \end{aligned}$$
Now we have the relationship between I_n and I_{n-1} we need some 'low' integral that we can evaluate easily: I_0 fits the bill since $I_0 = \int_0^1 e^{2x}\,dx = \frac{e^2-1}{2}$. So
$$\begin{aligned} I_6 &= \frac{e^2}{2} - \frac{6}{2} I_5 \\ &= \frac{e^2}{2} - \frac{6}{2}\left(\frac{e^2}{2} - \frac{5}{2} I_4\right) \\ &= \text{...work through for yourself...} \\ &= \frac{7e^2}{8} - \frac{45}{8}. \end{aligned}$$

- 'Snapping off' bits of trig functions often helps (i.e. writing $\sin^n x$ as either $\sin x \sin^{n-1} x$ or $\sin^2 x \sin^{n-2} x$). For example find a reduction formula for $I_n = \int \cos^n x\,dx$. Snap off a '$\cos x$' and then do parts, integrating the $\cos x$ and differentiating the $\cos^{n-1} x$.
$$\begin{aligned} I_n &= \int \cos^n x\,dx = \int \cos x \cos^{n-1} x\,dx \\ &= \sin x \cos^{n-1} x + (n-1)\int \sin^2 x \cos^{n-2} x\,dx \\ &= \sin x \cos^{n-1} x + (n-1)\int (1-\cos^2 x)\cos^{n-2} x\,dx \\ I_n &= \sin x \cos^{n-1} x + (n-1) I_{n-2} - (n-1) I_n. \end{aligned}$$
Isolating I_n we find $I_n = \frac{1}{n}\sin x \cos^{n-1} x + \frac{n-1}{n} I_{n-2}$. We can find I_0 and I_1 easily enough, which means we can evaluate $I_n = \int \cos^n x\,dx$ for any positive integer n.

[24] $\int u \frac{dv}{dx}\,dx = uv - \int v \frac{du}{dx}\,dx$.
[25] Remember the Alloway special!
[26] Hopefully you can see why letting $I_n = \int_0^1 x^7 e^{nx}\,dx$ is a dreadful idea!

Maclaurin Series

- The Maclaurin series/expansion for a function is given by

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \frac{f''''(0)}{4!}x^4 + \cdots = \sum_{r=0}^{\infty} \frac{f^{(r)}(0)}{r!}x^r.$$

  This is a remarkable formula; it implies that you can know a function completely over *all* values of x provided you know all the derivatives of a function at *one* value of x.

- You must know (and any good candidate ought to derive for themselves) the following standard expansions:

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots && \text{valid for all } x, \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots && \text{valid for all } x, \\ \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots && \text{valid for all } x, \\ \ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots && \text{valid for } -1 < x < 1. \end{aligned}$$

  It is good to note that the e^x series differentiates to itself and the $\sin x$ and $\cos x$ series differentiate twice to minus themselves (as they should). Also we note that if we differentiate $\ln(1+x)$ we get $\frac{1}{1+x}$ and the general binomial expansion of this (using C4 methods) is precisely what we get by differentiating our Maclaurin expansion[27].

- You rarely (if ever) need to derive a Maclaurin series from first principles[28]. What you need to do is apply the series in the 'formula booklet' to similar situations.

- For example find the Maclaurin series for $\frac{3\cos(2x)}{1+\ln(1-4x)}$. So

$$\begin{aligned} \frac{3\cos(2x)}{1+\ln(1-4x)} &= \frac{3\left(1 - \frac{(2x)^2}{2!} + \frac{(2x)^4}{4!} - \dots\right)}{1 + \left((-4x) - \frac{(-4x)^2}{2} + \frac{(-4x)^3}{3} - \dots\right)} \\ &= \frac{3 - 6x^2 + 2x^4 - \dots}{1 - 4x - 8x^2 + \dots} \\ &= 3 + 12x + 66x^2 + \dots \end{aligned}$$

  To do the last step you consider the general binomial expansion on $(1 - 4x - 8x^2 + \dots)^{-1}$.

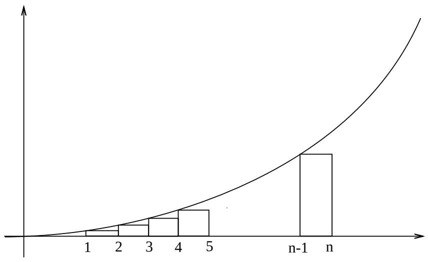
Series & Integrals

- You must be able to sandwich certain integrals between sums ($\text{sum}_1 < \text{integral} < \text{sum}_2$) and sandwich certain sums between integrals ($\text{integral}_1 < \text{sum} < \text{integral}_2$). The first of these is easier to formulate, but the second of these is more useful since you are currently better at integrals than sums.

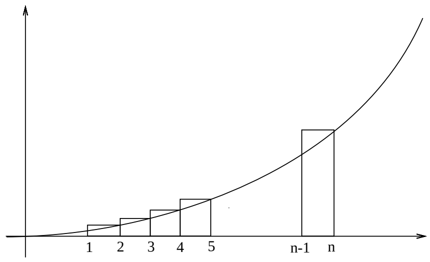
- You must be abundantly clear whether you are dealing with an *increasing* or *decreasing* function and you must always draw a sketch of the relevant curve and associated rectangles to make sure you are not writing gibberish (as I have occasionally done in class). Remember a function is increasing if $\frac{dy}{dx} \geqslant 0$ for all x-values in a range. Similarly a function is decreasing if $\frac{dy}{dx} \leqslant 0$.

[27] $1 - x + x^2 - x^3 + \dots$
[28] That means you McKelvie!

- For example sandwich $\int_1^n x^3\,dx$ between two sums. Firstly $y = x^3$ is an increasing function in the range stated, so if we want the sum below the integral we want the rectangles where the left height joins the curve. So $1^3 + 2^3 + \cdots + (n-1)^3 < \int_1^n x^3\,dx$. Similarly the sum above the integral is where the right height joins the curve.



and



So

$$
\begin{aligned}
1^3 + 2^3 + \cdots + (n-1)^3 \;&<\; \int_1^n x^3\,dx \;<\; 2^3 + 3^3 + \cdots + n^3,\\
\sum_{i=1}^{n-1} i^3 \;&<\; \int_1^n x^3\,dx \;<\; \sum_{i=2}^{n} i^3.
\end{aligned}
$$

- For example sandwich $\int_1^n \frac{1}{x^2}\,dx$ between two sums. Here we have a decreasing function in the range required, so the lower limit is now given by the rectangles whose right heights join the curve. You should therefore find

$$
\begin{aligned}
\left(\tfrac{1}{2}\right)^2 + \left(\tfrac{1}{3}\right)^2 + \cdots + \left(\tfrac{1}{n}\right)^2 \;&<\; \int_1^n \frac{1}{x^2}\,dx \;<\; \left(\tfrac{1}{1}\right)^2 + \left(\tfrac{1}{2}\right)^2 + \cdots + \left(\tfrac{1}{n-1}\right)^2,\\
\sum_{i=2}^{n} \left(\tfrac{1}{i}\right)^2 \;&<\; \int_1^n \frac{1}{x^2}\,dx \;<\; \sum_{i=1}^{n-1} \left(\tfrac{1}{i}\right)^2.
\end{aligned}
$$

- Notice the way the limits on the sums seem to *flip* between increasing and decreasing functions.

- Sandwiching a sum between two integrals is a little more fiddly. For example sandwich $\sqrt[3]{1} + \sqrt[3]{2} + \sqrt[3]{3} + \cdots + \sqrt[3]{n}$ between two integrals. Clearly the function we are considering here is $y = \sqrt[3]{x}$; this is an increasing function in the range. By considering two suitable sketches we find

$$
\begin{aligned}
\int_0^n \sqrt[3]{x}\,dx \;&<\; \sqrt[3]{1} + \sqrt[3]{2} + \sqrt[3]{3} + \cdots + \sqrt[3]{n} \;<\; \int_1^{n+1} \sqrt[3]{x}\,dx,\\
\frac{3n^{4/3}}{4} \;&<\; \sqrt[3]{1} + \sqrt[3]{2} + \sqrt[3]{3} + \cdots + \sqrt[3]{n} \;<\; \frac{3(n+1)^{4/3} - 3}{4}.
\end{aligned}
$$

- Similarly if you want to sandwich $f(1) + f(2) + \cdots + f(n)$ between two integrals where $y = f(x)$ is a *decreasing* function for $1 < x < n$ you should find

$$
\int_1^{n+1} f(x)\,dx \;<\; f(1) + f(2) + \cdots + f(n) \;<\; \int_0^n f(x)\,dx.
$$

Draw a sketch to see why.

Numerical Methods

- In C3 you considered iterations of the form $x_{n+1} = F(x_n)$ which give a progression of values $x_0, x_1, x_2, \ldots x_i, \ldots x_n \ldots$ which hopefully *converge* towards a solution of the equation $x = F(x)$. The 'true value' of the solution is denoted α. We define the *error* at any point of the iteration to be the difference between the 'true value' and the value of the iteration at that point; i.e.

$$e_n = \alpha - x_n \qquad \text{or} \qquad e_i = \alpha - x_i$$

  It is obviously to be hoped that e_i's get smaller as the iteration progresses.

- [Taylor series (which are not technically on the FP2 syllabus, but are needed for a full understanding of what follows) are a generalisation of Maclaurin series. Whereas Maclaurin series are 'centred around' $x = 0$ and provide increasingly good approximations to a function around the y-axis, Taylor series provide approximations to a function around any x-value you choose. The Taylor expansion around $x = a$ is

$$F(x) = F(a) + F'(a)(x-a) + \frac{F''(a)}{2!}(x-a)^2 + \frac{F'''(a)}{3!}(x-a)^3 + \cdots.$$

- When considering the iteration $x_{n+1} = F(x_n)$ we can Taylor expand $F(x_n)$ about the root α, so

$$x_{n+1} = F(x_n) = F(\alpha) + F'(\alpha)(x_n-\alpha) + \frac{F''(\alpha)}{2!}(x_n-\alpha)^2 + \frac{F'''(\alpha)}{3!}(x_n-\alpha)^3 + \cdots.$$

  But note that $F(\alpha) = \alpha$ because we are solving the equation $x = F(x)$. Therefore $x_{n+1} = \alpha + F'(\alpha)(x_n-\alpha) + \frac{F''(\alpha)}{2!}(x_n-\alpha)^2 + \frac{F'''(\alpha)}{3!}(x_n-\alpha)^3 + \cdots.$]

- If $\boxed{F'(\alpha) \neq 0}$ and we are in the neighbourhood of (i.e. close to) the root we can truncate the Taylor series at the $(x_n - \alpha)$ term to obtain

$$x_{n+1} \approx \alpha + F'(\alpha)(x_n - \alpha).$$

  Rearranging we find $\alpha - x_{n+1} \approx F'(\alpha)(\alpha - x_n)$ which gives $e_{n+1} \approx F'(\alpha)e_n$ so

$$\boxed{\frac{e_{n+1}}{e_n} \approx F'(\alpha) \approx \text{constant.}}$$

  This shows that we require $-1 < F'(\alpha) < 1$ to get the desired *convergence* because we need the errors to get smaller as we iterate.

- If $\boxed{F'(\alpha) = 0}$ the second term in the Taylor series vanishes which means we need the next term, so

$$x_{n+1} \approx \alpha + \frac{F''(\alpha)}{2!}(x_n - \alpha)^2.$$

  This rearranges to $\alpha - x_{n+1} \approx -\frac{F''(\alpha)}{2!}(\alpha - x_n)^2$ and so $e_{n+1} \approx -\frac{F''(\alpha)}{2!}e_n^2$ and therefore

$$\boxed{e_{n+1} \propto (e_n)^2;}$$

  this is called *quadratic convergence* and these iterations converges quickly because if e_n is small (close to the root) then $(e_n)^2$ is much smaller (e.g. $0.01^2 = 0.0001$).

- The Newton-Raphson Method for numerical solution of equations is an ingenious method which takes the tangent to a curve at a point and uses its x-axis intercept as the next value for the iteration. For a given start value $x = x_1$ the iteration is given by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

  This is derived thus:

  - Start with x_n,
  - Go to the curve at this point $(x_n, f(x_n))$,
  - Construct tangent using $f'(x_n)$ as the gradient and $y - y_1 = m(x - x_1)$,
  - $y - f(x_n) = f'(x_n)(x - x_n)$,
  - Put $y = 0$ to find where tangent crosses x-axis,
  - This x value is x_{n+1}.

  Newton-Raphson converges quadratically[29] explaining its speed.

[29] The N-R iteration $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ can be thought of as $x_{n+1} = F(x_n)$ and we wish to show that $F'(\alpha) = 0$ for quadratic convergence. So differentiating $F(x)$ we find

$$\begin{aligned} F'(x) &= \frac{d}{dx}(F(x)) \\ &= \frac{d}{dx}\left(x - \frac{f(x)}{f'(x)}\right) \\ &= 1 - \left(\frac{f'(x)f'(x) - f(x)f''(x)}{[f'(x)]^2}\right) \quad \text{(by the quotient rule)} \\ &= \frac{f(x)f''(x)}{[f'(x)]^2} \end{aligned}$$

Putting in $x = \alpha$ we obtain $F'(\alpha) = \frac{f(\alpha)f''(\alpha)}{[f'(\alpha)]^2}$. However $f(\alpha)$ must be zero because we are solving $f(x) = 0$ and α is a root. Therefore $F'(\alpha) = 0$ when using N-R $\Rightarrow$ quadratic convergence.

OCR Further Pure 3 Module Revision Sheet

The FP3 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J .M .S .

Differential Equations

First Order

- Given a first order differential equation, first see if it is separable *à la* C4.

$$\frac{dy}{dx} = f(x)g(y) \qquad \Rightarrow \qquad \int \frac{dy}{g(y)} = \int f(x)\,dx.$$

  Remember to add the "$+c$" to one side only. If you leave the "$+c$" unevaluated then your answer represents the general solution; otherwise your answer is a particular solution. (This will hardly ever happen in FP3.) In these notes $P \equiv P(x)$ and $Q \equiv Q(x)$.

- Differential equations can sometimes be changed by a substitution into a more accessible form for solution. Usually this will be trivial, but sometimes replacing the differentiable bit will require the chain/product rules, and can be a little fiddly. For example find the general solution of

$$\frac{dy}{dx} = \frac{x^2+y^2}{2xy} \text{ by the substitution } y = xu.$$

  Clearly the question is guiding us to get rid of all the y's in the original equation so we find

$$\begin{aligned} \frac{d(xu)}{dx} &= \frac{x^2+(xu)^2}{2x(xu)}, \\ u + x\frac{du}{dx} &= \frac{1+u^2}{2u}, \\ x\frac{du}{dx} &= \frac{1-u^2}{2u}. \end{aligned}$$

  This is separable, and we obtain the solution $\frac{1}{1-u^2} = Ax$ for some arbitrary A. Eliminating u from this using $y = xu$ we find $y^2 = x^2 - kx$ for some constant k.

- If it is not separable and there is no substitution to try then try to re-arrange into the form

$$\frac{dy}{dx} + P(x)y = Q(x).$$

  You will then need to multiply by an integrating factor (IF) which is defined as IF $= e^{\int P\,dx}$. Then we can[30] re-write the LHS in the form $d(\dots)/dx$.

[30] By understanding that $\frac{d}{dx}(ye^{\int P\,dx}) = e^{\int P\,dx}\frac{dy}{dx} + Pe^{\int P\,dx}y$.

- Example solve the differential equation $x\frac{dy}{dx} + 2y = \frac{4}{x}$. We rearrange to get $\frac{dy}{dx} + \frac{2}{x}y = \frac{4}{x^2}$ the we work out the IF to be $e^{\int \frac{2}{x}dx} = e^{2\ln x} = e^{\ln x^2} = x^2$ so the equation becomes $x^2\frac{dy}{dx} + 2xy = 4$. We can now write the LHS to obtain

$$\frac{d(x^2y)}{dx} = 4 \quad \Rightarrow \quad \int d(x^2y) = \int 4\,dx \quad \Rightarrow \quad x^2y = 4x + c \quad \Rightarrow \quad y = \frac{4}{x} + \frac{c}{x^2}.$$

  This represents the general solution of the differential equation. If we were told that $y = 6$ when $x = 1$ we would put this into the GS and find $c = 2$. This would give us a particular solution of $y = \frac{4}{x} + \frac{2}{x^2}$.

- *Homogeneous linear* differential equations are of the form $\frac{dy}{dx} + ay = 0$ or $\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = 0$ for constant a and b.

  Non-homogeneous linear differential equations are of the form $\frac{dy}{dx} + ay = f(x)$ or $\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = f(x)$ for constant a and b.

- Linear (and *only* linear) differential equations can be approached by the *auxiliary equation method*: we try a solution of the form $y = Ae^{\lambda x}$ in the original equation made homogeneous and construct an auxiliary equation in λ. We can then find λ and this will give us a complementary function (CF). For example $\frac{dy}{dx} - 3y = 2x$ would be modified to $\frac{dy}{dx} - 3y = 0$. Then the AE would be $\lambda Ae^{\lambda x} - 3Ae^{\lambda x} = 0$ so $\lambda = 3$ so the CF would be $y = Ae^{3x}$. If the original equation was homogeneous then the CF *is* the GS of the equation, but if it is non-homogeneous then you need to find particular integral (PI) which we add to the CF to get the GS. GS = CF + PI.

- In the above example we need to find a particular integral (PI) based on the form of $f(x)$ which in this case is $2x$. You guess the type of it depending on the type of $f(x)$. Here is a table of logical trials:

$f(x)$	TRIAL
linear	$lx + m$
polynomial order n	$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$
trig functions involving sin or cos px	$l\sin px + m\cos px$
exponential involving e^{px}	ce^{px}

  So here we try a PI of $y = mx + n$. Putting it in we find $m - 3(mx + n) = 2x$ and equating coefficients we find $m = -\frac{2}{3}$ and $n = -\frac{2}{9}$. The general solution (GS) would then be

$$\text{GS} = \text{CF} + \text{PI} \qquad \Rightarrow \qquad y = Ae^{3x} - \tfrac{2}{3}x - \tfrac{2}{9}.$$

Second Order

- Given a second order DE of the form

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = f(x)$$

  for constant a and b you construct the auxiliary equation[31] (AE) $\lambda^2 + a\lambda + b = 0$. The solution to the AE dictates the form of the complementary function (CF). There are three cases to consider depending on the type of solutions the AE has:

[31] This is not plucked out of thin air! The AE is obtained by trying a solution of the form $y = e^{\lambda x}$

	Type AE		CF
1.	Real and distinct roots, λ_1 and λ_2	$\Rightarrow$	$y = Ae^{\lambda_1 x} + Be^{\lambda_2 x}$
2.	Repeated real root, α.	$\Rightarrow$	$y = e^{\alpha x}(A + Bx)$
3.	Complex roots, $\alpha \pm i\beta$	$\Rightarrow$	$y = e^{\alpha x}(A \sin \beta x + B \cos \beta x)$

Notice that each CF has two arbitrary constants A and B. This makes sense, because we are effectively integrating twice and so would expect this[32].

- If you are dealing with a homogeneous equation $\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = 0$ then, as with first order, the GS is just the CF and you are done (subject to boundary/initial conditions!).

- If you are dealing with a non-homogeneous equation $\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = f(x)$ then we need to find a particular integral (PI) with the same principles as above. For example solve the equation
$$\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = \sin x + \cos x.$$
This gives AE of $\lambda^2 - 3\lambda + 2 = 0$ so $\lambda = 1$ or $\lambda = 2$. Therefore the CF is $y = Ae^x + Be^{2x}$. Looking at $\sin x + \cos x$ we would clearly try $y = m \sin x + n \cos x$ as our PI. Putting this in we discover
$$\begin{aligned}(-m \sin x - n \cos x) - 3(m \cos x - n \sin x) + 2(m \sin x + n \cos x) &= \sin x + \cos x,\\ (-m + 3n + 2m) \sin x + (-n - 3m + 2n) \cos x &= \sin x + \cos x.\end{aligned}$$
Equating coefficients of $\sin x$ and $\cos x$ we discover $m + 3n = 1$ and $n - 3m = 1$. This solves to $m = -\frac{1}{5}$ and $n = \frac{2}{5}$ giving us a GS of
$$y = Ae^x + Be^{2x} - \tfrac{1}{5} \sin x + \tfrac{2}{5} \sin x.$$

- The only *caveat* to the PI guesses is if the trial function for a PI is the same as one of CFs: you then multiply the trial function by x. For example find the general solution of
$$\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + 6y = e^{2x}.$$
The auxiliary equation gives $\lambda = 2$ or $\lambda = 3$. This gives a CF of $y = Ae^{2x} + Be^{3x}$. We would normally try a PI of $y = me^{2x}$, but we notice that e^{2x} is part of the CF, so we try a PI of $y = mxe^{2x}$ instead. Put this into the original differential equation and we find
$$\begin{aligned}(2me^{2x} + 2me^{2x} + 4mxe^{2x}) - 5(me^{2x} + 2mxe^{2x}) + 6mxe^{2x} &= e^{2x},\\ 2m + 2m + 4mx - 5m - 10mx + 6mx &= 1,\\ m &= -1.\end{aligned}$$
Therefore the GS is $y = Ae^{2x} + Be^{3x} - xe^{2x}$.

- If you are given conditions to be satisfied by the system at the start, then you find the GS and then put in the information to find the value of the arbitrary constants.

[32]The number of arbitrary constants should always match the order of the equation.

Vectors

- Recall that a line through position vector $\mathbf{a}$ and with direction $\mathbf{d}$ is, in vector form, $\mathbf{r} = \mathbf{a} + \lambda\mathbf{d}$. Recall also that the line through position vectors $\mathbf{a}$ and $\mathbf{b}$ is $\mathbf{r} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a})$ or $\mathbf{r} = \mathbf{b} + \lambda(\mathbf{a} - \mathbf{b})$ or equivalent. This is because when I subtract two position vectors $(\mathbf{b} - \mathbf{a})$ it yields the translation vector that travels from $\mathbf{a}$ to $\mathbf{b}$. For example find the line that passes through $(1, 3, 4)$ and $(3, 1, 8)$: This gives $\mathbf{r} = \begin{pmatrix}1\\3\\4\end{pmatrix} + \lambda\begin{pmatrix}2\\-2\\4\end{pmatrix}$ which would then simplify to give $\mathbf{r} = \begin{pmatrix}1\\3\\4\end{pmatrix} + \mu\begin{pmatrix}1\\-1\\2\end{pmatrix}$. Always look to simplify the direction vector if possible.

- A line can also be given in cartesian form
$$\frac{x-a}{l} = \frac{y-b}{m} = \frac{z-c}{n}.$$
To convert cartesian to vector form place each of the three elements equal to λ and 'unwrap'. For example $\frac{x-3}{2} = \frac{y+7}{2} = \frac{2-z}{3}$: so
$$\frac{x-3}{2} = \lambda \qquad \frac{y+7}{2} = \lambda \qquad \frac{2-z}{3} = \lambda.$$
So $\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}3+2\lambda\\-7+2\lambda\\2-3\lambda\end{pmatrix}$, therefore $\mathbf{r} = \begin{pmatrix}3\\-7\\2\end{pmatrix} + \lambda\begin{pmatrix}2\\2\\-3\end{pmatrix}$. To convert the other way should be trivial; unwind the above (practice for yourself).

- A plane in 3D can be given by $\mathbf{r} = \mathbf{a} + \lambda\mathbf{p} + \mu\mathbf{q}$. You can think of this as a point $\mathbf{a}$ being stretched along the line with direction $\mathbf{p}$ and then that line being stretched along in direction $\mathbf{q}$ to form a plane. This is the least helpful form for the plane (IMHO). [If I saw this I would cross $\mathbf{p}$ and $\mathbf{q}$ to get $\mathbf{n}$ and then use the form $ax + by + cz = k$: see below.]

- A plane in 3D through point $\mathbf{a}$ and normal $\mathbf{n}$ can be given by $(\mathbf{r} - \mathbf{a}).\mathbf{n} = 0$ where $\mathbf{r}$ represents all the points on the plane[33]. This is because if the dot product is zero then $(\mathbf{r} - \mathbf{a})$ and $\mathbf{n}$ must be at right angles and $(\mathbf{r} - \mathbf{a})$ is the vector that travels from $\mathbf{a}$ to any point $\mathbf{r}$. A nice sketch in the textbook P250. This can then be written $\mathbf{r}.\mathbf{n} = \mathbf{a}.\mathbf{n}$ so $\begin{pmatrix}x\\y\\z\end{pmatrix}.\begin{pmatrix}n_1\\n_2\\n_3\end{pmatrix} = n_1x + n_2y + n_3z = \text{constant}$. Therefore...

- ...a plane can most usefully be in the form $ax + by + cz = k$ where $\begin{pmatrix}a\\b\\c\end{pmatrix}$ is the normal vector and k is some constant for *that* plane. As k varies it will produce parallel planes, like pages in a book.

- To find the intersection of a line and a plane is easy. Best done by example, find intersection of $\mathbf{r} = \begin{pmatrix}2\\1\\3\end{pmatrix} + t\begin{pmatrix}-1\\1\\3\end{pmatrix}$ and $2x+y-3z = -24$. From the line we know that $x = 2-t$ etc. so we place all of these into the plane and solve for t, so $2(2-t) + (1+t) - 3(3+3t) = -24$ which solves to give $t = 2$. Place this back into the line and the point is $(0, 3, 9)$.

[33] $\mathbf{r} = (x, y, z)$

- You also need to be able to determine whether a line lies *in* a plane or parallel to it. If you try to find where the line crosses the plane (like above) you will either boil your equation for λ down to a consistency $(1 = 1)$ in which case the line lies in the plane, or an inconsistency $(0 = 1)$ in which case the line is parallel to the plane.

- To find the shortest distance between a line and a point, do a sketch of the line and the point, and construct the triangle between the point away from the line (P), the point $\mathbf{a}$ on the line (A) and the point which is closest to the point (F). We want the length PF in the right angled triangle APF. The angle $P\widehat{A}F$ can be found by dotting the direction vector of the line with the vector $\overrightarrow{AP}$ and we can work out the length AP by working out the magnitude of $\overrightarrow{AP}$. Then it's just sin in a right angled triangle to get length PF.

- To find shortest distance between a point and a plane is a trivial extension of the above. Construct the line through the point with direction vector normal to the plane. Find where this line crosses the plane and then find the distance between these two points. [Quick reminder: The distance between the points (x_1, y_1, z_1) and (x_2, y_2, z_2) is
$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2},$$
or, equivalently, the distance between points with position vectors $\mathbf{a}$ and $\mathbf{b}$ is the magnitude of $(\mathbf{b} - \mathbf{a})$.]

- Must be able to find the intersection of two planes. For example find intersection of $x + 2y - 2z = 2$ and $2x + 3y - 7z = 1$: Up to you which variable you would like to eliminate, but I'd do twice the first minus the second. This gives $y + 3z = 3$. Let $z = t$ (you'd be silly to let $y = t$, try it to see why!) and we find $y = 3 - 3t$ and then $x = 2 - 2y + 2z = 2 - 2(3 - 3t) + 2(t) = -4 + 8t$. Therefore
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -4 + 8t \\ 3 - 3t \\ t \end{pmatrix} \qquad \text{so} \qquad \mathbf{r} = \begin{pmatrix} -4 \\ 3 \\ 0 \end{pmatrix} + t \begin{pmatrix} 8 \\ -3 \\ 1 \end{pmatrix}.$$
You could also have solved this by crossing the normals of the planes to discover the direction vector of the line of intersection and then find any point where the planes cross.

- The angle between two planes is the same as the angle between their normals. The angle between a plane and a line is $\frac{\pi}{2}$ minus the angle between the line and the plane's normal.

- The vector product of two vectors is defined
$$\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}| \sin\theta \widehat{\mathbf{n}}, \qquad \begin{pmatrix} a \\ b \\ c \end{pmatrix} \times \begin{pmatrix} l \\ m \\ n \end{pmatrix} = \begin{pmatrix} bn - cm \\ cl - an \\ am - bl \end{pmatrix}.$$
In practice you very rarely need the $\sin\theta$ bit of this definition (any angles you need are always more easily accessible by the dot product). It is most useful in the way it constructs a vector perpendicular to both original vectors.

- For example: Find the equation of the plane through $A(1, 2, -1)$,$B(2, 3, 4)$ and $C(-2, 0, 1)$. We construct the normal vector to the plane by crossing $\overrightarrow{AB}$ and $\overrightarrow{AC}$.
$$\overrightarrow{AB} \times \overrightarrow{AC} = \left(\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}\right) \times \left(\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix} \times \begin{pmatrix} -3 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ -17 \\ 1 \end{pmatrix}.$$
Therefore the plane is $12x - 17y + z = \text{const}$. We find the constant by taking your favourite of A, B or C and plugging it in[34]. So $12x - 17y + z = -23$.

[34] A nice check is to put more than one point in and check that you get the same constant each time.

- The shortest distance between the lines $\mathbf{r} = \mathbf{a}_1 + \lambda\mathbf{d}_1$ and $\mathbf{r} = \mathbf{a}_2 + \mu\mathbf{d}_2$ is given by

$$d = \frac{| (\mathbf{a}_2 - \mathbf{a}_1).(\mathbf{d}_1 \times \mathbf{d}_2) |}{| \mathbf{d}_1 \times \mathbf{d}_2 |}.$$

Complex Numbers

- Recall that the complex number $z = a + ib$ has modulus $|z| \equiv r = \sqrt{a^2 + b^2}$ and argument $\arg(z) = \theta = \tan^{-1}\frac{b}{a}$, where argument is measured anti-clockwise from the positive real axis. Also recall that arguments can be of either convention $0 \leqslant \arg(z) < 2\pi$ or $-\pi < \arg(z) \leqslant \pi$. This is purely arbitrary and will be clear from the question what they want. [In my mind arguments are such that $-\infty < \arg(z) < \infty$ and they can keep twirling round, but OCR forces each complex number to have a unique argument.]

- By considering a right angled triangle like the one at the top of P295 we discover $z = a + ib = r(\cos\theta + i\sin\theta)$. This is the *polar form* of a complex number.

  It can also be shown[35] that $r(\cos\theta + i\sin\theta) \equiv re^{i\theta}$. This is the *exponential form* of a complex number and is unbelievably useful![36]

- Properties of complex number multiplication and division become immediately apparent:

$$re^{i\alpha} \times \rho e^{i\beta} = (r\rho)e^{i(\alpha+\beta)} \qquad \text{When multiplying, add arguments and multiply moduli,}$$
$$\frac{re^{i\alpha}}{\rho e^{i\beta}} = \left(\frac{r}{\rho}\right) e^{i(\alpha-\beta)} \qquad \text{When dividing, subtract arguments and divide moduli.}$$

  So given a complex number w, if I were to multiply w by the complex number $2e^{\frac{i\pi}{3}}$ (say) then the result would be the complex number with twice the length/modulus of w and rotated $\frac{\pi}{3}$ anti-clockwise from the original w. This is called a spiral-enlargement.

- De Moivre's Theorem[37] states that

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta \qquad \text{or} \qquad (e^{i\theta})^n = e^{in\theta}.$$

  For integer n this can be proven by induction. The reason for the truth of De Moivre's theorem should be obvious from the above two properties for multiplication and division. To raise complex number w to the power four (say) the modulus would be raised to the power four, but the argument would be made four times bigger... which is what De Moivre says.

  We notice the special case

$$(\cos\theta + i\sin\theta)^{-1} = \cos\theta - i\sin\theta \qquad \text{or} \qquad (e^{i\theta})^{-1} = e^{-i\theta}$$

  which comes up a lot; the inverse of a complex number with unit modulus is its complex conjugate.

- You can use De Moivre's theorem to derive certain trigonometric results. You need

$$\cos n\theta = \text{Re}((\cos\theta + i\sin\theta)^n) \qquad \text{and} \qquad \sin n\theta = \text{Im}((\cos\theta + i\sin\theta)^n).$$

  You can then use the standard relationship $\sin^2\theta + \cos^2\theta = 1$ to manipulate any raw results. Your algebra needs to be top-notch here; any slip at the start of a question can

[35] By mucking about with $\cos\theta + i\sin\theta = (1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots) + i(x - \frac{x^3}{3!} + \dots) = \text{lots of working} = e^{i\theta}$.

[36] It should be noted that an alternative notation for e^x at a higher level is $\exp(x)$. This is because the power on the e can sometimes become quite complicated.

[37] Abraham de Moivre, an 18th century statistician and consultant to gamblers. French...

cost you dearly on a multi-parter. For example express $\cos 5\theta$ in terms of powers of $\cos\theta$. So

$$\begin{aligned}
\cos 5\theta &= \mathrm{Re}[(\cos\theta + i\sin\theta)^5] \\
&= \mathrm{Re}[\cos^5\theta + 5\cos^4\theta(i\sin\theta) + 10\cos^3\theta(i\sin\theta)^2 + 10\cos^2(i\sin\theta)^3 + 5\cos\theta(i\sin\theta)^4 + (i\sin\theta)^5] \\
&= \cos^5\theta - 10\cos^3\theta\sin^2\theta + 5\cos\theta\sin^4\theta \\
&= \cos^5\theta - 10\cos^3\theta(1-\cos^2\theta) + 5\cos\theta(1-\cos^2\theta)^2 \\
&= 16\cos^5\theta - 20\cos^3\theta + 5\cos\theta.
\end{aligned}$$

- For problems involving $\tan n\theta$ you would consider the expansions of $\sin n\theta$ and $\cos n\theta$ and use $\tan n\theta \equiv \dfrac{\sin n\theta}{\cos n\theta}$. For example

  $$\tan 3\theta = \frac{\sin 3\theta}{\cos 3\theta} = \frac{\mathrm{Im}[(\cos\theta + i\sin\theta)^3]}{\mathrm{Re}[(\cos\theta + i\sin\theta)^3]} = \frac{3\cos^2\theta\sin\theta - \sin^3\theta}{\cos^3\theta - 3\cos\theta\sin^2\theta}.$$

  You could then divide both numerator and denominator by $\cos^3\theta$ to obtain the 'nicer'

  $$\frac{3\tan\theta - \tan^3\theta}{1 - 3\tan^2\theta}.$$

- Given any complex number of unit modulus $z = e^{i\theta} = \cos\theta + i\sin\theta$ it can easily be shown that

  $$\frac{1}{2}\left(z + \frac{1}{z}\right) = \cos\theta \qquad \text{and} \qquad \frac{1}{2i}\left(z - \frac{1}{z}\right) = \sin\theta.$$

  Similarly we can derive (by using De Moivre on the unit complex number; $z^n = (\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$) the very useful relations

  $$\frac{1}{2}\left(z^n + \frac{1}{z^n}\right) = \cos n\theta \qquad \text{and} \qquad \frac{1}{2i}\left(z^n - \frac{1}{z^n}\right) = \sin n\theta.$$

  So whereas we can use De Moivre to find multiple angle expressions (such as $\sin 6\theta$) in terms of powers of sin and cos, we can use the above to write powers of sin and cos (such as $\sin^7\theta$) in terms of multiple angles. For example express $\cos^6\theta$ as a sum of multiple angles of $\cos\theta$.

  $$\begin{aligned}
  \cos^6\theta &= \left[\frac{1}{2}\left(z + \frac{1}{z}\right)\right]^6, \\
  &= \frac{1}{64}\left(z^6 + 6z^4 + 15z^2 + 20 + \frac{15}{z^2} + \frac{6}{z^4} + \frac{1}{z^6}\right), \\
  &= \frac{1}{64}\left(\left(z^6 + \frac{1}{z^6}\right) + 6\left(z^4 + \frac{1}{z^4}\right) + 15\left(z^2 + \frac{1}{z^2}\right) + 20\right), \\
  &= \frac{1}{64}\left(2\cos 6\theta + 12\cos 4\theta + 30\cos 2\theta + 20\right), \\
  &= \frac{1}{32}\left(\cos 6\theta + 6\cos 4\theta + 15\cos 2\theta + 10\right).
  \end{aligned}$$

- You could then use the above to help you with integrals:

  $$\begin{aligned}
  \int \cos^6\theta\, d\theta &= \frac{1}{32}\int (\cos 6\theta + 6\cos 4\theta + 15\cos 2\theta + 10)\ d\theta, \\
  &= \tfrac{1}{32}(\tfrac{1}{6}\sin 6\theta + \tfrac{3}{2}\sin 4\theta + \tfrac{15}{2}\sin 2\theta + 10\theta + c), \\
  &= \frac{\sin 6\theta}{192} + \frac{3\sin 4\theta}{64} + \frac{15\sin 2\theta}{64} + \frac{5}{16} + c'.
  \end{aligned}$$

- You need to be able (using the ideas of roots of unity) to solve any equation of the form $z^n = a + ib$ (i.e. to get all n solutions to this equation). To get the *primary solution* you convert the $a + ib$ into the form $Re^{i\theta}$ and then obtain the first solution by raising both sides to the power $\frac{1}{n}$. So

$$\begin{aligned} z^n &= a + ib \\ z^n &= Re^{i\theta} \\ z &= (Re^{i\theta})^{1/n} \\ z &= \sqrt[n]{R}\, e^{\frac{i\theta}{n}}. \end{aligned}$$

  This is the primary solution and then you find the others by using the fact they are all evenly spaced around the circle of radius $\sqrt[n]{R}$, like spokes on a bike. So you keep adding $\frac{2\pi}{n}$ on to the argument of $\sqrt[n]{R}e^{\frac{i\theta}{n}}$, $n-1$ times to get all n solutions to the equation.

- For example solve $z^4 = -16$. We rewrite in the form $z^4 = 16e^{i\pi}$. Therefore the primary solution is $z = 2e^{\frac{i\pi}{4}}$. Adding on $2\pi/4 = \pi/2$ to the arguments we find the four solutions

$$z = 2e^{\frac{i\pi}{4}}, 2e^{\frac{i3\pi}{4}}, 2e^{\frac{i5\pi}{4}}, 2e^{\frac{i7\pi}{4}}.$$

  If need be you could then convert these back to $a + ib$ form and get:

$$z = (\sqrt{2} + i\sqrt{2}), (-\sqrt{2} + i\sqrt{2}), (-\sqrt{2} - i\sqrt{2}), (\sqrt{2} - i\sqrt{2}).$$

  By taking the roots in complex conjugate pairs you could then factorise $z^4 + 16$ into the product of two real quadratic factors as

$$\begin{aligned} z^4 + 16 &= (z - (\sqrt{2} + i\sqrt{2}))(z - (\sqrt{2} - i\sqrt{2}))(z - (-\sqrt{2} + i\sqrt{2}))(z - (-\sqrt{2} - i\sqrt{2})) \\ &= (z^2 - 2\sqrt{2}z + 4)(z^2 + 2\sqrt{2}z + 4). \end{aligned}$$

Groups

- A group $(G, \circ)$ is a non-empty set G with a binary operation $\circ$ which
  - is **closed**, (for every a and b in G, $a \circ b$ also lies in G),
  - is **associative**, (for every a, b and c in G, $(a \circ b) \circ c = a \circ (b \circ c)$),
  - has a unique **identity** element, (an element e such that $e \circ a = a \circ e = a$ for all a in G),
  - every element has its own **inverse**, (for every a in G there exists a^{-1} such that $a \circ a^{-1} = a^{-1} \circ a = e$).

  A group can be represented as a Latin square. For example

	e	a	b
e	$e \circ e$	$e \circ a$	$e \circ b$
a	$a \circ e$	$a \circ a$	$a \circ b$
b	$b \circ e$	$b \circ a$	$b \circ b$

  Each row and column must contain every element of G once only. You can find the identity easily from this by looking for the row or column which is unchanged. Inverses are easy to find from a Latin Square; you merely look for which other element makes it the identity.

- If you are asked to show that something is a group in an exam you must tick off each of the above criteria one-by-one. For example show that the set $\{1, -1, i, -i\}$ forms a group under complex number multiplication. Firstly create table

	1	-1	i	$-i$
1	1	-1	i	$-i$
-1	-1	1	$-i$	i
i	i	$-i$	-1	1
$-i$	$-i$	i	1	-1

  So we can see that it is closed. Complex number multiplication is associative[38]. Identity element: 1. Inverses: $1^{-1} = 1$, $(-1)^{-1} = -1$, $i^{-1} = -i$ and $(-i)^{-1} = i$.

- A group is commutative or Abelian if $a \circ b = b \circ a$ for all a and b in G. If you have a Latin Square for the group you can see if it is Abelian by seeing if it symmetrical along the leading diagonal.

- The *order of a group* is the number of elements the group contains. If a group contains an infinite number of elements it is said to be of *infinite order*.

- The *order of an element* a of G is the *smallest* n such that $a^n = e$. If no such n exists then the element is said to have *infinite order*. A group is *cyclic* if every element of a group can be generated by powers of a single element.

- A *subgroup* (of a group) is any non-empty subset of G which also forms a group under the same binary operation $\circ$. (A subgroup includes the subset containing just e and the subset G itself.) A *proper subgroup* is any subgroup with order not one or the same as the original group.

- A good way to find subgroups (beyond the cases where it is obvious) is to consider the powers of the elements of the original group; if you get back to e then the set of elements gone through will be a subgroup. For example in a group of order 16, if you take an element a and discover that $a^4 = e$ (i.e. the order of a is 4) then the set $\{e, a, a^2, a^3\}$ will form a subgroup.

- *Lagrange's theorem* states that the order of any subgroup must divide the order of the original group. For example a group of order 8 could potentially only have subgroups of order 1, 2, 4 or 8. It could therefore potentially only have proper subgroups of order 2 or 4. Some useful corollaries of Lagrange's Theorem include:
  - The order of an element *must* divide the order of the group.
  - A group of prime order *must* be cyclic.

- Two groups $(G, \circ)$ and $(H, \bullet)$ are isomorphic if there exists a one-to-one mapping between them which preserves their structure, i.e.

$$a \leftrightarrow x \text{ and } b \leftrightarrow y \qquad \Leftrightarrow \qquad a \circ b \leftrightarrow x \bullet y.$$

  A good way to show that groups are not isomorphic is to consider the orders of the elements of G and H: If they are different, then they *cannot* be isomorphic. In an exam you must make the mappings (something) $\leftrightarrow$ (something else) *very* clear; i.e. list them out!

[38]Do be careful what you can assume here! In the question it should tell you what you can assume. Beware of assuming anything that you are not told in the question. To assume makes an ass out of u and me!

- You need to know the structure of groups up to order 7. Groups of order 2, 3, 5 and 7 must be cyclic (prime order) and therefore every group of order p (say), must be isomorphic to every other group of order p. These groups are all isomorphic to $(\mathbb{Z}_p, +)$, the group of $\{0, 1, 2, p-1\}$ under addition mod p.

- There are two groups of order 4:

$(\mathbb{Z}_4, +)$

	e	a	a^2	a^3
e	e	a	a^2	a^3
a	a	a^2	a^3	e
a^2	a^2	a^3	e	a
a^3	a^3	e	a	a^2

and the Klein four-group

	e	a	b	ba
e	e	a	b	ba
a	a	e	ba	b
b	b	ba	e	a
ba	ba	b	a	e

Whereas the cyclic group is generated by a single element a, the Klein four-group is generated by two elements, a and b with $a^2 = b^2 = e$ and $ab = ba$. In the Klein four-group every element is self inverse (i.e. has order 2).

- For groups of order 6 there are two fundamental types, the cyclic group isomorphic to $(\mathbb{Z}_6, +)$ and the dihedral group D_3 which represents the symmetries of the regular triangle under rotation and reflection. The group is generated by the rotation $\frac{2\pi}{3}$ (a) and reflection (b) with $a^3 = b^2 = e$ and $ab = ba^2$. The table is:

	e	a	a^2	b	ba	ba^2
e	e	a	a^2	b	ba	ba^2
a	a	a^2	e	ba^2	b	ba
a^2	a^2	e	a	ba	ba^2	b
b	b	ba	ba^2	e	a	a^2
ba	ba	ba^2	b	a^2	e	a
ba^2	ba^2	b	ba	a	a^2	e

OCR Mechanics 1 Module Revision Sheet

The M1 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J .M .S .

Preliminaries

- Most mistakes in Mechanics tend to be "sign" issues. You must have a clear sense of which way you are 'defining' to be positive. It doesn't matter which way you define positive to be, but you must then be *consistent* through the whole question and interpret your answer in the context of that definition.
- Sometimes it will not be obvious which direction a quantity points; don't worry too much about it, just put an arrow and variable name next to your sketch. If the variable turns out positive then your guess was correct. If negative then it turns out it points the other way to which you guessed.
- This is going to sound somewhat 'new-age', but in Mechanics you must have a "feel" for what your equations are saying. You should be saying to yourself things like "So as t gets bigger, s gets bigger until the quadratic peak and then starts to fall". Quite often you can spot mistakes because the equations you construct simply have the wrong properties.
- In some problems involving surfaces at an angle, instead of explicitly giving the angle ($30°$, say) they will tell you that the angle α is such that $\tan\alpha = \frac{4}{3}$. *Do not* work out α by going to your calculator and typing $\tan^{-1}\frac{4}{3}$. Instead draw a right angle triangle such that $\tan\alpha = \frac{4}{3}$; using Pythagoras' Theorem we see that $\sin\alpha = \frac{4}{5}$ and $\cos\alpha = \frac{3}{5}$. Much nicer!

Forces

- You should not think of it as $F = ma$ and I don't ever want to see it written down. You should think of it as
$$\textit{Resultant}\ \text{Force} = ma.$$

Friction

- Friction always points in the direction opposing the motion or the potential motion (i.e. the motion that *would* occur if there was no friction).
- If a particle is moving or "on the point of moving" or "in limiting equilibrium" then friction is maximal such that
$$F_{\text{max}} = \mu R.$$
This is lovely in questions because it gives you a concrete equation to work with, rather than a slippery inequality.

"SUVAT" Equations (a.k.a. Kinematics Equations)

- You must know (and be able to use proficiently) the SUVAT equations. Each equation involves four of the five variables s, u, v, a, t. They state (in roughly decreasing order of importance IMHO):

$$v = u + at,$$
$$s = ut + \tfrac{1}{2}at^2,$$
$$v^2 = u^2 + 2as,$$
$$s = \left(\frac{u+v}{2}\right)t,$$
$$s = vt - \tfrac{1}{2}at^2.$$

  I have heard it stated by a person I respect that the last of those five is not a SUVAT equation; for the sake of elegance and symmetry I respectfully disagree.

- You can only use the SUVAT equations if you have a *constant* acceleration. Enough students (even moderately bright students with my poor teaching) balls this up at some point or another that I will say it again:

  "You can only use the SUVAT equations if you have a *constant* acceleration."

- I always insist students draw a table (no matter how simple the question is) and fill in what they know and what they *want* to know. I also demand an arrow to show which way they are defining *positive* to be in the context of the question. For example work out the maximum height reached by a ball thrown from ground level upwards with speed 20ms^{-1} I would expect:

	s	u	v	a	t
$(\uparrow)$	x	20	0	-9.8	-

  We have defined up to be positive (therefore $a = -9.8$) and we don't care about time so the equation we need is $v^2 = u^2 + 2as$. You must also start by writing down the equation you are going to use to show anyone reading/marking your work what you are trying to do[39].

General Motion

- Blah

Momentum

- The momentum of a particle is its velocity times its mass. It is a vector quantity.

importance of velocity time graphs
combining and splitting forces
good force diagram vital to success of mechanics.... don't mix force diagrams between objects....draw them as rectangles with forces coming out of the object.
don't simplify within the diagram....take components in the second.
reactions always at right angles to surface (not always equal to the weight!)

[39]Mechanics work, in particular, can turn into a bit of a bombsite if you're not careful.

friction always opposes motion or potential motion. Example ti figure out which out which way in ambiguous case.
strings distribute tensions
forces in equilibrium

1 Motion

MECHANICS 1 is all about the motion of objects. For all examination questions, one or more of the following modelling assumptions will be made:

- The object is modelled as a particle.
- No air resistance.
- No wind.
- All strings are light and inextensible.
- All pulleys are smooth.

1.1 Terminology

A scalar is a quantity with magnitude only, such as speed. A vector is a quantity with both magnitude and direction, such as velocity. In order to find the scalar from the vector (e.g. speed from velocity), just find the magnitude. For example

$$\text{Velocity} = (-3\mathbf{i} + 4\mathbf{j})ms^{-1} \qquad \Rightarrow \qquad \text{Speed} = \sqrt{(-3)^2 + 4^2} = 5ms^{-1}.$$

Displacement (often known as **position**) is the vector related to distance. For instance, if I walk 4 miles North, then 3 miles East, then the total distance I have walked is 7 miles, but my displacement from my starting point is 5 miles (by Pythagoras' Theorem).

1.2 Graphs of Motion

We can draw position/time, velocity/time, or acceleration/time graphs to represent the motion of an object. When drawing graphs like these the key is to label all axes correctly, mark on the points we know and then join them up sensibly. Things to remember about graphs are:

- The gradient of a position/time graph is the velocity.
- The gradient of a velocity/time graph is the acceleration.
- The area under a velocity/time graph is the displacement.
- The area under a speed/time graph is the distance.

1.3 Equations

Equations that you might need to know include

$$\begin{aligned} \text{Average speed} &= \frac{\text{total distance travelled}}{\text{time taken}}, \\ \text{Average velocity} &= \frac{\text{displacement}}{\text{time taken}}. \end{aligned}$$

Don't forget that to find the gradient of a straight line you need to do

$$\frac{\text{A little bit of } y}{\text{A little bit of } x} = \frac{dy}{dx}.$$

2 Constant Acceleration Formulae

Otherwise known as the SUVAT equations, these are the equations you will need to use whenever you are modelling the motion of something with constant acceleration. They provide a link between the following things:

- s – displacement (m)
- u – initial velocity (ms^{-1})
- v – final velocity (ms^{-1})
- a – acceleration (ms^{-2})
- t – time (s)

2.1 The Equations

The equations you should know by heart are:

$$\begin{aligned} v &= u + at \\ s &= \frac{u+v}{2} \times t \\ s &= ut + \tfrac{1}{2}at^2 \\ v^2 &= u^2 + 2as \\ s &= vt - \tfrac{1}{2}at^2 \end{aligned}$$

2.2 Vertical Motion under Gravity

Whenever a problem occurs which involves an object falling or flying through the air (in any direction) then there are only three things to remember:

- DRAW A DIAGRAM!
- Decide which way is positive and where your origin is.
- Acceleration is gravity (usually taken to be $9.8ms^{-2}$).

The easiest thing to forget here is to add the original height of the object back on to the answer obtained using SUVAT, for instance when finding the greatest height above the ground of an object which didn't start on the ground.

3 Basic Forces

A force is a vector; it has both magnitude (size, i.e. how strong it is) and direction (i.e. whether it pushes or pulls, and in which direction). Newton's Laws of Motion help to explain the behaviour of forces.

3.1 Newton's Laws

If you can only remember one of these three laws, then make it Newton's Second, which is used throughout the Mechanics 1 course. It can be remembered as $F = ma$ but you *must* remember that F is resultant force and not just force! The laws in full are listed here:

3.1.1 Newton's First Law

Every particle continues in a state of rest or uniform motion in a straight line unless acted on by a resultant external force.

3.1.2 Newton's Second Law

The change in motion is proportional to the force.

3.1.3 Newton's Third Law

Every action has an equal and opposite reaction.

3.2 Types of Force

Remember that resultant force in any particular direction can be found using

$$\text{Resultant force} = (\text{Forces in that direction}) - (\text{Forces in opposite direction})$$

Once you know this, all you need to remember about forces can be covered in the following few points:

- Forces are measured in newtons (N).
- The weight (mg) always acts vertically downwards.
- The normal reaction exists only when the object is on a surface. It always acts perpendicularly upwards from the surface.
- Friction is always in the opposite direction to motion.
- A driving force goes in the direction of motion, a braking force goes in the opposite direction.

3.3 Tension and Thrust

Remember Newton's Third Law? If you push inwards on the ends of a pencil, you are putting it in compression (sometimes called thrust). It follows therefore that the pencil will be pushing outwards on your hands:

Similarly, if the pencil is in tension, the forces in the pencil will be inwards towards each other (and away from your hands):

If in doubt, just remember that tensions always go towards one another. That is, tension in a string or a rod always goes *away* from the object it is attached to. Also remember that tensions on either side of a pulley are equal, regardless of whether they both go in the same direction or otherwise.

4 Applying Newton's Second Law Along a Line

Writing down Newton's Second Law for an object with some numbers in gives you an **equation of motion** for the object.

For one object, this is simple. However, when objects are connected to one another, things become a little more tricky. There are only two cases though, and you need to remember the following points for each:

4.1 Transport

Here, you will be given a situation like a train with various carriages, or a car with a caravan.

- To find acceleration, consider the whole system.
- Once you have acceleration, you can use SUVAT if you need to.
- To find the tension in a rope or in a rod, use the last object in the chain only (i.e. the caravan or the last carriage). Remember that the only forces applying to the last object in the line of its motion should be tension and resistance for that one object. Use acceleration as found for the whole system.
- If the acceleration changes again for whatever reason (e.g. brakes applied), then you will need to go back to Step 1 and find acceleration using the whole system!

4.2 Pulleys

Here, you will be given two or more objects connected by light inextensible strings over smooth pulleys. You must remember that tensions are *towards* the pulleys (away from the objects, as explained before) and are equal. If you have one pulley, you will have one tension, T, on your diagram (in two different places). If you have two pulleys, you should expect to have two tensions, T_1 and T_2.

- The system moves in the direction of the heaviest dangling object. Draw each object's direction of acceleration.
- Write down an equation of motion for each object.
- Add these together. This will eliminate *all* the tensions!
- Find acceleration, and substitute back in for tension.
- Now you have the acceleration, you can use SUVAT if required.

5 Vectors

Vectors have both magnitude and direction. That is, they are in two or three dimensions. If they are expressed in terms of **i**'s, **j**'s and possibly **k**'s, then they are in component form. Vectors can also be expressed as columns. Don't forget to underline all your vectors!

5.1 Magnitude and Direction

Let

$$v = a_1\mathbf{i} + a_2\mathbf{j},$$

then

$$\text{Magnitude} = \sqrt{(a_1)^2 + (a_2)^2}.$$

For direction

$$\tan\theta = \frac{a_2}{a_1}.$$

Be careful, as direction can be tricky. Always draw a diagram of your **i**'s and then your **j**'s, and remember that direction is taken from the positive **i** direction as follows:

Basic facts about vectors include:

- To add vectors, add the **i**'s and **j**'s separately. This gives you the resultant of those vectors.
- Two vectors are equal if their **i** bits are equal, and their **j** bits are equal.
- A unit vector has magnitude 1. To find a unit vector in a direction, you will need to divide by the magnitude.

5.2 More Complicated Problems

These are ones involving current, or wind, or the real world. Follow the steps as listed:

- **Draw a diagram**. Ensure each vector starts at the end of the previous one. Try to sketch all angles as accurately as you can and make your diagram large; it will make your life easier!
- Draw a straight line between the very beginning of your chain of vectors and the end. Draw a double arrowhead on it. This is the resultant.
- To nd the resultant, you will need to add the **i** components of the original vectors, and the **j** components of the original vectors.
- Use the resultant vector's **i** and **j** components to find its magnitude and direction (draw a diagram for the direction!).
- Don't forget, bearings are three-figure angles, taken clockwise from North.

The biggest difficulty with vector problems comes when then question asks on what course the captain SHOULD steer, or where the pilot *ought* to go. For these types of problems, the should / ought bit is the resultant, although we don't know the length (magnitude of it). Draw on the other two vectors (usually current / wind and the boat / plane itself) and use the sine rule for absolutely everything.

6 Projectiles

The key to every projectile question is to SPLIT HORIZONTAL AND VERTICAL MOTION. I advise you to follow the following steps, no matter what the style of the question:

- Write down initial velocity as a column vector.
- Write down acceleration (constant!) as a column vector.
- Write down expressions in terms of t for x and y (horizontal and vertical displacement) using $s = ut + \frac{1}{2}at^2$. Don't forget to add the initial height on to your y equation, if appropriate.

You should now be able to answer any projectile question, no matter how complex, as long as you remember the following few simple facts:

- The ball reaches its greatest height when the *vertical* component of the velocity is zero ($v = u + at$ vertically will give you what you need…).
- The ball lands when $y = 0$.
- If you find the time taken to land (i.e. solve a quadratic in t) then you can find the range (x). This just means the horizontal distance the ball travels before it lands.
- If the question involves a Cartesian equation for the path (or trajectory) of the ball, this is just a case of eliminating t as we have done so many times before.

7 Newton in 2D

Newton's 2$^{\text{nd}}$ Law is fairly simple when applied in a straight line (car and caravan problems and so on). However, it becomes a little more complicated in two or more dimensions, but we are OK as long as we remember the following:

- If there is no resultant force (i.e. body is at rest or moving with constant velocity) then the forces are in equilibrium.
- When the resultant is not zero, then there is an acceleration, and we can use Newton's 2$^{\text{nd}}$ all over again. Lucky us.

Once we know what the system is doing, we can solve problems in the following manner:

- Draw a diagram. Make sure it includes weight (vertically downwards), normal reaction if appropriate, and any tensions which need to be there.
- Write down an equation of motion for each item. *Always* resolve horizontally and vertically, unless…
- If an item is on a slope, *always* resolve in two directions, parallel and perpendicular to the slope.

8 General Motion

Don't forget:

- Constants of integration; if you are given initial velocity or position then just stick these in but you must somehow show that you are aware of their existence!
- Time constraints still apply even after differentiation or integration.
- Greatest displacement is when velocity is zero.
- Greatest velocity is when acceleration is zero.

OCR Mechanics 2 Module Revision Sheet

The M2 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J.M.S.

Moments

- Blah

Projectiles

- Blah

Circular Motion

- ω is the angular speed of a particle; it is the rate of change of the angle at the centre of the circle $\omega \equiv \frac{d\theta}{dt}$. It is the arc length formula differentiated wrt time.

$$\begin{aligned} S &= r\theta \\ \frac{d}{dt}(S) &= \frac{d}{dt}(r\theta) \\ \frac{d}{dt}(S) &= r\frac{d}{dt}(\theta) \text{ because } r \text{ is constant} \\ v &= r\omega. \end{aligned}$$

  Note that the formula $S = r\theta$ is only valid for radians, so ω is measured in *radians* per second.[40]

- The acceleration of a particle travelling in a circle is

$$a = \frac{v^2}{r} = r\omega^2.$$

  The acceleration is towards the centre of the circle. [Make sure you can derive $a = \frac{v^2}{r}$ if you are going for Oxbridge.] Since $F = ma$ we also have

$$F = \frac{mv^2}{r} = mr\omega^2.$$

[40] Also worth noting $T = \frac{2\pi}{\omega}$ and $T = \frac{1}{f}$.

OCR Mechanics 3 Module Revision Sheet

The M3 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted. . . always an area prone to error.

J.M.S.

Blah

- Blah

OCR Mechanics 4 Module Revision Sheet

The M4 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J.M.S.

Blah

- Blah

OCR Statistics 1 Module Revision Sheet

The S1 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted. . . always an area prone to error.

J.M.S.

"Without data, all you are is just another person with an opinion."

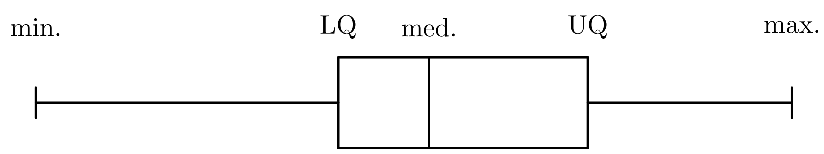
Representation Of Data

- You must be happy constructing unordered, back-to-back and ordered stem and leaf diagrams. They show the overall distribution of the data and back-to-back diagrams allow you to compare two sets of data.
- Cumulative frequency graphs. The cumulative frequency is a "running total" of the frequencies as you go up the values. For example

x	f				x (upper limit of)	cum. freq
$0 \leqslant x < 5$	8	$\Rightarrow$	Create	$\Rightarrow$	5	8
$5 \leqslant x < 10$	13	$\Rightarrow$	Cumulative	$\Rightarrow$	10	21
$10 \leqslant x < 15$	17	$\Rightarrow$	Frequency	$\Rightarrow$	15	38
$15 \leqslant x < 20$	10				20	48

  Plot the second of these tables and join it with a smooth curve to form the *cumulative frequency curve*. From this the median and the two quartiles can be found.

- Once these values are found we can draw a *box and whisker diagram*. The box and whisker diagram uses five values: the minimum, the maximum, the lower quartile, the upper quartile and the median. It is good for showing spread and comparing two quantities.



- Histograms are usually drawn for continuous data in classes. If the classes have equal widths, then you merely plot amount against frequency.
- If the classes do *not* have equal widths then we need to create a new column for *frequency density*. Frequency density is defined by f.d. $= \frac{\text{frequency}}{\text{class width}}$. The *area* of the bars are what represents the frequency, *not* the height.
- Frequency polygons are made by joining together the mid-points of the bars of a histogram with a ruler.

Measures Of Location

- The *mean* (arithmetic mean) of a set of data $\{x_1, x_2, x_3 \dots x_n\}$ is given by
$$\overline{x} = \frac{\text{sum of all values}}{\text{the number of values}} = \frac{\sum x}{n}.$$
When finding the mean[41] of a frequency distribution the mean is given by
$$\frac{\sum(xf)}{\sum f} = \frac{\sum(xf)}{n}.$$

- If a set of numbers is arranged in ascending (or descending) order the *median* is the number which lies half way along the series. It is the number that lies at the $\left(\frac{n+1}{2}\right)^{\text{th}}$ position. Thus the median of $\{13, 14, 15, 15\}$ lies at the $2\frac{1}{2}$ position $\Rightarrow$ average of 14 and 15 $\Rightarrow$ median = 14.5.

- The *mode* of a set of numbers is the number which occurs the most frequently. Sometimes no mode exists; for example with the set $\{2, 4, 7, 8, 9, 11\}$. The set $\{2, 3, 3, 3, 4, 5, 6, 6, 6, 7\}$ has two modes 3 and 6 because each occurs three times. One mode $\Rightarrow$ "unimodal". Two modes $\Rightarrow$ "bimodal". More than two modes $\Rightarrow$ "multimodal".

	Advantages	Disadvantages
Mean	⋆ The best known average. ⋆ Can be calculated exactly. ⋆ Makes use of all the data. ⋆ Can be used in further statistical work.	⋆ Greatly affected by extreme values. ⋆ Can't be obtained graphically. ⋆ When the data are discrete can give an impossible figure (2.34 children).
Median	⋆ Can represent an actual value in the data. ⋆ Can be obtained even if some of the values in a distribution are unknown. ⋆ Unaffected by irregular class widths and unaffected by open-ended classes. ⋆ Not influenced by extreme values.	⋆ For grouped distributions its value can only be estimated from an ogive. ⋆ When only a few items available or when distribution is irregular the median may not be characteristic of the group. ⋆ Can't be used in further statistical calculations.
Mode	⋆ Unaffected by extreme values. ⋆ Easy to calculate. ⋆ Easy to obtain from a histogram.	⋆ May exist more than one mode. ⋆ Can't be used for further statistical work. ⋆ When the data are grouped its value cannot be determined exactly.

Measures Of Spread

- The simplest measure of spread is the *range*. Range = $x_{\text{max}} - x_{\text{min}}$.

- The interquartile range is simply the upper quartile take away the lower quartile. Both of these values are usually found from a cumulative frequency graph (above).

- The *sum of squares from the mean* is called the *sum of squares* and is denoted
$$S_{xx} = \sum(x - \overline{x})^2 = \sum x^2 - n\overline{x}^2.$$

[41] Statistics argues that the average person has one testicle and that 99.999% of people have more than the average number of arms...

For example given the data set $\{3, 6, 7, 8\}$ the mean is 6; $\sum x^2 = 9 + 36 + 49 + 64 = 158$; so $S_{xx} = \sum x^2 - n\overline{x}^2 = 158 - 4 \times 6^2 = 14$.[42]

- The *standard deviation* (σ) is defined: $\sigma = \sqrt{\text{variance}} = \sqrt{\dfrac{S_{xx}}{n}} = \sqrt{\dfrac{\sum x^2}{n} - \overline{x}^2}$.

- *Example*: Given the set of data $\{5, 7, 8, 9, 10, 10, 14\}$ calculate the standard deviation. Firstly we note that $\overline{x} = 9$.

$$\sigma = \sqrt{\frac{\sum x^2}{n} - \overline{x}^2} = \sqrt{\frac{(5^2 + \cdots + 14^2)}{7} - 9^2}$$
$$= \sqrt{\frac{615}{7} - 81} = 2.6186\ldots$$

- When dealing with frequency distributions such as

x	1	2	3	4	5
f	4	5	7	5	4

, we *could* calculate σ by writing out the data[43] and carrying out the calculations as above, but this is clearly slow and inefficient. To our rescue comes a formula for σ that allows direct calculation from the table. This is

$$\sigma = \sqrt{\frac{\sum(x^2 f)}{n} - \overline{x}^2}.$$

- *Example*: Calculate mean and sd for the above frequency distribution. For easy calculation we need to add certain columns to the usual x and f columns thus;

x	f	xf	$x^2 f$
1	4	4	4
2	5	10	20
3	7	21	63
4	5	20	80
5	4	20	100
	$n = \sum f = 25$	$\sum(xf) = 75$	$\sum(x^2 f) = 267.$

So $\overline{x} = \dfrac{\sum(xf)}{n} = \dfrac{75}{25} = 3$ and $\sigma = \sqrt{\dfrac{\sum(x^2 f)}{n} - \overline{x}^2} = \sqrt{\dfrac{267}{25} - 3^2} = 1.2961\ldots$

- *Linear Coding*. Given the set of data $\{2, 3, 4, 5, 6\}$ we can see that $\overline{x} = 4$ and it can be calculated that $\sigma = 1.414$ (3dp). If we add 20 to all the data points we can see that the mean becomes 24 and the standard deviation will be unchanged. If the data set is multiplied by 3 we can see that the mean becomes 12 and the standard deviation would become three times as large (4.743 (3dp)).

- If, instead of being given $\sum x$ and $\sum x^2$, you were given $\sum(x - a)$ and $\sum(x - a)^2$ for some constant a, you just use the substitution $u = x - a$ and use $\sum u$ and $\sum u^2$ to work out the mean of u and the standard deviation of u. Then, using the above paragraph, we know $\bar{x} = \bar{u} + a$ and $\sigma_x = \sigma_u$.

[42] Or we could have done $S_{xx} = \Sigma(x - \overline{x})^2 = (3-6)^2 + (6-6)^2 + (7-6)^2 + (8-6)^2 = 14$.
[43] $\{1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5\}$!!!

Probability

- An *independent event* is one which has no effect on subsequent events. The events of spinning a coin and then cutting a pack of cards are independent because the way in which the coin lands has no effect on the cut. For two *independent* events A & B

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(A) \times \mathbb{P}(B).$$

For example a fair coin is tossed and a card is then drawn from a pack of 52 playing cards. Find the probability that a head and an ace will result.

$$\mathbb{P}(\text{head}) = \tfrac{1}{2}, \qquad \mathbb{P}(\text{ace}) = \tfrac{4}{52} = \tfrac{1}{13}, \qquad \text{so } \mathbb{P}(\text{head and ace}) = \tfrac{1}{2} \times \tfrac{1}{13} = \tfrac{1}{26}.$$

- *Mutually Exclusive Events.* Two events which cannot occur at the same time are called mutually exclusive. The events of throwing a 3 or a 4 in a single roll of a fair die are mutually exclusive. For any two mutually exclusive events

$$\mathbb{P}(A \text{ or } B) = \mathbb{P}(A) + \mathbb{P}(B).$$

For example a fair die with faces of 1 to 6 is rolled once. What is the probability of obtaining either a 5 or a 6?

$$\mathbb{P}(5) = \tfrac{1}{6}, \qquad \mathbb{P}(6) = \tfrac{1}{6}, \qquad \text{so } \mathbb{P}(5 \text{ or } 6) = \tfrac{1}{6} + \tfrac{1}{6} = \tfrac{1}{3}.$$

- *Non-Mutually Exclusive Events.* When two events can both happen they are called non-mutually exclusive events. For example studying English and studying Maths at A Level are non-mutually exclusive. By considering a Venn diagram of two events A & B we find

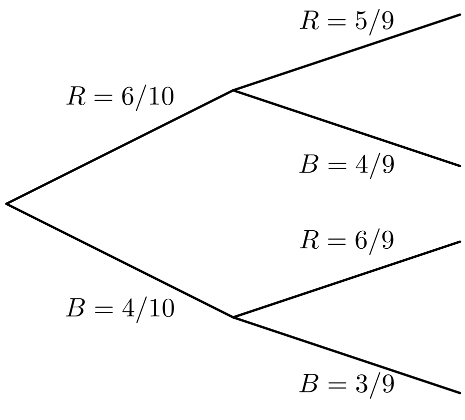
$$\mathbb{P}(A \text{ or } B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \text{ and } B),$$
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

- *Tree Diagrams.* These may be used to help solve probability problems when more than one event is being considered. The probabilities on any branch section must sum to one. You multiply along the branches to discover the probability of that branch occurring.

For example a box contains 4 black and 6 red pens. A pen is drawn from the box and it is not replaced. A second pen is then drawn. Find the probability of

(i) two red pens being obtained.
(ii) two black pens being obtained.
(iii) one pen of each colour being obtained.
(iv) two red pens *given* that they are the same colour.

Draw tree diagram to discover:



(i) $\mathbb{P}(\text{two red pens}) = \frac{6}{10} \times \frac{5}{9} = \frac{30}{90} = \frac{1}{3}.$

(ii) $\mathbb{P}(\text{two black pens}) = \frac{4}{10} \times \frac{3}{9} = \frac{12}{90} = \frac{2}{15}.$

(iii) $\mathbb{P}(\text{one of each colour}) = 1 - \frac{30}{90} - \frac{12}{90} = \frac{8}{15}.$

(iv) $\mathbb{P}(\text{two reds} \mid \text{same colour}) = \frac{1/3}{1/3+2/15} = \frac{5}{7}.$

- *Conditional Probability.* In the above example we see that the probability of two red pens is $\frac{1}{3}$, but the probability of two red pens *given that both pens are the same colour* is $\frac{5}{7}$. This is known as conditional probability. $\mathbb{P}(A \mid B)$ mean the probability of A *given* that B has happened. It is governed by

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

  For example if there are 120 students in a year and 60 study Maths, 40 study English and 10 study both then

$$\mathbb{P}(\text{study English} \mid \text{study Maths}) = \frac{\mathbb{P}(\text{study Maths \& English})}{\mathbb{P}(\text{study Maths})} = \frac{10/120}{60/120} = \frac{1}{6}.$$

- A is independent of B if $\mathbb{P}(A) = \mathbb{P}(A \mid B) = \mathbb{P}(A \mid B')$. (i.e. whatever happens in B the probability of A remains unchanged.) For example flicking a coin and then cutting a deck of cards to try and find an ace are independent because

$$\mathbb{P}(\text{cutting ace}) = \mathbb{P}(\text{cutting ace} \mid \text{flick head}) = \mathbb{P}(\text{cutting ace} \mid \text{flick tail}) = \tfrac{1}{13}.$$

Permutations And Combinations

- Factorials are defined $n! \equiv n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$. Many expressions involving factorials simplify with a bit of thought. For example $\frac{n!}{(n-2)!} = n(n-1)$. Also there is a convention that $0! = 1$.

- The number of ways of arranging n different objects in a line is $n!$ For example how many different arrangements are there if 4 different books are to be placed on a bookshelf? There are 4 ways in which to select the first book, 3 ways in which to choose the second book, 2 ways to pick the third book and 1 way left for the final book. The total number of different ways is $4 \times 3 \times 2 \times 1 = 4!$

- Permutations. The number of ways of selecting r objects from n when *the order of the selection matters* is ${}^nP_r$. It can be calculated by

$${}^nP_r = \frac{n!}{(n-r)!}.$$

  For example in how many ways can the gold, silver and bronze medals be awarded in a race of ten people? The order in which the medals are awarded matters, so the number of ways is given by ${}^{10}P_3 = 720$.

  In another example how many words of four letters can be made from the word CONSIDER? This is an arrangement of four out of eight different objects where the order matters so there are ${}^8P_4 = \frac{8!}{4!} = 1680$ different words.

- Combinations. The number of ways of selecting r objects from n when *the order of the selection does not matter* is ${}^nC_r$. It can be calculated by

$${}^nC_r \equiv \binom{n}{r} = \frac{n!}{r!\,(n-r)!}.$$

  For example in how many ways can a committee of 5 people be chosen from 8 applicants? Solution is given by ${}^8C_5 = \frac{8!}{5! \times 3!} = 56$.

  In another example how many ways are there of selecting your lottery numbers (where one selects 6 numbers from 49)? It does not matter which order you choose your numbers, so there are ${}^{49}C_6 = 13\,983\,816$ possible selections.

- If letters are repeated in a 'word', then you just divide through by the factorials of each repeat. Therefore there are $\frac{11!}{4!\times 4!\times 2!}$ arrangements of the word 'MISSISSIPPI'.[44]

- You must be good at 'choosing committee' questions [be on the lookout, they can be in disguise]. For example how many ways are there of choosing a committee of 3 women and 4 men from a group containing 10 women and 5 men? There are $\binom{10}{3}$ ways of choosing the women (the order doesn't matter) and $\binom{5}{4}$ ways of choosing the men. Therefore overall there are $\binom{5}{4} \times \binom{10}{3}$ ways of choosing the committee.

- Example: If I deal six cards from a standard deck of cards, in how many ways can I get exactly four clubs? Well there are $\binom{13}{4}$ ways of getting the clubs, and $\binom{39}{2}$ ways of getting the non-clubs, so therefore the answer to the original question is $\binom{13}{4} \times \binom{39}{2}$.

- When considering lining things up in a line we start from the principle that there are $n!$ ways of arranging n objects. In the harder examples you need to be a cunning.

  For example three siblings join a queue with 5 other people making 8 in total.

  1. How many way are there of arranging the 8 in a queue? Easy; 8!
  2. How many ways are there of arranging the 8, such that the siblings are together? We, we imaging the three siblings tied together. There are therefore 6! ways of arranging the 5 and the bundle of siblings and then there are 3! ways of arranging the siblings in the bundle. Therefore the answer is $6! \times 3!$
  3. How many ways are there of arranging the siblings so they are not together? There are 5! ways of arranging the five without the siblings. There are then 6 places for the first sibling to go, 5 for the second, and 4 for the third. Therefore $5! \times 6 \times 5 \times 4$.

- To calculate probabilities we go back to first principles and remember that probability is calculated from the number of ways of getting what we want over the total number of possible outcomes. So in the above example, if the 8 are arranged randomly in a line, what is the probability of the siblings being together? $\mathbb{P}(\text{together}) = \frac{6!\times 3!}{8!}$.

  Going back to the four club question if it asked for the probability of getting exactly four clubs if I dealt exactly six cards from the pack, the answer would be $\frac{\binom{13}{4}\times\binom{39}{2}}{\binom{52}{6}}$. The $\binom{52}{6}$ represents the total number of ways I can deal six cards from the 52.

Probability Distributions

- A random variable is a quantity whose value depends on chance. The outcome of a random variable is usually denoted by a capital letter (e.g. X). We read $\mathbb{P}(X = 2)$ as the probability that the random variable takes the value 2. For a fair die, $\mathbb{P}(X = 5) = \frac{1}{6}$.

- For discrete random variables they are usually presented in a table. For example for a fair die:

x	1	2	3	4	5	6
$\mathbb{P}(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

- In general, for any event, the probability distribution is of the form

x	x_1	x_2	x_3	x_4	x_5	x_6	$\cdots$
$\mathbb{P}(X = x)$	p_1	p_2	p_3	p_4	p_5	p_6	$\cdots$

[44] As an non-mathematical aside, find two fruits that are anagrams of each other.

- As before, it is crucial that we remember the probabilities sum to one. This can be useful at the start or problems where a constant must be evaluated. For example in:

x	1	2	3	4
$\mathbb{P}(X = x)$	k	k	$2k$	$4k$

We discover $k + k + 2k + 4k = 1$, so $k = \frac{1}{8}$.

Binomial And Geometric Distributions

- The **binomial distribution** is applicable when you have fixed number of *repeated, independent* 'trials' such that each trial can be viewed as 'success' (p) or 'fail' ($q = 1 - p$). In justifying a binomial distribution you must not just quote the previous sentence; you must apply it to the situation in the question. For example: "Binomial is applicable because the probability of each tulip flowering is independent of each other tulip and the probability of flowering is a constant".

- For example if I throw darts at a dart board and my chance of hitting a double is 0.1 and I throw 12 darts at the board and my chance of hitting a double is independent of all the other throws then a binomial distribution will be applicable. We let X be the number of doubles I hit. X can therefore take the values $\{0, 1, 2 \ldots, 11, 12\}$; i.e. there are 13 possible outcomes. $p = 0.1$; the probability of success and $q = 1 - p = 0.9$; the probability of failure. We write $X \sim B(n, p)$ which here is $X \sim B(12, 0.1)$.

- I would always advise you to visualise the tree diagram. From this we can 'see' that $\mathbb{P}(X = 12) = 0.1^{12}$ and $\mathbb{P}(X = 0) = 0.9^{12}$. In general

$$\mathbb{P}(X = x) = \binom{n}{x} \times p^x \times q^{n-x}.$$

So in the example, the probability I hit exactly 7 doubles is $\mathbb{P}(X = 7) = \binom{12}{7} \times 0.1^7 \times 0.9^5$.

- For questions such as $\mathbb{P}(X \leqslant 5)$ or $\mathbb{P}(X \geqslant 8)$ you must be able to use the tables in the formula book. The tables always give $\mathbb{P}(X \leqslant \text{something})$. You must be able to convert probabilities to this form and then read off from the table. For $X \sim B(10, 0.35)$.

$$\begin{aligned}
&\mathbb{P}(X \leqslant 7) = 0.9952,\\
&\mathbb{P}(X < 5) = \mathbb{P}(X \leqslant 4) = 0.7515,\\
&\mathbb{P}(X \geqslant 7) = 1 - \mathbb{P}(X \leqslant 6) = 1 - 0.9740 = 0.0260,\\
&\mathbb{P}(X > 3) = 1 - \mathbb{P}(X \leqslant 3) = 1 - 0.5138 = 0.4862.
\end{aligned}$$

- The **geometric distribution** is applicable when you are looking for how long you wait until an event has occurred. The events must be *repeated, independent and success/fail.* Potentially you could wait forever until a success occurs; something to look for if you are unsure what distribution to apply. Similar to the binomial you must justify *in the context of the question.*

- Going back to the darts example, we could rephrase it as how long must I wait until I hit a double? Let X be the number of throws until I hit a double. We write $X \sim \text{Geo}(0.1)$. X can take the values $\{1, 2, 3, \ldots\}$.

- Obviously $\mathbb{P}(X = 1) = 0.1$. Less obviously $\mathbb{P}(X = 4) = 0.9^3 \times 0.1$ (I must have three failures and *then* my success). In general

$$\mathbb{P}(X = x) = q^{x-1} \times p.$$

- There are no tables for the geometric distribution because there does not need to be. To calculate $\mathbb{P}(X \geqslant 5)$ we must have had 4 failures. Therefore $\mathbb{P}(X \geqslant 5) = q^4 = (1-p)^4$. Also to calculate $\mathbb{P}(X \leqslant 6)$ we use the fact that $\mathbb{P}(X \leqslant 6) = 1-\mathbb{P}(X \geqslant 7) = 1-q^6 = 1-(1-p)^6$. In general

$$\mathbb{P}(X \geqslant x) = (1-p)^{x-1} \qquad \text{and} \qquad \mathbb{P}(X \leqslant x) = 1-(1-p)^x.$$

Expectation And Variance Of A Random Variable

- The expected value of the event is denoted $\mathbb{E}(X)$ or μ. It is defined

$$\mathbb{E}(X) = \mu = \boxed{\sum x\mathbb{P}(X = x)}.$$

For example for a fair die with

x	1	2	3	4	5	6
$\mathbb{P}(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

we find:

$$\begin{aligned}\mathbb{E}(X) &= \left(1 \times \tfrac{1}{6}\right) + \left(2 \times \tfrac{1}{6}\right) + \left(3 \times \tfrac{1}{6}\right) + \left(4 \times \tfrac{1}{6}\right) + \left(5 \times \tfrac{1}{6}\right) + \left(6 \times \tfrac{1}{6}\right)\\ &= 3\tfrac{1}{2}.\end{aligned}$$

- The variance of an event is denoted $\text{Var}(X)$ or σ^2 and is defined

$$\text{Var}(X) = \sigma^2 = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \mathbb{E}(X^2) - \mu^2 = \boxed{\sum x^2\mathbb{P}(X = x) - \mu^2}.$$

So for the *biased* die with distribution

x	1	2	3	4	5	6
$\mathbb{P}(X = x)$	$\frac{1}{3}$	$\frac{1}{6}$	0	0	$\frac{1}{6}$	$\frac{1}{3}$

we find that

$$\mathbb{E}(X) = \left(1 \times \tfrac{1}{3}\right) + \left(2 \times \tfrac{1}{6}\right) + (3 \times 0) + (4 \times 0) + \left(5 \times \tfrac{1}{6}\right) + \left(6 \times \tfrac{1}{3}\right) = 3\tfrac{1}{2}$$

and

$$\begin{aligned}\text{Var}(X) &= \sum x^2\mathbb{P}(X = x) - \mu^2\\ &= \left(1^2 \times \tfrac{1}{3}\right) + \left(2^2 \times \tfrac{1}{6}\right) + \left(3^2 \times 0\right) + \left(4^2 \times 0\right) + \left(5^2 \times \tfrac{1}{6}\right) + \left(6^2 \times \tfrac{1}{3}\right) - 3\tfrac{1}{2}^2\\ &= 17\tfrac{1}{6} - 3\tfrac{1}{2}^2 = 4\tfrac{11}{12}.\end{aligned}$$

- The expectation of a binomial distribution $B(n,p)$ is np. The variance of $B(n,p)$ is npq.
- The expectation of a geometric distribution $\text{Geo}(p)$ is $\frac{1}{p}$.

Correlation

- The Product Moment Correlation Coefficient is a number (r) calculated on a set of bivariate data that tells us how correlated two data sets are.

- The value of r is such that $-1 < r < 1$. If $r = 1$ you have perfect positive linear correlation. If $r = -1$ you have perfect negative linear correlation. If $r = 0$ then there exists no correlation between the data sets.

- It is defined
$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$
where we define the individual components as
$$S_{xx} = \sum x^2 - \tfrac{1}{n}\left(\sum x\right)^2,$$
$$S_{yy} = \sum y^2 - \tfrac{1}{n}\left(\sum y\right)^2,$$
$$S_{xy} = \sum xy - \tfrac{1}{n}\sum x \sum y.$$

- So to calculate r for the data set

x	14	12	16	18	21	13	15	17
y	1	2	4	5	2	8	5	6

we write the data in columns and add extra ones. We then sum the columns and calculate from these sums. Note that in the above example $n = 8$ (i.e. the number of pairs, not the number of individual data pieces).

x	y	x^2	y^2	xy
14	1	196	1	14
12	2	144	4	24
16	4	256	16	64
18	5	324	25	90
21	2	441	4	42
13	8	169	64	104
15	5	225	25	75
17	6	289	36	102
126	**33**	**2044**	**175**	**515**

Therefore
$$S_{xx} = \sum x^2 - \tfrac{1}{n}\left(\sum x\right)^2 = 2044 - \tfrac{126^2}{8} = 59.5,$$
$$S_{yy} = \sum y^2 - \tfrac{1}{n}\left(\sum y\right)^2 = 175 - \tfrac{33^2}{8} = 38.875,$$
$$S_{xy} = \sum xy - \tfrac{1}{n}\sum x \sum y = 515 - \tfrac{126 \times 33}{8} = -4.75.$$
Therefore
$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{-4.75}{\sqrt{59.5 \times 38.875}} = -0.09876\ldots$$
Therefore the data has very, very weak negative correlation. Basically it has no *meaningful* correlation.

- It can be shown that if one (or both) of the variables are transformed in a linear fashion i.e. if we replace the x values by, say, $\frac{x-4}{3}$ (or any transformation formed by $+$, $-$, $\div$ or $\times$ with constants) then the value of r will be unchanged.

- You need to be able to calculate Spearman's rank correlation coefficient (r_s). You will be given a table and you will need to (in the next 2 columns) rank the data. If two data points are tied then you (e.g. the 2nd and 3rd are tied) then you rank them both 2.5.

%	IQ	Rank %	Rank IQ	d	d^2
89	143	2.5	1	1.5	2.25
55	89	7	8	-1	1
72	102	5	6	-1	1
91	136	1	2	-1	1
89	126	2.5	3	-0.5	0.25
30	60	9	9	0	0
71	115	6	4	2	4
53	100	8	7	1	1
78	103	4	5	-1	1

Now $\boxed{r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}}$. $\sum d^2$ is just the sum of the d^2 column in the table and n is the number of pairs of data; here $n = 9$. We therefore find $r_s = 1 - \frac{6\times 11.5}{9(81-1)} = 0.9041\dot{6}$. Therefore we see a strong degree of positive association.

- If r_s is close to -1 then strong negative association. If close to zero then no meaningful association/agreement.

Regression

- For any set of bivariate data (x_i, y_i) there exist two possible regression lines; 'y on x' and 'x on y'.
- If neither is controlled (see below) then if you want to predict y from a given value of x, you use the 'y on x' line. If you want to predict x from a given value of y, you use the 'x on y' line.
- The 'y on x' line is defined

$$y = a + bx \quad \text{where} \quad b = \frac{S_{xy}}{S_{xx}} \quad \text{and} \quad a = \bar{y} - b\bar{x}.$$

- The 'x on y' line is defined

$$x = a' + b'y \quad \text{where} \quad b' = \frac{S_{xy}}{S_{yy}} \quad \text{and} \quad a' = \bar{x} - b'\bar{y}.$$

- Both regression lines pass through the average point $(\bar{x}, \bar{y})$.
- In the example in the book (P180) the height of the tree is the dependent variable and the circumference of the tree is the independent variable. This is because the experiment has been constructed to see how the height of the tree depends on its circumference.
- If one variable is being controlled by the experimenter (e.g. x), it is called a controlled variable. If x is controlled you would never use the 'x on y' regression line. Only use the 'y on x' line. You would use this to predict y from x (expected) *and* x from y (not-expected)

OCR Statistics 2 Module Revision Sheet

The S2 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J.M.S.

Continuous Random Variables

- A continuous random variable (crv) is usually described by means of a probability density function (pdf) which is defined for all real x. It must satisfy

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \quad \text{and} \quad f(x) \geqslant 0 \text{ for all } x.$$

- Probabilities are represented by areas under the pdf. For example the probability that X lies between a and b is

$$\mathbb{P}(a < X < b) = \int_a^b f(x)\,dx.$$

  It is worth noting that for any specific value of X, $\mathbb{P}(X = \text{value}) = 0$ because the area of a single value is zero.

- The median is the value m such that

$$\int_{-\infty}^{m} f(x)\,dx = \frac{1}{2}.$$

  That is; the area under the curve is cut in half at the value of the median. Similarly the lower quartile (Q_1) and upper quartile (Q_3) are defined

$$\int_{-\infty}^{Q_1} f(x)\,dx = \frac{1}{4} \quad \text{and} \quad \int_{-\infty}^{Q_3} f(x)\,dx = \frac{3}{4}.$$

- The expectation of X is defined

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x)\,dx.$$

  Compare this to the discrete definition of $\sum x\mathbb{P}(X = x)$. Always be on the lookout for symmetry in the distribution before carrying out a long integral; it could save you a lot of time. You should therefore always sketch the distribution if you can.

- The variance of X is defined

$$\text{Var}(X) = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2.$$

  Again, compare this to the discrete definition of $\sum x^2\mathbb{P}(X = x) - \mu^2$. Don't forget to subtract μ^2 at the end; someone always does!

- The main use for this chapter is to give you the basics you may need for the normal distribution. The normal distribution is by far the most common crv.

The Normal Distribution

- The normal distribution (also known as the Gaussian distribution[45]) is the most common crv. It is found often in nature; for example daffodil heights, human IQs and pig weights can all be modelled by the normal curve. A normal distribution can be summed up by two parameters; its mean (μ) and its variance (σ^2). For a random variable X we say $X \sim N(\mu, \sigma^2)$.

- As with all crvs probabilities are given by areas; i.e. $\mathbb{P}(a < X < b) = \int_a^b f(x)\,dx$. However the $f(x)$ for a normal distribution is complicated and impossible to integrate exactly. We therefore need to use tables to help us. Since there are an infinite number of $N(\mu, \sigma^2)$ distributions we use a special one called the standard normal distribution. This is $Z \sim N(0, 1^2)$.

- The tables given to you work out the areas to the left of a value. The notation used is $\Phi(z) = \int_{-\infty}^{z} f(z)\,dz$. So $\Phi(0.2)$ is the area to the left of 0.2 in the standard normal distribution. The tables do not give Φ(negative value) so there are some tricks of the trade you must be comfortable with. These and they are always helped by a sketch and remembering that the area under the whole curve is one. For example

$$\Phi(z) = 1 - \Phi(-z)$$
$$\mathbb{P}(Z > z) = 1 - \Phi(z)$$

- Real normal distributions are related to the standard distribution by

$$Z = \frac{X - \mu}{\sigma} \qquad (\dagger).$$

So if $X \sim N(30, 16)$ and we want to answer $\mathbb{P}(X > 24)$ we convert $X = 24$ to $Z = (24 - 30)/4 = -1.5$ and answer $\mathbb{P}(Z > -1.5) = \mathbb{P}(Z < 1.5) = 0.9332$.

- Another example; If $Y \sim N(100, 5^2)$ and we wish to calculate $\mathbb{P}(90 < Y < 105)$. Converting to $\mathbb{P}(-2 < Z < 1)$ using †. Then finish off with

$$\mathbb{P}(-2 < Z < 1) = \Phi(1) - \Phi(-2) = \Phi(1) - (1 - \Phi(2)) = 0.8413 - (1 - 0.9772) = 0.8185.$$

- You must also be able to do a 'reverse' lookup from the table. Here you don't look up an area from a z value, but look up a z value from an area.

  For example find a such that $\mathbb{P}(Z < a) = 0.65$. Draw a sketch as to what this means; to the left of some value a the area is 0.65. Therefore, reverse looking up we discover $a = 0.385$.

- Harder example; Find b such that $\mathbb{P}(Z > b) = 0.9$. Again a sketch shows us that the area to the right of b must be 0.9, so b must be negative. Considering the sketch carefully, we discover $\mathbb{P}(Z < -b) = 0.9$, so reverse look up tells us $-b = 1.282$, so $b = -1.282$.

- Reverse look up is then combined with † in questions like this. For $X \sim N(\mu, 5^2)$ it is known $\mathbb{P}(X < 20) = 0.8$; find μ. Here you will find it easier if you draw both a sketch for the X and also for Z and marking on the important points. The z value by reverse look up is found to be 0.842. Therefore by † we obtain, $0.842 = (20 - \mu)/5$, so $\mu = 15.79$.

[45] I *do* wish we would call it the Gaussian distribution. Carl Friedrich Gauss. Arguably the greatest mathematician ever. German...

- Harder example; $Y \sim (\mu, \sigma^2)$ you know $\mathbb{P}(Y < 20) = 0.25$ and $\mathbb{P}(Y > 30) = 0.4$. You should obtain two † equations;

$$-0.674 = \frac{20-\mu}{\sigma} \quad \text{and} \quad 0.253 = \frac{30-\mu}{\sigma} \quad \Rightarrow \quad \mu = 27.27 \text{ and } \sigma = 10.79.$$

- The binomial distribution can sometimes be approximated by the normal distribution. If $X \sim B(n,p)$ and $np > 5$ and $nq > 5$ then we can use $V \sim N(np, npq)$ as an approximation. Because we are going from a discrete distribution to a continuous, a continuity correction must be used.

- For example if $X \sim B(90, \frac{1}{3})$ we can see $np = 30 > 5$ and $nq = 60 > 5$ so we can use $V \sim N(30, 20)$. Some examples of the conversions:

$$\begin{aligned} \mathbb{P}(X = 29) &\approx \mathbb{P}(28.5 < V < 29.5), \\ \mathbb{P}(X > 25) &\approx \mathbb{P}(V > 25.5), \\ \mathbb{P}(5 \leqslant X < 40) &\approx \mathbb{P}(4\tfrac{1}{2} < V < 39\tfrac{1}{2}). \end{aligned}$$

The Poisson Distribution

- The Poisson distribution is a discrete random variable (like the binomial or geometric distribution). It is defined

$$\mathbb{P}(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}.$$

X can take the values $0, 1, 2, \ldots$ and the probabilities depend on only one parameter, λ. Therefore we find

x	0	1	2	3	...
$\mathbb{P}(X = x)$	$e^{-\lambda}\frac{\lambda^0}{0!}$	$e^{-\lambda}\frac{\lambda^1}{1!}$	$e^{-\lambda}\frac{\lambda^2}{2!}$	$e^{-\lambda}\frac{\lambda^3}{3!}$	...

- For a Poisson distribution $\mathbb{E}(X) = \text{Var}(X) = \lambda$. We write $X \sim \text{Po}(\lambda)$.

- As for the binomial we use tables to help us and they are given (for various different λs) in the form $\mathbb{P}(X \leqslant x)$. So if $\lambda = 5$ and we wish to discover $\mathbb{P}(X < 8)$ we do $\mathbb{P}(X < 8) = \mathbb{P}(X \leqslant 7) = 0.8666$. Also note that if we want $\mathbb{P}(X \geqslant 4)$ we would use the fact that probabilities sum to one, so $\mathbb{P}(X \geqslant 4) = 1 - \mathbb{P}(X \leqslant 3) = 1 - 0.2650 = 0.7350$.

- The Poisson distribution can be used as an approximation to the binomial distribution provided $n > 50$ and $np < 5$. If these conditions are met and $X \sim B(n,p)$ we use $W \sim \text{Po}(np)$. [No continuity correction required since we are approximating a discrete by a discrete.]

- For example with $X \sim B(60, \frac{1}{30})$ both conditions are met and we use $W \sim \text{Po}(2)$. Therefore some example of some calculations:

$$\begin{aligned} \mathbb{P}(X \leqslant 3) &\approx \mathbb{P}(W \leqslant 3) = 0.8571 \text{ (from tables)} \\ \mathbb{P}(3 < X \leqslant 7) &\approx \mathbb{P}(3 < W \leqslant 7) = \mathbb{P}(W \leqslant 7) - \mathbb{P}(W \leqslant 3) = 0.9989 - 0.8571 = 0.1418. \end{aligned}$$

- The normal distribution can be used as an approximation to the to the Poisson distribution if $\lambda > 15$. So if $X \sim \text{Po}(\lambda)$ we use $Y \sim N(\lambda, \lambda)$. However, here we *are* approximating a discrete by a continuous, so a continuity correction must be applied.

- For example if $X \sim \text{Po}(50)$ we can use $Y \sim N(50, 50)$ since $\lambda > 15$. To calculate $\mathbb{P}(X = 49)$ we would calculate (using $Z = (X - \mu)/\sigma$)

$$\begin{aligned}\mathbb{P}(X = 49) \approx \mathbb{P}(48.5 < Y < 49.5) &= \mathbb{P}(-0.212 < Z < -0.071)\\ &= \mathbb{P}(0.071 < Z < 0.212)\\ &= \Phi(0.212) - \Phi(0.071)\\ &= 0.5840 - 0.5283 = 0.0557.\end{aligned}$$

Similarly

$$\begin{aligned}\mathbb{P}(X < 55) &\approx \mathbb{P}(Y < 54.5)\\ &= \mathbb{P}\left(Z < \frac{54.5 - 50}{\sqrt{50}}\right)\\ &= \mathbb{P}(Z < 0.6364)\\ &= 0.738.\end{aligned}$$

Sampling

- If a sample is taken from an underlying population you can view the mean of this sample as a random variable in its own right. This is a subtle point and you should dwell on it! If you can't get to sleep sometime, you should lie awake thinking about it. (I had to.)

- If the underlying population has $\mathbb{E}(X) = \mu$ and $\text{Var}(X) = \sigma^2$, then the distribution of the mean of the sample, $\bar{X}$, is

$$\mathbb{E}(\bar{X}) = \mu \text{ (the same as the underlying) and } \text{Var}(\bar{X}) = \frac{\sigma^2}{n}.$$

This means that the larger your sample, the less likely it is that the mean of this sample is a long way from the population mean. So if you are taking a sample, make it as big as you can!

- If your sample is sufficiently large (roughly > 30) the central limit theorem (CLT) states that the distribution of the sample mean is approximated by

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

no matter what the underlying distribution is.

- If the underlying population is discrete you need to include a $\frac{1}{2n}$ correction factor when using the CLT. For example $\mathbb{P}(\bar{X} > 3.4)$ for a discrete underlying with a sample size of 45 would mean you calculate $\mathbb{P}(\bar{X} > 3.4 + \frac{1}{90})$.

- If the underlying population is a normal distribution then no matter how large the sample is (e.g. just 4) we can say

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

- If you have the whole population data available to you then to calculate the mean you use $\mu = \frac{\sum x}{n}$ and to calculate the variance you use

$$\sigma^2 = \frac{\sum x^2}{n} - \bar{x}^2 = \frac{\sum x^2 - n\bar{x}^2}{n}.$$

However you do not usually have all the data. It is more likely that you merely have a sample from the population. From this sample you may want to estimate the population mean and variance. As you would expect your best estimate of the population mean is the mean of the sample $\frac{\sum x}{n}$. However the best estimate of the population variance is not the variance of the sample. You must calculate s^2 where

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

Some textbooks use $\hat{\sigma}$ to mean s; they both mean 'the unbiased estimator of the population σ'. So

$$\text{(Estimate of population variance)} = \frac{n}{n-1} \times \text{(Sample variance)}.$$

- You could be given raw data ($\{x_1, x_2, \ldots x_n\}$) in which you just do a direct calculation. Or summary data ($\sum x^2, \sum x$ and n). Or you could be given the sample variance and n. From all of these you should be able to calculate s^2. It should be clear from the above section how to do this.

Continuous Hypothesis Testing

- In *any* hypothesis test you will be testing a 'null' hypothesis H_0 against an 'alternative' hypothesis H_1. In S2, your H_0 will only *ever* be one of these three:

  $H_0 : p =$something

  $H_0 : \lambda =$something

  $H_0 : \mu =$something

  Don't deviate from this and you can't go wrong. Notice that it does *not* say $H_0 = p =$something.

- The book gives three approaches to continuous hypothesis testing, but they are all essentially the same. You always compare the probability of what you have seen (under H_0) and anything more extreme, and compare this probability to the significance level. If it is less than the significance level, then you reject H_0 and if it is greater, then you accept H_0.

- Remember we connect the real (X) world to the standard (Z) world using $Z = \frac{X-\mu}{\sigma}$.

- You can do this by:

  1. Calculating the probability of the observed value and anything more extreme and comparing to the significance level.
  2. Finding the critical Z-values for the test and finding the Z-value for the observed event and comparing. (e.g. critical Z-values of 1.96 and -1.96; if observed Z is 1.90 we accept H_0; if observed is -2.11 the reject H_0.)
  3. Finding the critical values for $\bar{X}$. For example critical values might be 17 and 20. If X lies between them then accept H_0; else reject H_0.

- Example: P111 Que 8. Using method 3 from above.

  Let X be the amount of magnesium in a bottle. We are told $X \sim N(\mu, 0.18^2)$. We are taking a sample of size 10, so $\bar{X} \sim N(\mu, \frac{0.18^2}{10})$. Clearly

$$\begin{aligned} H_0 &: \mu = 6.8 \\ H_1 &: \mu \neq 6.8. \end{aligned}$$

We proceed assuming H_0 is correct. Under H_0, $\bar{X} \sim N(6.8, \frac{0.18^2}{10})$. This is a 5% two-tailed test, so we need $2\frac{1}{2}$% at each end of our normal distribution. The critical Z values are (by reverse lookup) $Z_{\text{crit}} = \pm 1.960$. To find how these relate to $\bar{X}_{\text{crit}}$ we convert thus

$$\begin{aligned} Z_{\text{crit}} &= \frac{\bar{X}_{\text{crit}} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \\ 1.960 &= \frac{\bar{X}_{\text{crit}} - 6.8}{\sqrt{\frac{0.18^2}{10}}} \\ \text{and} \quad -1.960 &= \frac{\bar{X}_{\text{crit}} - 6.8}{\sqrt{\frac{0.18^2}{10}}} \end{aligned}$$

These solve to $\bar{X}_{\text{crit}} = 6.912$ and $\bar{X}_{\text{crit}} = 6.688$. The observed $\bar{X}$ is 6.92 which lies just outside the acceptance region. We therefore reject H_0 and conclude that the amount of magnesium per bottle is probably *different* to 6.8. [The book is in error in claiming that we conclude it is bigger than 6.8.]

Discrete Hypothesis Testing

- For any test with discrete variables, it is usually best to find the critical value(s) for the test you have set and hence the critical region. The critical value is the first value *at which you would reject* the null hypothesis.

- For example if testing $X \sim B(16, p)$ we may test (at the 5% level)

$$\begin{aligned} \mathrm{H}_0 &: p = \tfrac{5}{6} \\ \mathrm{H}_1 &: p < \tfrac{5}{6}. \end{aligned}$$

  We are looking for the value at the lower end of the distribution (remember the "<" acts as an arrow telling us where to look in the distribution). We find $\mathbb{P}(X \leqslant 11) = 0.1134$ and $\mathbb{P}(X \leqslant 10) = 0.0378$. Therefore the critical value is 10. Thus the critical region is $\{0, 1, 2 \ldots 9, 10\}$. So when the result for the experiment is announced, if it lies in the critical region, we reject H_0, else accept H_0.

- Another example: If testing $X \sim B(20, p)$ at the 10% level with

$$\begin{aligned} \mathrm{H}_0 &: p = \tfrac{1}{6} \\ \mathrm{H}_1 &: p \neq \tfrac{1}{6}. \end{aligned}$$

  Here we have a two tailed test with 5% at either end of the distribution. At the lower end we find $\mathbb{P}(X = 0) = 0.0261$ and $\mathbb{P}(X \leqslant 1) = 0.1304$ so the critical value is 0 at the lower end. At the upper end we find $\mathbb{P}(X \leqslant 5) = 0.8982$ and $\mathbb{P}(X \leqslant 6) = 0.9629$. Therefore

$$\begin{aligned} \mathbb{P}(X \geqslant 6) &= 1 - \mathbb{P}(X \leqslant 5) = 1 - 0.8982 = 0.1018 \\ \mathbb{P}(X \geqslant 7) &= 1 - \mathbb{P}(X \leqslant 6) = 1 - 0.9629 = 0.0371 \end{aligned}$$

  So at the upper end we find $X = 7$ to be the critical value. [Remember that at the upper end, the critical value is always one more than the upper of the two values where the gap occurs; here the gap was between 5 and 6 in the tables, so 7 is the critical value.] The critical region is therefore $\{0, 7, 8 \ldots 20\}$.

- There is a Poisson example in the 'Errors in hypothesis testing' section.

Errors In Hypothesis Testing

- A Type I error is made when a true null hypothesis is rejected.
- A Type II error is made when a false null hypothesis is accepted.
- For continuous hypothesis tests, the $\mathbb{P}$(Type I error) is just the significance level of the test. [This fact should be obvious; if not think about it harder!]
- For a Type II error, you must consider something like the example on page 140/1 which is superbly explained. From the original test, you will have discovered the acceptance and the rejection region(s). When you are told the real mean of the distribution and asked to calculate the $\mathbb{P}$(Type II error), you must use the new, real mean and the old standard deviation (with a new normal distribution; e.g. $N(\mu_{\text{new}}, \sigma^2_{\text{old}}/n)$) and work out the probability that the value lies within the old acceptance region. [Again, the book is *very* good on this and my explanation is poor.]
- For discrete hypothesis tests, the $\mathbb{P}$(Type I error) is not merely the stated significance level of the test. The stated value (e.g. 5%) is merely the 'notional' value of the test. The true significance level of the test (and, therefore, the $\mathbb{P}$(Type I error)) is the probability of all the values in the rejection region, given the truth of the null hypothesis.

  For example in a binomial hypothesis test we might have discovered the rejection region was $X \leqslant 3$ and $X \geqslant 16$. If the null hypothesis was "H_0: $p = 0.3$", then the true significance level of the test would be $\mathbb{P}(X \leqslant 3 \text{ or } X \geqslant 16 \mid p = 0.3)$.
- To calculate $\mathbb{P}$(Type II error) you would, given the true value for p (or λ for Poisson), calculate the probability of the *complementary* event. So in the above example, if the true value of p was shown to be 0.4, you would calculate $\mathbb{P}(3 < X < 16 \mid p = 0.4)$.
- Worked example for Poisson: A hypothesis is carried out to test the following:

$$\begin{aligned} \text{H}_0 &: \lambda = 7 \\ \text{H}_1 &: \lambda \neq 7 \\ \alpha &= 10\% \\ &\quad \text{Two tailed test.} \end{aligned}$$

  Under H_0, $X \sim \text{Po}(7)$. We discover the critical values are $X = 2$ and $X = 13$. The critical region is therefore $X \leqslant 2$ and $X \geqslant 13$.

  Therefore $\mathbb{P}$(Type I error) and the true value of the test is therefore

$$\begin{aligned} \mathbb{P}(X \leqslant 2 \text{ or } X \geqslant 13 \mid \lambda = 7) &= \mathbb{P}(X \leqslant 2) + \mathbb{P}(X \geqslant 13) \\ &= \mathbb{P}(X \leqslant 2) + 1 - \mathbb{P}(X \leqslant 12) \\ &= 0.0296 + 1 - 0.9730 \\ &= 0.0566 = 5.66\%. \end{aligned}$$

  Given that the true value of λ was shown to be 10, then $\mathbb{P}$(Type II error) would be

$$\begin{aligned} \mathbb{P}(2 < X < 13 \mid \lambda = 10) &= \mathbb{P}(X \leqslant 12) - \mathbb{P}(X \leqslant 2) \\ &= 0.7916 - 0.0028 \\ &= 0.7888 = 78.88\%. \end{aligned}$$

OCR Statistics 3 Module Revision Sheet

The S3 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J .M .S .

Preliminaries

- In S1 when calculating the variance you will mostly have used $\sigma^2 = \frac{\sum x^2}{n} - \bar{x}^2$. This was for ease of calculation. However in S3 the equivalent formula $\sigma^2 = \frac{\sum(x-\bar{x})^2}{n}$ appears to make a storming comeback. You will often be given $\sum(x-\bar{x})^2$ summary data and you must know how to handle it.
- The unbiased estimator of variance from a sample (s^2) simplifies to

$$s^2 \equiv \frac{n}{n-1}\left(\frac{\sum x^2}{n} - \bar{x}^2\right) = \frac{n}{n-1}\left(\frac{\sum(x-\bar{x})^2}{n}\right) = \frac{\sum(x-\bar{x})^2}{n-1}.$$

- Because of this, if you need to calculate (for a two sample t-test) $s_p^2 = \frac{(n_x-1)s_x^2+(n_y-1)s_y^2}{n_x+n_y-2}$ and are given $\sum(x-\bar{x})^2$ and $\sum(y-\bar{y})^2$ then s_p^2 simplifies thus

$$\begin{aligned} s_p^2 &= \frac{(n_x-1)s_x^2+(n_y-1)s_y^2}{n_x+n_y-2} \\ &= \frac{(n_x-1)\left(\frac{\sum(x-\bar{x})^2}{n_x-1}\right)+(n_y-1)\left(\frac{\sum(y-\bar{y})^2}{n_y-1}\right)}{n_x+n_y-2} \\ &= \frac{\sum(x-\bar{x})^2+\sum(y-\bar{y})^2}{n_x+n_y-2}. \end{aligned}$$

Continuous Random Variables

- In S2 you met probability density functions (pdf) $f(x)$. They measured where events were more likely to occur than others. To find $\mathbb{P}(a < X < b)$ we needed to calculate the area between $x = a$ and $x = b$; i.e. $\int_a^b f(x)\,dx$. In S3 we have cumulative distribution functions (cdf) $F(x)$ which are defined $F(x) \equiv \mathbb{P}(X \leqslant x)$. We can think of $F(x)$ as the area to the left of x in the pdf. So $F(4)$ is the area to the left of 4 and $F(3)$ is the area to the left of 3. Therefore $\mathbb{P}(3 < X < 4) = F(4) - F(3)$. This is an example of:

$$\mathbb{P}(a < X < b) = F(b) - F(a).$$

- Cdfs make calculating the median (M) very easy. You just solve $F(M) = \frac{1}{2}$. Likewise the upper (Q_3) and lower (Q_1) quartiles are very easy to calculate; $F(Q_1) = \frac{1}{4}$ and $F(Q_3) = \frac{3}{4}$.

You must understand the concept of percentiles and how to get them from a cdf. The 85th percentile (say) is such that 85% of the data lies to the left of that point. Therefore $F(P_{85}) = \frac{85}{100}$.

- You cannot write
$$\int_1^x x^2\,dx = \left[\frac{x^3}{3}\right]_1^x = \frac{x^3}{3} - \frac{1}{3}.$$
You must use a dummy variable thus:
$$\int_1^x t^2\,dt = \left[\frac{t^3}{3}\right]_1^x = \frac{x^3}{3} - \frac{1}{3}.$$
Basically whenever you find yourself putting an x on the upper limit of an integral, change all future x's to t's.

- To calculate $f(x)$ from $F(x)$ is easy; just differentiate $F(x)$. For example given
$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^3}{27} & 0 \leqslant x \leqslant 3 \\ 1 & x > 3. \end{cases}$$
When we differentiate the constants 0 and 1 they become 0. The $\frac{x^3}{27}$ becomes $\frac{x^2}{9}$ so the pdf is
$$f(x) = \begin{cases} \frac{x^2}{9} & 0 \leqslant x \leqslant 3 \\ 0 & \text{otherwise.} \end{cases}$$

- To calculate $F(x)$ from $f(x)$ is a little trickier. You must remember that $F(x)$ is the *entire* area to the left of a point. Therefore given
$$f(x) = \begin{cases} k & 0 \leqslant x < 2 \\ k(x-1) & 2 \leqslant x \leqslant 3 \\ 0 & \text{otherwise.} \end{cases}$$
Firstly we calculate[46] $k = \frac{2}{7}$. For the section $0 \leqslant x < 2$ we do the expected $\int_0^x \frac{2}{7}\,dt = \left[\frac{2}{7}t\right]_0^x = \frac{2}{7}x$. However, for the next region we *do not* just do $\int_2^x \frac{2}{7}(x-1)\,dt$. We need to *add in* the contribution from the first part (i.e. the value of $F(2)$ from the first result; $\frac{4}{7}$ in this case). So we do $\frac{4}{7} + \int_2^x \frac{2}{7}(t-1)\,dt = \frac{1}{7}(x^2 - 2x + 4)$. Therefore
$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{2}{7}x & 0 \leqslant x < 2 \\ \frac{1}{7}(x^2 - 2x + 4) & 2 \leqslant x \leqslant 3 \\ 1 & x > 3. \end{cases}$$

- Once you have calculated your $F(x)$ a nice check to see whether your cdf is correct is to see if your $F(x)$ is continuous[47] *which it must be.* For example let's say you discovered that
$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{3}x & 0 \leqslant x < 1 \\ x^2 - \frac{5}{2}x + 2 & 1 \leqslant x \leqslant 2 \\ 1 & x > 2. \end{cases}$$
You then check the 'boundary' values where the functions are being joined; here they are $x = 0$, $x = 1$ and $x = 2$. In this case there is no problem for $x = 0$ nor $x = 2$, but when we look at $x = 1$ there is a problem. $\frac{1}{3}x$ gives $\frac{1}{3}$ but $x^2 - \frac{5}{2}x + 2$ gives $\frac{1}{2}$. Therefore we must have made a mistake which must be fixed.

[46] By remembering $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

[47] A function is continuous if you can draw it without taking your pen off the paper... basically.

- Given a cdf $F(x)$ you can find a related cdf $F(y)$ where X and Y are related; i.e $Y = g(X)$. The idea here is that $F(x) \equiv \mathbb{P}(X \leqslant x)$. Start with the original cdf. Then write $F(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(X \leqslant g^{-1}(y))$. Then replace *every* x in the original cdf by $g^{-1}(y)$ (even the ones in the limits).

  For example given

$$F(x) = \begin{cases} 0 & x < 2 \\ \frac{1}{8}(x^2 - 2x) & 2 \leqslant x \leqslant 4 \\ 1 & x > 4. \end{cases}$$

  and $Y = 4X^2$ find $F(y)$.

  So, $F(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(4X^2 \leqslant y) = \mathbb{P}(X \leqslant \frac{\sqrt{y}}{2})$ (we don't have to worry about $\pm$ when square rooting because the cdf is only defined for positive x). Therefore

$$F(y) = \begin{cases} 0 & \frac{\sqrt{y}}{2} < 2 \\ \frac{1}{8}((\frac{\sqrt{y}}{2})^2 - 2(\frac{\sqrt{y}}{2})) & 2 \leqslant \frac{\sqrt{y}}{2} \leqslant 4 \\ 1 & \frac{\sqrt{y}}{2} > 4. \end{cases}$$

  And so

$$F(y) = \begin{cases} 0 & y < 16 \\ \frac{1}{32}(y - 4\sqrt{y}) & 16 \leqslant y \leqslant 64 \\ 1 & y > 64. \end{cases}$$

- You must be a little careful if (say) $Y = \frac{1}{X}$. You start $F(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(\frac{1}{X} \leqslant y) = \mathbb{P}(X \geqslant \frac{1}{y}) = 1 - \mathbb{P}(X \leqslant \frac{1}{y})$. Notice this reversal of the inequality sign; this is because if $\frac{a}{b} > \frac{c}{d}$ then $\frac{b}{a} < \frac{d}{c}$.

- In S2 expectation for a pdf $f(x)$ is $\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)\,dx$. In S3 you can find the expectation of any function $g(X)$ of the pdf $f(x)$ by the formula

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)\,dx.$$

  For example find $\mathbb{E}(X^2 + 1)$ of

$$f(x) = \begin{cases} \frac{e^{x-1}}{e-1} & 1 \leqslant x \leqslant 2 \\ 0 & \text{otherwise.} \end{cases}$$

  So,

$$\begin{aligned} \mathbb{E}(X^2 + 1) &= \int_1^2 (x^2 + 1)\frac{e^{x-1}}{e-1}\,dx \\ &= \frac{1}{e-1}\int_1^2 x^2e^{x-1} + e^{x-1}\,dx \\ &= \text{int by parts twice on first bit\dots good exercise for you to do\dots} \\ &= \frac{3e-2}{e-1}. \end{aligned}$$

Linear Combinations Random Variables

- Any random variable X can be transformed to become a new random variable $Y = aX + b$ where a and b are constants. It can be shown that

$$\mathbb{E}(Y) = \mathbb{E}(aX + b) = a\mathbb{E}(X) + b.$$

It can also be shown that

$$\mathrm{Var}(Y) = \mathrm{Var}(aX + b) = a^2\mathrm{Var}(X).$$

The b 'disappears' because it only has the effect of moving X up or down the number line and does not therefore alter the spread (i.e. variance). Note also that the a gets squared when one 'pulls it out' of the variance. Therefore $\mathrm{Var}(-2X) = (-2)^2\mathrm{Var}(X) = 4\mathrm{Var}(X)$. It also makes sense with $\mathrm{Var}(-X) = (-1)^2\mathrm{Var}(X) = \mathrm{Var}(X)$ because if one makes all the values of X negative from where they were they are just as spread out.

- Take any two random variables X and Y. If they are combined in a linear fashion $aX+bY$ for constant a and b then it is ***always true*** (even when X and Y are not independent) that

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y).$$

If X and Y are *independent* then

$$\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y).$$

It is particularly useful to note that $\mathrm{Var}(X-Y) = \mathrm{Var}(X)+\mathrm{Var}(Y)$. These results extend (rather obviously) to more than two variables

$$\mathbb{E}(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = a_1\mathbb{E}(X_1) + a_2\mathbb{E}(X_2) + + \cdots + a_n\mathbb{E}(X_n),$$
$$\mathrm{Var}(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = a_1^2\mathrm{Var}(X_1) + a_2^2\mathrm{Var}(X_2) + \cdots + a_n^2\mathrm{Var}(X_n).$$

The second (of course) true if all independent.

- If X and Y are *independent* and normally distributed then $aX + bY$ is also normally distributed. Because $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$ and $\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y)$ we find

$$X \sim N(\mu_1, \sigma_1^2) \quad \text{and} \quad Y \sim N(\mu_2, \sigma_2^2) \qquad \Rightarrow \qquad aX + bY \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2).$$

For example when Jon throws a shot put his distance is $J \sim N(11, 4)$. When Ali throws a shot his distance is $A \sim N(12, 9)$. Find the probability on one throw that Jon beats Ali. So we need $J - A \sim N(11 - 12, 4 + 9)$ which gives $J - A \sim N(-1, 13)$. Notice the variances have been added and that the expected value is negative (on average Jon will lose to Ali). Now

$$\begin{aligned}\mathbb{P}(J - A > 0) &= \mathbb{P}\left(Z > \frac{0 - (-1)}{\sqrt{13}}\right)\\ &= \mathbb{P}(Z > 0.277)\\ &= 1 - \mathbb{P}(Z < 0.277) = 0.3909\end{aligned}$$

- Given a random variable X you must fully appreciate the difference between two *independent* samplings of this random variable (X_1 and X_2) and two times this random variable ($2X$). For example given a random variable X such that

x	1	2
$\mathbb{P}(X = x)$	$\frac{1}{2}$	$\frac{1}{2}$

The random variable $2X$ is doubling the outcome of *one* sampling of X, but $X_1 + X_2$ is adding *two* independent samplings of X. Thus $2X$ can *only* take values 2 and 4 with probabilities $\frac{1}{2}$ each. But $X_1 + X_2$ can take values 2, 3 and 4 with probabilities $\frac{1}{4}$, $\frac{1}{2}$ and

$\frac{1}{4}$ respectively. Note that the expected values for $2X$ and $X_1 + X_2$ are the same (because $\mathbb{E}(2X) = 2\mathbb{E}(X)$ and $\mathbb{E}(X_1 + X_2) = \mathbb{E}(X_1) + \mathbb{E}(X_2) = 2\mathbb{E}(X)$), but that the variances are *not* the same; i.e. $\text{Var}(2X) \neq \text{Var}(X_1 + X_2)$. This is because $\text{Var}(2X) = 4\text{Var}(X)$ and $\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) = 2\text{Var}(X)$.

For example given the above shot put example $J \sim N(11, 4)$. If Jon was to throw the shot put three times (independently) and the total of all three throws recorded we would need $J_1 + J_2 + J_3 \sim N(33, 3 \times 4)$ and **not** $3J \sim N(33, 9 \times 4)$.

- Given Poisson distributed X and Y it is even simpler. Here $aX + bY$ is not distributed Poisson[48]. However the special case of $X + Y$ *is* distributed Poisson.

$$X \sim Po(\lambda_1) \quad \text{and} \quad Y \sim Po(\lambda_2) \qquad \Rightarrow \qquad X + Y \sim Po(\lambda_1 + \lambda_2).$$

For example if Candy makes on average 3 typing errors per hour and Tiffany makes 4 typing errors per hour find the probability of fewer than 12 errors in total in a two hour period. Here we have $Po(14)$ so $\mathbb{P}(X < 12) = \mathbb{P}(X \leqslant 11) = 0.2600$ (tables).

Student's *t*-Distribution

- In S2 you learnt that if you take a sample from a normal population of *known variance* σ^2 then no matter what the sample size $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ <u>exactly</u>.

  The test statistic for $\text{H}_0 : \mu = c$ is $Z = \dfrac{\bar{X} - c}{\sqrt{\frac{\sigma^2}{n}}}$.

- You also learnt that if you take a sample of size $n > 30$ from *any* population distribution where you know σ^2 then (by CLT) $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ <u>approximately</u>.

  The test statistic for $\text{H}_0 : \mu = c$ is $Z = \dfrac{\bar{X} - c}{\sqrt{\frac{\sigma^2}{n}}}$.

- You also learnt that if you take a sample of size $n > 30$ from *any* population distribution with unknown σ^2 then you estimate σ^2 by calculating s^2 and (by CLT) $\bar{X} \sim N\left(\mu, \frac{s^2}{n}\right)$ <u>approximately</u>.

  The test statistic for $\text{H}_0 : \mu = c$ is $Z = \dfrac{\bar{X} - c}{\sqrt{\frac{s^2}{n}}}$.

- You would therefore think that if you were drawing from a normal population with unknown σ^2 then you would estimate σ^2 by calculating s^2 and $\bar{X} \sim N\left(\mu, \frac{s^2}{n}\right)$. But ***this is not the case!!!*** In fact $\bar{X}$ is <u>exactly</u> described by Student's t-distribution[49].

  The test statistic for $\text{H}_0 : \mu = c$ is $T = \dfrac{\bar{X} - c}{\sqrt{\frac{s^2}{n}}}$.

[48]Because with the Poisson we require the expectation and the variance to be the same and given $X \sim Po(\lambda_1)$ and $Y \sim Po(\lambda_2)$ we have $\mathbb{E}(aX+bY) = a\mathbb{E}(X)+b\mathbb{E}(Y) = a\lambda_1 + b\lambda_2$ and $\text{Var}(aX+bY) = a^2\text{Var}(X)+b^2\text{Var}(Y) = a^2\lambda_1 + b^2\lambda_2$ and the only time $aX + bY = a^2X + b^2Y$ is when $a = b = 1$.

[49]Named after W.S.Gosset who wrote under the pen name 'Student'. Gosset devised the t-test as a way to cheaply monitor the quality of stout. Good bloke.

- (You will notice the apparent contradiction between the last two bullet points. If a large sample ($n > 30$) is taken from a normal population with unknown variance then how can $\bar{X}$ be distributed *both* normally and as a t-distribution? Well, as the sample size gets larger, the t-distribution converges to the normal distribution. Just remember that *technically* if you have a normal population with unknown variance then $\bar{X}$ is *exactly* a t-distribution, but if $n > 30$ then CLT lets us *approximate* $\bar{X}$ as a normal. In practice the t-distribution is used only with small sample sizes.)

- There is the new concept of the degree of freedom (denoted ν) of the t-distribution. As ν gets larger the t-distribution tends towards the standard normal distribution. However if ν is small enough, then the difference between t and z becomes quite marked (as you can see yourself from the tables).

- We can do hypothesis tests here just like we did in S2, only instead of using the normal tables we use the t tables (with correct degrees of freedom ν) to find t_{crit} and compare the test statistic $\dfrac{\bar{X} - c}{\sqrt{\frac{s^2}{n}}}$ against t_{crit}. Here $\nu = n - 1$.

- For example a machine is producing circular disks whose radius is normally distributed. Their radius historically has been 5cm. The factory foreman believes that the machine is now producing disks that are too small. A sample of 9 disks are taken and their radii are

  $$4.8, \quad 4.9, \quad 4.5, \quad 5.2, \quad 4.9, \quad 4.8, \quad 5.0, \quad 4.8, \quad 5.0$$

  Test at the 10% level whether the foreman has a case.

  Let $\mu =$ the population mean radii of the disks.

  $\text{H}_0 : \mu = 5$,

  $\text{H}_1 : \mu < 5$.

  $n = 9$, so $\nu = 9 - 1 = 8$.

  $\alpha = 10\%$. Therefore in t_8 we lookup 90% (because one tailed) and discover 1.397. But because it is a "<" test t_{crit} must be negative to $t_{\text{crit}} = -1.397$.

  $\bar{x} = \frac{\sum x}{n} = \frac{43.9}{9} = 4.8\dot{7}$.

  $s^2 = \frac{n}{n-1}\left(\frac{\sum x^2}{n} - \bar{x}^2\right) = \frac{9}{8}\left(\frac{214.43}{9} - 4.8\dot{7}^2\right) = 0.03694$.

  $t_{\text{obs}} = \frac{\bar{x}-c}{\sqrt{\frac{s^2}{n}}} = \frac{4.8\dot{7}-5}{\sqrt{\frac{0.03694}{9}}} = -1.908$.

  $-1.908 < -1.397$. This value lies in the rejection region of the test and therefore at the 10% level we have sufficient evidence to reject H_0 and conclude that the machine is probably not working fine.

Testing For Difference Between Means

- The central pillar in this section is that if $\bar{X} \sim N\left(\mu_x, \frac{\sigma_x^2}{n_x}\right)$ (which is either exactly true if X is itself normal, or approximately true if $n_x > 30$ from CLT) and $\bar{Y} \sim N\left(\mu_y, \frac{\sigma_y^2}{n_y}\right)$ then (provided X and Y are independent)

  $$\bar{X} - \bar{Y} \sim N\left(\mu_x - \mu_y, \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}\right).$$

- If X and Y are *normally* distributed with *known* variances (σ_x^2 and σ_y^2) and we are testing $\mathrm{H}_0 : \mu_x - \mu_y = c$ the test statistic is

$$Z = \frac{\bar{X} - \bar{Y} - c}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}.$$

For example[50] it is known that French people's heights (in cm) are normally distributed $N(\mu_f, 25)$. It is also known that German people's heights are normally distributed $N(\mu_g, 20)$. It is wished to test whether or not German people are taller than French people (at the $2\frac{1}{2}\%$ level). A random sample of 10 French people's heights are and their mean height recorded ($\bar{f}$). Similarly 8 German people's heights are taken and their mean recorded ($\bar{g}$).

1. State appropriate null and alternative hypotheses.
2. Find the set of values for $\bar{g} - \bar{f}$ for which we would reject the null hypothesis.
3. If in fact Germans are 7cm taller on average then find the probability of a Type II error.

1. $\mathrm{H}_0 : \mu_g - \mu_f = 0$,
   $\mathrm{H}_1 : \mu_g - \mu_f > 0$.
2. Given $Z = \frac{\bar{X} - \bar{Y} - c}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}$ we obtain

$$Z_{\text{crit}} = 1.960 = \frac{(\bar{G} - \bar{F})_{\text{crit}}}{\sqrt{\frac{25}{10} + \frac{20}{8}}}.$$

   Therefore critical value is $(\bar{g} - \bar{f})_{\text{crit}} = 4.383$. We therefore reject the null hypothesis if $\bar{g} - \bar{f} \geqslant 4.383$.
3. For a Type II error we must lie in the *acceptance region* of the original test given the new information. Here we require $\mathbb{P}(\bar{g} - \bar{f} < 4.383 \mid \mu_g - \mu_f = 7)$, so

$$\begin{aligned}\mathbb{P}(\bar{g} - \bar{f} < 4.383 \mid \mu_g - \mu_f = 7) &= P\left(Z < \frac{4.383 - 7}{\sqrt{\frac{25}{10} + \frac{20}{8}}}\right)\\ &= \mathbb{P}(Z < -1.170)\\ &= 1 - \mathbb{P}(Z < 1.170)\\ &= 1 - 0.8790 = 0.121\end{aligned}$$

- If X and Y are *not* normally distributed we need the samples to be *large* (then CLT applies). If the variances are *known* then the above is still correct. However if the population variances are unknown we replace the σ_x and σ_y by their estimators s_x and s_y.

For example, Dr. Evil believes that people's attention spans are different in Japan and America. He samples 80 Japanese people and finds their attention spans are described (in minutes) $\sum j = 800$ and $\sum j^2 = 12000$. He samples 100 people in America and finds $\sum a = 850$ and $\sum a^2 = 11200$. Test at the 5% level whether Dr Evil is justified in his claim. So

$\mathrm{H}_0 : \mu_j - \mu_a = 0$.

$\mathrm{H}_1 : \mu_j - \mu_a \neq 0$.

[50] It's well worth thinking very hard about this example. It stumped me the first time I saw a similar question.

$\alpha = 5\%$.

$\bar{j} = 10$, $\bar{a} = 8.5$.

$s_j^2 = \frac{80}{79}\left(\frac{12000}{80} - 10^2\right) = 50.63$.

$s_a^2 = \frac{100}{99}\left(\frac{11200}{100} - 8.5^2\right) = 40.15$.

$Z_{\text{obs}} = \frac{\bar{X}-\bar{Y}-c}{\sqrt{\frac{s_j^2}{n_j}+\frac{s_a^2}{n_a}}} = \frac{10-8.5}{\sqrt{\frac{50.63}{80}+\frac{40.15}{100}}} = 1.475$.

$Z_{\text{crit}} = \pm 1.960$. Therefore we reject if $|Z_{\text{obs}}| > 1.960$.

$1.475 < 1.960$, so at the 5% level we have no reason to reject H_0 and conclude that Dr Evil is probably mistaken in his claim that the two countries have different attention spans.

- If X and Y are *normally* distributed with an *unknown,* ***common*** variance and we are testing $H_0 : \mu_x - \mu_y = c$ we use a two-sample t-test. The test statistic here is

$$T = \frac{\bar{X} - \bar{Y} - c}{\sqrt{s_p^2\left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}.$$

Here s_p^2 is the unbiased pooled estimate of the *common* variance, defined

$$s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2}.$$

Also $\nu = n_x + n_y - 2$. For example a scientist wishes to test whether new heart medication reduces blood pressure. 10 patients with high blood pressure were given the medication and their summary data is $\sum x = 1271$ and $\sum(x-\bar{x})^2 = 640.9$. 8 patients with high blood pressure were given a placebo and their summary data is $\sum y = 1036$ and $\sum(y-\bar{y})^2 = 222$. Carry out a hypothesis test at the 10% level to see if the medication is working.

$\text{H}_0 : \mu_x - \mu_y = 0$.

$\text{H}_1 : \mu_x - \mu_y < 0$.

$\alpha = 10\%$.

$\bar{x} = 127.1$, $\bar{y} = 129.5$.

$s_x^2 = \frac{10}{9}\left(\frac{640.9}{10}\right) = 71.2\dot{1}$.

$s_y^2 = \frac{8}{7}\left(\frac{222}{8}\right) = 31.71$.

$s_p^2 = \frac{9\times 71.2\dot{1}+7\times 31.71}{16} = 53.93$.

$T_{\text{obs}} = \frac{\bar{X}-\bar{Y}-c}{\sqrt{s_p^2\left(\frac{1}{n_x}+\frac{1}{n_y}\right)}} = \frac{127.1-129.5}{\sqrt{53.93\left(\frac{1}{10}+\frac{1}{8}\right)}} = -0.689$.

$\nu = 16$ so $T_{\text{crit}} = -1.337$.

$-0.689 > -1.337$, so at the 10% level we have no reason to reject H_0 and conclude that the medication is probably not lowering blood pressure.

- Also look for 'paired' data. This can only happen if $n_x = n_y$ *and* if every piece of data in x is somehow linked to a piece of data in y. Ask yourself 'would it matter if you changed the ordering of the x_i but not the y_i?' If yes, then paired. If the data is paired then you create a new set of data $d_i = x_i - y_i$.

1. If the *population of differences* is distributed normally (or assumed to be distributed normally) then the test statistic for $H_0 : \mu_d = c$ is

$$T = \frac{\bar{D} - c}{\sqrt{\frac{s_d^2}{n}}} \qquad \text{with } \nu = n - 1.$$

For example, Dwayne believes that his mystical crystals can boost IQs. He takes 10 students and records their IQs before and after they have been 'blessed' by the crystals. The results are

Victim	1	2	3	4	5	6	7	8	9	10
IQ Before	107	124	161	89	96	120	109	98	147	89
IQ After	108	124	159	100	101	119	110	101	146	94

Test at the 5% level Dwayne's claim. The data is clearly paired and thus we create $d_i = IQ_{\text{after}} - IQ_{\text{before}}$ giving

$$1, \quad 0, \quad -2, \quad 11, \quad 5, \quad -1, \quad 1, \quad 3, \quad -1, \quad 5.$$

$\text{H}_0 : \mu_d = 0,$
$\text{H}_1 : \mu_d > 0.$
$\alpha = 5\%$
$\nu = 10 - 1 = 9.$
$\bar{d} = \frac{22}{10} = 2.2$
$s_d^2 = \frac{n}{n-1}\left(\frac{\sum d^2}{n} - \bar{d}^2\right) = \frac{10}{9}\left(\frac{188}{10} - 2.2^2\right) = 15.5\dot{1}.$
$T_{\text{obs}} = \frac{\bar{D}-c}{\sqrt{\frac{s_d^2}{n}}} = 1.766.$
$T_{\text{crit}} = 1.833$ (tables)
$1.766 < 1.833$ therefore at the 5% level no reason to reject H_0 and conclude that the crystals probably don't significantly increase IQ.

2. If the *population of differences* is not distributed normally, but the sample size is large, then CLT applies and the test statistic for $H_0 : \mu_d = c$ is

$$Z = \frac{\bar{D} - c}{\sqrt{\frac{s_d^2}{n}}}.$$

- If testing for differences in population *proportions* there are two cases, each requiring *independent, large* samples (CLT).

1. For $H_0 : p_x = p_y$ (i.e. no difference in population proportions) the test statistic is

$$Z = \frac{P_{sx} - P_{sy}}{\sqrt{pq\left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}.$$

Here p is the value of the *common* population proportion $p = \dfrac{x + y}{n_x + n_y}$. Also $p_{sx} = \frac{x}{n_x}$ and $p_{sy} = \frac{y}{n_y}$.

2. For $H_0 : p_x - p_y = c$ the test statistic is

$$Z = \frac{P_{sx} - P_{sy} - c}{\sqrt{\frac{P_{sx}Q_{sx}}{n_x} + \frac{P_{sy}Q_{sy}}{n_y}}}.$$

Here $q_{sy} = 1 - p_{sx}$ and $q_{sy} = 1 - p_{sy}$.

Confidence Intervals

- It has been described to me by someone I respect that a confidence interval is like an 'egg-cup' of a certain width that we throw down onto the number-line. Of all possible 'egg-cups' we want 90% (or some other percentage) of those egg cups to contain the true mean μ. This does not mean that a confidence interval has a 90% chance of containing the mean; it either contains the mean or it doesn't.

- A confidence interval is denoted $[a, b]$ which means $a < x < b$. In S3 we only consider symmetric confidence intervals about the sample mean (because $\bar{x}$ is an unbiased estimate of μ). They basically represent the acceptance region of a hypothesis test where $H_0 : \mu = \bar{x}$.

- To find the required z or t values in all of the following confidence intervals is easy. If you want (say) a 90% confidence interval then you (sort of) want to contain 90% of the data, so you must have 10% not contained which means that there must be 5% at each end of the distribution. Therefore you look up, either in the little table *beneath* the big normal table or in the correct line of the t table, 95%. This then gives you the z or t value to the left of which 95% of the data lies.

- This is fine for certain special values (90%, 95%, 99% etc.) and for the t-distribution this is all you can do. However for z values we can also do a 'reverse look-up' in the main normal tables to find more 'exotic' values. For example if I wanted a 78% confidence interval with z, then 11% would be in each end. Therefore I would reverse look-up 0.8900 *within* the main body of the table to find $z = 1.226$.

- If you are drawing from a normal *of known variance* σ^2 then the confidence interval will be
$$\left[\bar{x} - z\frac{\sigma}{\sqrt{n}}, \bar{x} + z\frac{\sigma}{\sqrt{n}}\right].$$
This result is true even for small sample sizes.

  For example, an α% confidence interval is calculated from a normal population whose variance is known to be 9. The sample size is 16 and the confidence interval is $[19.68675, 22.31325]$. Find α. The midpoint of the interval is 21. Therefore the confidence interval is $[21 - z\frac{3}{\sqrt{16}}, 21 + z\frac{3}{\sqrt{16}}]$. We can then solve $21 + z\frac{3}{\sqrt{16}} = 22.31325$ to find $z = 1.751$. A forward lookup in the table reveals 0.96. Therefore there exists 4% at either end, so $\alpha = 8$; i.e. it is an 92% confidence interval.

- If you are drawing from a normal *of unknown variance* then the confidence interval will be
$$\left[\bar{x} - t\frac{s}{\sqrt{n}}, \bar{x} + t\frac{s}{\sqrt{n}}\right].$$
The degrees of freedom here will be $\nu = n - 1$.

- If you are drawing from an unknown distribution then (provided $n > 30$ to invoke the CLT) then the confidence interval will be
$$\left[\bar{x} - z\frac{s}{\sqrt{n}}, \bar{x} + z\frac{s}{\sqrt{n}}\right].$$

- If, instead of means, we are taking a sample proportion then the confidence interval will be
$$\left[p_s - z\sqrt{\frac{p_s q_s}{n}}, p_s + z\sqrt{\frac{p_s q_s}{n}}\right].$$

- If instead of single samples we are looking for a confidence interval for the difference between two populations we use the following, depending on the situation.

1. Difference in means being zero from two normals of *known* variances

$$\left[\bar{x} - \bar{y} - z\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}, \bar{x} - \bar{y} + z\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}\right].$$

Or for difference in means $\bar{x} - \bar{y}$ being c,

$$\left[\bar{x} - \bar{y} - c - z\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}, \bar{x} - \bar{y} - c + z\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}\right].$$

This can also be used for non-normal populations of known variance if the samples are *large* (CLT).

2. The above can be altered if the samples are *large* (CLT) and the variances are not known to

$$\left[\bar{x} - \bar{y} - z\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}, \bar{x} - \bar{y} + z\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}\right].$$

3. Difference in means being zero from two normals of *the same, unknown* variance

$$\left[\bar{x} - \bar{y} - ts_p\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}, \bar{x} - \bar{y} + ts_p\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}\right].$$

Or for difference in means $\bar{x} - \bar{y}$ being c,

$$\left[\bar{x} - \bar{y} - c - ts_p\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}, \bar{x} - \bar{y} - c + ts_p\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}\right].$$

Here s_p is the unbiased pooled estimate of the *common* variance $s_p^2 = \frac{(n_x-1)s_x^2+(n_y-1)s_y^2}{n_x+n_y-2}$. The degrees of freedom is $\nu = n_x + n_y - 2$.

4. If dealing with difference in population proportions we use

$$\left[p_{sx} - p_{sy} - z\sqrt{\frac{p_{sx}q_{sx}}{n_x} + \frac{p_{sy}q_{sy}}{n_y}}, p_{sx} - p_{sy} + z\sqrt{\frac{p_{sx}q_{sx}}{n_x} + \frac{p_{sy}q_{sy}}{n_y}}\right].$$

χ^2-Tests

- χ^2 tests measure how good data fits a given distribution. The test statistic here is

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

Here O is the observed frequency and E the expected frequency. The larger X^2 becomes the more likely it is that the observed data does *not* come from the expected values that we have calculated.

- As with the t-distribution, the χ^2 distribution has a degree of freedom associated with it still denoted ν. This is calculated

$$\nu = \text{number of classes} - \text{number of constraints}.$$

- Given observed frequencies you need to calculate expected frequencies from theoretical probabilities. Expected frequencies are the expected probability times the total number of trials. The convention is that if an expected value is less than 5, then you combine with a larger expected value such that all values end up greater than 5. For example if you had

OBS	22	38	24	18	9	2	1	0
EXP	23.4	35.1	27.2	16.1	7.2	3.1	0.9	0.2

you would combine the final four columns to get

OBS	22	38	24	18	12
EXP	23.4	35.1	27.2	16.1	11.4

Because of this combining the total number of classes would be 5 and *not* 8.

- Fitting a Distribution
  - As with any hypothesis tests, the expected values are computed supposing that H_0 is correct. For example given the data

Outcome	0	1	2	3	4	5
Obs Frequency	22	37	23	10	6	2

test at the 5% level the hypotheses
H_0 : The data is well modelled by $B(5, \frac{1}{4})$,
H_1 : The data is not well modelled by $B(5, \frac{1}{4})$.
So, under H_0 we have $B(5, \frac{1}{4})$. We calculate the probabilities of the six outcomes from S1:

x	0	1	2	3	4	5
$\mathbb{P}(X = x)$	$\frac{243}{1024}$	$\frac{405}{1024}$	$\frac{135}{512}$	$\frac{45}{512}$	$\frac{15}{1024}$	$\frac{1}{1024}$

Then we note that the total number in the observed data is 100, so we multiply the expected probabilities by 100 to obtain expected frequencies (to 1dp).

Outcome	0	1	2	3	4	5
Exp Frequency	23.7	39.6	26.3	8.8	1.5	0.1

We see that the expected frequencies have dropped below five, so we combine the last 3 columns to obtain:

OBS	22	37	23	18
EXP	23.7	39.6	26.3	10.4

So $X^2 = \frac{2.89}{23.7} + \frac{6.76}{39.6} + \frac{10.89}{26.3} + \frac{57.76}{10.4} = 6.26$.
Now the only constraint here is the total observed frequencies of 100, so $\nu = 4-1 = 3$. In the tables we observe $\mathbb{P}(\chi^2_3 \leqslant 7.815) = 0.95$. Therefore the critical X^2 value is 7.815. So $6.26 < 7.815$ and we therefore have no reason to reject H_0 and conclude that $B(5, \frac{1}{4})$ is probably a good model for the data.
  - Parameter Estimation. It is important to note that there is a difference in ν in the following situations:
    - H_0 : The data can be modelled by a Poisson distribution with $\lambda = 3.1$.
    - H_0 : The data can be modelled by a Poisson distribution.

The second has an extra constraint because you will need to estimate the value of λ from your observed data. In general just remember that if you estimate a parameter from observed data then this provides another constraint.

* If you need to estimate p from a frequency table for testing the goodness of fit of a binomial distribution you calculate $\bar{x}$ from the data in the usual way and equate this with np because that is the expectation of a binomial. For example, estimate p from the following observed data:

x	0	1	2	3	4
Obs frequency	12	16	6	2	1

So $np = \bar{x} = \frac{0\times12+1\times16+2\times6+3\times2+4\times1}{37} = \frac{38}{37}$. Therefore $p = \frac{38}{37\times4} = 0.257$ (to 3dp).

* If you need to estimate λ from a frequency table for testing the goodness of fit of a Poisson distribution you calculate $\bar{x}$ from the data in the usual way and equate this with λ. The only potential difficulty lies in the fact that the Poisson distribution has an infinite number of outcomes $\{0, 1, 2, 3 \dots\}$. However, the examiners will take pity and give you a scenario such as

x	0	1	2	3	4 or more
Obs frequency	5	11	10	3	0

where the "4 or more" frequency will be zero. Therefore $\lambda = \frac{0\times5+1\times11+2\times10+3\times3}{29} =$ 1.38 (to 2dp).

* Likewise the geometric distribution takes an infinite number of possible outcomes $\{1, 2, 3, 4 \dots\}$, and $\mathbb{E}(X) = \frac{1}{p}$, so to estimate p we calculate $\frac{1}{\mathbb{E}(X)}$. For example given

x	1	2	3	4	5 or more
Obs frequency	26	20	13	6	0

So, $\bar{x} = \frac{1\times26+2\times20+3\times13+4\times6}{65} = \frac{129}{65}$. Therefore $p = \frac{65}{129}$.

– For example for the following, test at the 1% level the following hypotheses:

H_0 : The data is well modelled by a Poisson,

H_1 : The data is not well modelled by a Poisson.

x	0	1	2	3	4	5or more
Obs frequency	14	23	14	7	2	0

So we estimate from the data (as above) $\lambda = \frac{4}{3}$. Now we calculate the first five expected values using total $\times\, \mathbb{P}(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}$. The final total we calculate by 60 subtract the other five totals.

x	0	1	2	3	4	5or more
Exp frequency	15.8	21.1	14.1	6.2	2.1	0.7

So combining columns so that the expected values equal at least five we obtain.

OBS	14	23	14	9
EXP	15.8	21.1	14.1	9.0

Now $X^2 = 0.377$. $\nu = 4 - 2 = 2$ (2 constraints because of 60 total and estimation of λ).

From tables $\mathbb{P}(\chi_2^2 < 9.210) = 0.99$. $0.377 < 9.210$ and therefore at the 1% level we have no reason to reject H_0 and conclude that the data is probably well described by a Poisson.

• Contingency Tables

- we are looking for *independence* (or, equivalently, dependence) between two variables. Remember that two events (A and B) are independent if $\mathbb{P}(A|B) = \mathbb{P}(A|B') = \mathbb{P}(A)$. Coupling this with the formula $\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$ (which drops out easily from a Venn diagram with A and B overlapping) we discover that independence *implies* $\mathbb{P}(A) \times \mathbb{P}(B) = \mathbb{P}(A \cap B)$. Therefore given any contingency table showing observed values we wish to calculate the values that would be expected *if they were* independent. Then carry out the analysis as before.
- For example 81 children are asked which of football, rugby or netball is their favourite.

OBS	Football	Rugby	Netball	TOTAL
Boy	17	25	3	45
Girl	9	3	24	36
Total	26	28	27	81

  Now, *if* the sex and choice of favourite were independent then $\mathbb{P}(\text{rugby and girl}) = \mathbb{P}(\text{rugby}) \times \mathbb{P}(\text{girl}) = \frac{28}{81} \times \frac{36}{81}$. Therefore the number of girls who like rugby best should be $81 \times \frac{28}{81} \times \frac{36}{81}$. The 81 cancels to give an expected number of $\frac{28\times36}{81}$. This is an example of the general result

$$\text{expected number} = \frac{\text{column total} \times \text{row total}}{\text{grand total}}.$$

  Therefore in our example we have

EXP	Football	Rugby	Netball	TOTAL
Boy	$\frac{26\times45}{81} = 14\frac{4}{9}$	$\frac{28\times45}{81} = 15\frac{5}{9}$	$\frac{27\times45}{81} = 15$	45
Girl	$\frac{26\times36}{81} = 11\frac{5}{9}$	$\frac{28\times36}{81} = 12\frac{4}{9}$	$\frac{27\times36}{81} = 12$	36
Total	26	28	27	81

  None of the expected values are less than 5, so no need to combine columns. Therefore $X^2 = \sum \frac{(O-E)^2}{E} = 35.52$ (to 2 dp). Make sure you can get my answer. A table often helps you build up to the answer. Use columns O, E, $(O-E)^2$, $\frac{(O-E)^2}{E}$.
- In an $m \times n$ contingency table the degrees of freedom is

$$\nu = (m-1)(n-1).$$

  So in the above example $\nu = (3-1) \times (2-1) = 2$. So if we were to carry out a hypothesis test (at the 5% level) of

  H_0 : The variables 'sex' and 'favourite sport' are independent;

  H_1 : The variables 'sex' and 'favourite sport' are not independent.

  We would use the correct row in the χ^2 tables to discover that $\mathbb{P}(\chi^2_2 > 5.991) = 0.05$. Now $35.52 > 5.991$ so we reject H_0 and conclude that 'sex' and 'favourite sport' are not independent.
- If you have a 2×2 contingency table you must apply Yates's correction. Here you reduce each value of $|O-E|$ by $\frac{1}{2}$. Again a table helps you build up to the answer. Use columns O, E, $|O-E|$, $\left(|O-E| - \frac{1}{2}\right)^2$, $\frac{(|O-E|-\frac{1}{2})^2}{E}$.

  For example carry out a hypothesis test to see if hair colour and attractiveness are independent.

OBS	Blonde	Not blonde	TOTAL
Fit	24	16	40
Minging	14	46	60
Total	38	62	100

Expected values are calculated as before.

EXP	Blonde	Not blonde	TOTAL
Fit	$\frac{38\times40}{100}=15.2$	$\frac{62\times40}{100}=24.8$	40
Minging	$\frac{38\times60}{100}=22.8$	$\frac{62\times60}{100}=37.2$	60
Total	38	62	100

Therefore the table would be

O	E	$\lvert O-E\rvert$	$\left(\lvert O-E\rvert-\frac{1}{2}\right)^2$	$\frac{(\lvert O-E\rvert-\frac{1}{2})^2}{E}$
24	15.2	8.8	68.89	4.532
16	24.8	8.8	68.89	2.778
14	22.8	8.8	68.89	3.021
46	37.2	8.8	68.89	1.852
				12.183

$X^2 = 12.183$ and $\nu = 1$ and you use these values in any subsequent hypothesis test. (Note that X^2 is pretty high here and for any significance level in the tables we would reject the hypothesis that hair colour and fitness were independent. Blondes are hot.)

OCR Statistics 4 Module Revision Sheet

The S4 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J.M.S.

Preliminaries

- Your pure maths needs to be far stronger for S4 than in any other Statistics module.
- You must be strong on general binomial expansion from C4.

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \frac{n(n-1)(n-2)(n-3)}{2}x^4 + \cdots$$

  This is valid only for $|x| < 1$. This is important for probability/moment generating functions.

- In particular you must be good at 'plucking out' specific coefficients (which may represent probabilities). For example find the x^8 coefficient in $\frac{x(3+x^2)}{\sqrt{4+2x}}$.

$$\begin{aligned}\frac{x(3+x^2)}{\sqrt{4+2x}} &= (3x+x^3)(4+2x)^{-\frac{1}{2}} \\ &= (3x+x^3)\left(4(1+\frac{x}{2})\right)^{-\frac{1}{2}} \\ &= \frac{1}{2}(3x+x^3)\left(1+\frac{x}{2}\right)^{-\frac{1}{2}} \\ &= \frac{1}{2}(3x+x^3)\left(1-\frac{x}{4}+\cdots-\frac{63}{8192}x^5+\cdots-\frac{429}{262,144}x^7+\dots\right)\end{aligned}$$

  So the x^8 coefficient will be $-\frac{1}{2}(3\times\frac{429}{262,144}+1\times\frac{63}{8192}) = -\frac{3303}{524,288}$. It helps hugely to be thinking ahead about what coefficients you are going to need.

- Recall from S3 that $\mathbb{E}(g(X)) = \sum g(x_i)p_i$ for discrete random variables and $\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)\,dx$ for continuous random variables.
- Recall also that $\text{Var}(X) \equiv \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.

Probability

- There are three very useful ways of representing information in probability questions. Venn diagrams, tree diagrams and two-way tables. You must think hard about which approach is going to be most helpful in the question you are to answer. Read the whole question before you start!
- Set theory is very important in probability. Know the following

- '$A \cap B$' is the intersection of the sets A and B. The overlap between the two sets. "AND"
- '$A \cup B$' is the union of the sets A and B. Anything that lies in either A or B (or both). "OR"
- A' means 'not A'. Everything outside A.
- $\{\,\}$ (or $\emptyset$) denotes the empty set. For example $A \cap A' = \{\,\}$

- Events A and B are *mutually exclusive* if both A and B cannot both happen. Represented by a Venn diagram of non-overlapping circles. Here

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

- However in the general case where A and B are not mutually exclusive we have

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

  This is because we are overcounting the overlap. It is called the *addition law*.

  For three events the addition law becomes A, B and C we have (in general)

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C).$$

  Again this drops out easily from a Venn diagram.

- Events $A_1, A_2, \ldots$ are said to be exhaustive if $\mathbb{P}(A_1 \cup A_2 \cup \ldots) = 1$. In other words the events $A_1, A_2, \ldots$ contain all the possibilities.

- If A and B are *independent* events then

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B).$$

- We read $\mathbb{P}(A|B)$ as the probability of A given that B has occured. It is defined

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

  However this formula is not always easy to apply, so Mr Stone's patented 'collapsing universes' approach from a Venn or tree diagram is often more intuitive.

- Using $\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$ and $\mathbb{P}(B|A) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(A)}$ we discover

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A) = \mathbb{P}(B)\mathbb{P}(A|B).$$

  This is called the *multiplication law* of probability and is incredibly useful in converting $\mathbb{P}(A|B)$ into $\mathbb{P}(B|A)$ and vice versa. The multiplication law drops out readily from a tree diagram.

- Bayes' Theorem[51] states

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(A_j)\mathbb{P}(B|A_j)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_j)\mathbb{P}(B|A_j)}{\sum_{\text{all } i} \mathbb{P}(A_i)\mathbb{P}(B|A_i)}.$$

  This looks scary, but drops out from a tree diagram. The formal statement is not required for S4, but is very important.

[51] Reverend Thomas Bayes from my home town of Tunbridge Wells. Wrote a document defending Newton's calculus hence a rather good bloke.

Non-Parametric Tests

- All of the hypothesis tests studied in Stats 2 & 3 required knowledge (or at the very least an assumption) of some kind of underlying distribution for you to carry out the test. However sometimes you have no knowledge about the underlying population. Statisticians therefore developed a series of *non-parametric* tests for situations where you have no knowledge of the underlying population.

- The *sign test* is a test about the *median* (i.e. the point at which you have an equal number of data points either side). If H_0 : median = 10, say, then under H_0, whether a random piece of data lies above or below 10 has probability $\frac{1}{2}$. For n pieces of data we therefore have a binomial $B(n, \frac{1}{2})$. Rather than work out critical values, the best approach is probably to calculate (under H_0) the probability of what you have observed and anything more extreme. For example test at the 5% level whether the median of the data

  $$1,\ 1,\ 2,\ 3,\ 6,\ 7,\ 8,\ 9,\ 9,\ 9,\ 10,\ 10,\ 11,\ 13$$

  is 5. Note that there are four pieces of data less than 5.

  H_0 : The median of the data is 5.

  H_1 : The median of the data is not 5.

  $\alpha = 5\%$. Two tailed test.

  Under H_0, $X \sim B(14, \frac{1}{2})$.

  $\mathbb{P}(X \leqslant 4) = 0.0898 > 0.025$, so at the 5% level there is insufficient evidence to reject H_0 and we conclude that the median of the data is probably 5. [You could have also gone through the rigmarole of demonstrating that the critical value is 2 (or 12) but my way is quicker and life's short.]

- Although there is no example in your textbook I see no reason why they couldn't ask a question where you had a large enough sample to require the normal approximation to $B(n, \frac{1}{2})$... don't forget your continuity correction.

- The sign test is a very crude test because it takes absolutely no account of how far away the data lies on either side of the median. If you want to take account of the magnitude of the deviations you need to use...

- ... the *Wilcoxon signed-rank test*. Here it is assumed that the data is *symmetric*; therefore it is a test about both the median or the mean because for symmetric data the median and mean are the same.

  You calculate the deviations from the median/mean, rank the size of the deviations and then sum the positive ranks to get P and sum the negative ranks to get Q. The test statistic is T, where T is the smaller of P or Q. For example test at the 5% level whether the mean of

  $$1.3,\ 2.1,\ 7.3,\ 4.9,\ 3.2,\ 1.6,\ 5.6,\ 5.7$$

  is 3.

  The data sort of looks symmetric, so OK to proceed with Wilcoxon.

  H_0 : The mean of the data is 3.

  H_1 : The mean of the data is greater than 3.

  $\alpha = 5\%$. One tailed test.

Data	1.3	2.1	7.3	4.9	3.2	1.6	5.6	5.7
Deviation	-1.7	-0.9	$+4.3$	$+1.9$	$+0.2$	-1.4	$+2.6$	$+2.7$
Rank	4	2	8	5	1	3	6	7
Signed Rank	-4	-2	$+8$	$+5$	$+1$	-3	$+6$	$+7$

So $P = 27$, $Q = 9$, so $T_{\text{obs}} = 9$. The lower T is, the worse it is for H_0 and the tables give the *largest* value at which you would reject H_0. $T_{\text{crit}} = 5$. $9 > 5$, so at the 5% level we have insufficient evidence to reject H_0 and conclude that the mean is probably 3.

- For large samples (i.e. when the tables don't give the values you want; running out of values) a normal approximation can be used where

$$Z = \frac{T + 0.5 - \frac{1}{4}n(n+1)}{\sqrt{\frac{1}{24}n(n+1)(2n+1)}}.$$

Note that because T is the smaller of P and Q that Z will always be negative (both Z_{crit} and Z_{obs}). For example if you had 100 pieces of data and you were testing at the 1% level whether the mean was some value (against H_1 of the mean not being some value) and $P = 2000$ and $Q = 3050$ then $T = 2000$. So

$$\begin{aligned} Z_{\text{obs}} &= \frac{T_{\text{obs}} + 0.5 - \frac{1}{4}n(n+1)}{\sqrt{\frac{1}{24}n(n+1)(2n+1)}} \\ &= \frac{2000 + 0.5 - \frac{1}{4} \times 100 \times 101}{\sqrt{\frac{1}{24} \times 100 \times 101 \times 201}} \\ &= -1.803 \end{aligned}$$

Because it is a two-tailed 1% test we reverse look-up 0.995 to obtain $Z_{\text{crit}} = -2.576$. Finally $-1.803 > -2.576$, so at the 1% level there is insufficient evidence to reject H_0 and conclude that the mean is probably whatever we thought it was under H_0.

- The *Wilcoxon rank-sum test* is the non-parametric equivalent of the two-sample t-test from S3. It tests whether two different sets of data are drawn from identical populations. The central idea for the theory is that if X and Y are drawn from identical distributions, then $P(X < Y) = \frac{1}{2}$. The tables are then constructed from tedious consideration of all the possible arrangements of the ranks (called the 'sampling distribution').

Given two sets of data, let m be the number of pieces of data from the smaller data set and n be the number of pieces of data from the larger data set (if they are both the same size it's up to you which is m and which n). Then rank *all* the data and sum the ranks of the 'm' population; call this total R_m. Also calculate $m(n+m+1) - R_m$ and let the test statistic W be the smaller of R_m and $m(n+m+1) - R_m$. The smaller W is, the more likely we are to reject H_0 and the tables give the largest W at which we reject H_0.

For example test at the 5% level whether the following are drawn from identical populations.

A	23	14	42	12	30	40
B	16	21	9	35		

H_0 : Data drawn from identical distributions.

H_1 : Data not drawn from identical distributions.

$\alpha = 5\%$. Two tailed test.

Data	$\boxed{9}$	12	14	$\boxed{16}$	$\boxed{21}$	23	30	$\boxed{35}$	40	42
Rank	$\boxed{1}$	2	3	$\boxed{4}$	$\boxed{5}$	6	7	$\boxed{8}$	9	10

So $m = 4$, $n = 6$, $R_m = 18$, $m(n + m + 1) - R_m = 26$, $W_{\text{obs}} = 18$. Looking at the tables we see $W_{\text{crit}} = 12$, and $18 > 12$, so at the 5% level there is insufficient evidence to reject H_0 and we conclude that the data is probably drawn from identical distributions.

- For large samples (i.e. when the tables don't give the values you want; running out of values) a normal approximation can be used where

$$Z = \frac{W + 0.5 - \frac{1}{2}m(m + n + 1)}{\sqrt{\frac{1}{12}mn(m + n + 1)}}.$$

Probability Generating Functions

- In Stats 1 & 2 you met discrete random variables (DRVs) such that each outcome had a probability attached. Sometimes there were rules which related the probability to the outcome (binomial, geometric, Poisson). However, in general we had:

x	x_1	x_2	x_3	x_4	$\dots$
$\mathbb{P}(X = x)$	p_1	p_2	p_3	p_4	$\dots$

Recall that $\sum p_i = 1$ because the sum of all the probabilities must total 1 and that $\mathbb{E}(X) = \sum p_i x_i$. Also $\mathbb{E}(f(X)) = \sum p_i f(x_i)$ from Stats 3 and $\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \sum p_i x_i^2 - (\sum p_i x_i)^2$ from Stats 2.

- At some point some bright spark decided to consider the properties of

$$G_X(t) = \mathbb{E}(t^X) = \sum p_i t^{x_i} = p_1 t^{x_1} + p_2 t^{x_2} + p_3 t^{x_3} + p_4 t^{x_4} + \cdots$$

where t is a 'dummy variable' unrelated to x. You can see that this will create either a finite or infinite series. This is called the probability generating function of X. It is a single function that contains within it all of the (potentially infinite) probabilities of X.

For example given

x	-2	-1	0	1	2
$\mathbb{P}(X = x)$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{8}$	$\frac{1}{8}$

the generating function is $G_X(t) = p_1 t^{x_1} + p_2 t^{x_2} + p_3 t^{x_3} + p_4 t^{x_4} + \cdots = \frac{1}{6}t^{-2} + \frac{1}{4}t^{-1} + \frac{1}{3} + \frac{1}{8}t + \frac{1}{8}t^2$. We can therefore see that if (say) we saw a term $\frac{5}{24}t^6$, then we can see that $\mathbb{P}(X = 6) = \frac{5}{24}$. Note that if you see a constant term then that tells you $\mathbb{P}(X = 0)$ because $t^0 = 1$.

- An important property is that $G_X(1) = 1$ because $G_X(1)$ is just the sum of all the probabilities of X, i.e. $\sum p_i$.

- Another useful thing to do is consider the derivative $G'_X(t)$ with respect to t;

$$G'_X(t) = \sum p_i x_i t^{x_i - 1} = p_1 x_1 t^{x_1 - 1} + p_2 x_2 t^{x_2 - 1} + p_3 x_3 t^{x_3 - 1} + \cdots$$

Again, if we consider $G'(1)$ we obtain

$$G'(1) = \sum p_i x_i = p_1 x_1 + p_2 x_2 + p_2 x_2 + p_3 x_3 + \cdots = \mathbb{E}(X).$$

- Variances can also be calculated by

$$\text{Var}(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2.$$

- Some standard pgfs are given in the formula book:

Distribution	$B(n,p)$	$Po(\lambda)$	$Geo(p)$
pgf	$(1-p+pt)^n$	$e^{\lambda(t-1)}$	$\frac{pt}{1-(1-p)t}$

  Any good candidate should be able to derive these...

- For two *independent* random variables X and Y (with pgfs $G_X(t)$ and $G_Y(t)$ respectively) the pgf of $X+Y$ is $G_{X+Y}(t) = G_X(t) \times G_Y(t)$. This extends to three or more *independent* random variables.

Moment Generating Functions

- You will recall from FP2 that the Maclaurin expansion for e^x is

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots.$$

  This is valid for all values of x (and you should know why from your Pure teachings). An alternative notation used is $e^x \equiv \exp(x)$.

- The nth moment of a distribution is $\mathbb{E}(X^n)$. So the first moment is just $\mathbb{E}(X)$. The second moment is $\mathbb{E}(X^2)$, which is useful in calculating variances. The zeroth moment is $\mathbb{E}(X^0) = \mathbb{E}(1) = 1$.

- The moment generating function (mgf) is defined for

x	x_1	x_2	x_3	x_4	$\cdots$
$\mathbb{P}(X=x)$	p_1	p_2	p_3	p_4	$\cdots$

  by

$$\begin{aligned}
M_X(t) = \mathbb{E}(e^{tX}) = \sum p_i e^{x_i t} &= p_1 e^{x_1 t} + p_2 e^{x_2 t} + p_3 e^{x_3 t} + p_4 e^{x_4 t} + \cdots \\
&= \quad p_1 + p_1 x_1 t + p_1 \frac{x_1^2 t^2}{2!} + p_1 \frac{x_1^2 t^3}{3!} + \cdots \\
&\quad + p_2 + p_2 x_2 t + p_2 \frac{x_2^2 t^2}{2!} + p_2 \frac{x_2^2 t^3}{3!} + \cdots \\
&\quad + p_3 + p_3 x_3 t + p_3 \frac{x_3^2 t^2}{2!} + p_3 \frac{x_3^2 t^3}{3!} + \cdots \\
&\quad + p_4 + p_4 x_2 t + p_4 \frac{x_4^2 t^2}{2!} + p_4 \frac{x_4^2 t^3}{3!} + \cdots \\
&= \quad (p_1 + p_2 + p_3 + p_4 + \cdots) \\
&\quad + (p_1 x_1 + p_2 x_2 + p_3 x_3 + p_4 x_4 + \cdots)\ t \\
&\quad + \left(p_1 x_1^2 + p_2 x_2^2 + p_3 x_3^2 + p_4 x_4^2 + \cdots\right) \frac{t^2}{2!} \\
&\quad + \left(p_1 x_1^3 + p_2 x_2^3 + p_3 x_3^3 + p_4 x_4^3 + \cdots\right) \frac{t^3}{3!} \\
&\quad + \cdots \\
&= \mathbb{E}(1) + \mathbb{E}(X)t + \mathbb{E}\left(X^2\right) \frac{t^2}{2!} + \mathbb{E}\left(X^3\right) \frac{t^3}{3!} + E\left(X^4\right) \frac{t^4}{4!} + \cdots
\end{aligned}$$

So you can see that the constant term of $M_X(t)$ should always be $\mathbb{E}(1) = 1$ because it represents the sum of the probabilities. The coefficient of t will be $\mathbb{E}(X)$ and the coefficient of $\frac{t^2}{2!}$ (*not* just the coefficient of t^2) will be $\mathbb{E}(X^2)$. In general the coefficient of $\frac{t^n}{n!}$ will be $\mathbb{E}(X^n)$, that is, the nth moment.

- As with pgfs, differentiating mgfs (with respect to t) is a 'good thing'. However, instead of letting $t = 1$ we let $t = 0$ (because $a^0 = 1$). So differentiating $M_X(t)$ we find:

$$\begin{aligned} M_X(t) &= p_1e^{x_1t} + p_2e^{x_2t} + p_3e^{x_3t} + p_4e^{x_4t} + \cdots \\ M'_X(t) &= p_1x_1e^{x_1t} + p_2x_2e^{x_2t} + p_3x_3e^{x_3t} + p_4x_4e^{x_4t} + \cdots \\ M'_X(0) &= p_1x_1 + p_2x_2 + p_3x_3 + p_4x_4 + \cdots \\ &= \sum x_ip_i = \mathbb{E}(X). \end{aligned}$$

So $M'_X(0) = \mathbb{E}(X)$.

Differentiating again we find:

$$\begin{aligned} M'_X(t) &= p_1x_1e^{x_1t} + p_2x_2e^{x_2t} + p_3x_3e^{x_3t} + p_4x_4e^{x_4t} + \cdots \\ M''_X(t) &= p_1x_1^2e^{x_1t} + p_2x_2^2e^{x_2t} + p_3x_3^2e^{x_3t} + p_4x_4^2e^{x_4t} + \cdots \\ M''_X(0) &= p_1x_1^2 + p_2x_2^2 + p_3x_3^2 + p_4x_4^2 + \cdots \\ &= \sum x_i^2p_i = \mathbb{E}(X^2). \end{aligned}$$

So using $\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ we find $\mathrm{Var}(X) = M''_X(0) - (M'_X(0))^2$.

- Notice that with mgfs there are two ways to obtain the expectation and variance of your random variable. All things being equal I would choose the differentiation method, but you must ensure that your mgf is defined for $t = 0$. Also read the question carefully to see what they are wanting.

- Moment generating functions can also be defined for continuous random variables:

$$M_X(t) = \int_{-\infty}^{\infty} f(x)e^{tx}\,dx.$$

As before $M_X(0) = 1$, $M'_X(0) = \mathbb{E}(X)$, $M''_X(0) = \mathbb{E}(X^2)$. Convergence issues can arise EXAMPLE!!!!!!

- Some standard mgfs are given in the formula book:

Distribution	Uniform on $[a, b]$	Exponential	$N(\mu, \sigma^2)$
mgf	$\frac{e^{bt}-e^{at}}{(b-a)t}$	$\frac{\lambda}{\lambda-t}$	$e^{\mu t+\frac{1}{2}\sigma^2t^2}$

Any good candidate should be able to derive these too...

- As with pgfs, for two *independent* random variables X and Y (with mgfs $G_X(t)$ and $G_Y(t)$ respectively) the mgf of $X + Y$ is $M_{X+Y}(t) = M_X(t) \times M_Y(t)$. This extends to three or more *independent* random variables.

Estimators

- It is vital to recall here that $\mathbb{E}(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$ (by definition).

- Given a population there may be many parameters that we may wish to know. For example we might like to know the mean μ, the variance σ^2, the median M, the maximum or minimum, the IQR, etc. In general we shall call this parameter θ.

  Usually we will never know θ because we won't have the whole population. But we will be able to take a random sample from the population. From this sample we can calculate a quantity U which we shall use to estimate θ. We call U an estimator of θ.

- U is said to be an *unbiased estimator* of θ if

$$\mathbb{E}(U) = \theta.$$

  i.e. if we take an average of *all possible* U (remember that U is a random variable) we will get the desired θ. If $\mathbb{E}(U) \neq \theta$ then the estimator is said to be biased (not giving the desired result on average).

  For example to show that $K = \frac{X_1+2X_2+5X_3}{8}$ is an unbiased estimator of μ we merely consider $\mathbb{E}(K)$ and keep whittling down as far as we can go (using S3 expectation and variance algebra)

$$\begin{aligned}\mathbb{E}(K) &= \mathbb{E}\left(\frac{X_1+2X_2+5X_3}{8}\right)\\ &= \mathbb{E}\left(\frac{X_1}{8}+\frac{X_2}{4}+\frac{5X_3}{8}\right)\\ &= \frac{1}{8}\mathbb{E}(X_1)+\frac{1}{4}\mathbb{E}(X_2)+\frac{5}{8}\mathbb{E}(X_3)\\ &= \frac{1}{8}\mu+\frac{1}{4}\mu+\frac{5}{8}\mu = \mu.\end{aligned}$$

- For continuous random variables just remember that $\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)\,dx$. For example find the value of k which makes $L = k(X_1 + X_2)$ an unbiased estimator of θ for

$$f(x) = \begin{cases} \frac{2}{\theta}\left(1-\frac{x}{\theta}\right) & 0 \leqslant x \leqslant \theta \\ 0 & \text{otherwise} \end{cases}$$

  First calculate $\mathbb{E}(X)$ *à la* S2:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_0^{\theta} x\frac{2}{\theta}\left(1-\frac{x}{\theta}\right)\,dx = \left[\frac{x^2}{\theta}-\frac{2x^3}{3\theta^2}\right]_0^{\theta} = \frac{\theta}{3}.$$

  So for L to be unbiased we need $\mathbb{E}(L) = \theta$, so

$$\begin{aligned}\mathbb{E}(L) &= \theta\\ \mathbb{E}(k(X_1+X_2) &= \theta\\ k(\mathbb{E}(X_1)+\mathbb{E}(X_2)) &= \theta\\ k(2\times\mathbb{E}(X)) &= \theta\\ k\left(\frac{2\theta}{3}\right) &= \theta\\ k &= \frac{3}{2}.\end{aligned}$$

- Given two *unbiased* estimators the *most efficient* estimator (of the two) is the one where $\mathrm{Var}(U)$ is smaller. A smaller variance is a 'good thing'.

- Sometimes you may need calculus to work out the most efficient estimator from an infinite family. For example X_1, X_2 and X_3 are three independent measurements of X.

$$S = \frac{aX_1 + 2X_2 + 4X_3}{a+6} \quad \text{(with } a \neq -6\text{)}$$

is suggested as an estimator for μ. Prove that S is unbiased whatever the value of a and find the value of a which makes S most efficient. So

$$\begin{aligned}
\mathbb{E}(S) &= \mathbb{E}\left(\frac{aX_1 + 2X_2 + 4X_3}{a+6}\right) \\
&= \frac{1}{a+6}\mathbb{E}(aX_1 + 2X_2 + 4X_3) \\
&= \frac{1}{a+6}[a\mathbb{E}(X_1) + 2\mathbb{E}(X_2) + 4\mathbb{E}(X_3)] \\
&= \frac{1}{a+6}[a\mu + 2\mu + 4\mu] \\
&= \frac{\mu}{a+6}(a+6) = \mu.
\end{aligned}$$

So S is unbiased for all values of a. Now consider

$$\begin{aligned}
\text{Var}(S) &= \text{Var}\left(\frac{aX_1 + 2X_2 + 4X_3}{a+6}\right) \\
&= \frac{1}{(a+6)^2}\text{Var}(aX_1 + 2X_2 + 4X_3) \\
&= \frac{1}{(a+6)^2}[a^2\text{Var}(X_1) + 4\text{Var}(X_2) + 16\text{Var}(X_3)] \\
&= \frac{a^2+20}{(a+6)^2}\sigma^2.
\end{aligned}$$

To minimise $\text{Var}(S)$ we need $\frac{d}{da}\text{Var}(S) = 0$. So

$$\begin{aligned}
0 &= \frac{d}{da}\left(\frac{a^2+20}{(a+6)^2}\sigma^2\right) = \frac{2a(a+6)^2 - 2(a+6)(a^2+20)}{(a+6)^4}\sigma^2 \\
\text{So } 0 &= 2a(a+6)^2 - 2(a+6)(a^2+20) \\
0 &= 2(a+6)[a(a+6) - (a^2+20)] \\
0 &= (a+6)(6a-20).
\end{aligned}$$

So $a = -6$ or $a = \frac{10}{3}$, but $a \neq -6$ so $a = \frac{10}{3}$ is the value of a that makes S most efficient[52].

- Here is a tough type of problem that caught me out the first two (or three (or four (...))) times I saw it. Slot away the method just in case. For example consider

$$f(x) = \begin{cases} \frac{2x}{\theta^2} & 0 \leqslant x \leqslant \theta \\ 0 & \text{otherwise} \end{cases}$$

An estimate of θ is required and a suggestion is made to calculate $\frac{5L}{4}$ where L is the maximum of two independent observations of X (X_1 and X_2). Show that this estimator is unbiased.

The thing to remember is that for L to be the maximum of X_1 and X_2, then X_1 *and* X_2 must *both* be less than or equal to L; i.e. we are going to calculate a cdf. So

$$\mathbb{P}(L \leqslant l) = \mathbb{P}(X_1 \leqslant l) \times \mathbb{P}(X_2 \leqslant l).$$

[52] I suppose we should consider the second derivative to show that this value of a minimises rather than maximises the variance, but life's too short...

(This can be extended to three or more independent samplings of X.)

By sketching $f(x)$ we can see that the probability that one observation is less than or equal to l is given by a triangle in this case of area $\frac{l^2}{\theta^2}$ (or by the integral $\int_0^l f(x)\,dx$ for a more general $f(x)$). So $\mathbb{P}(L \leqslant l) = \mathbb{P}(X_1 \leqslant l) \times \mathbb{P}(X_2 \leqslant l) = \frac{l^2}{\theta^2} \times \frac{l^2}{\theta^2} = \frac{l^4}{\theta^4}$. Differentiating wrt to l we find the pdf of l to be

$$f(l) = \begin{cases} \frac{4l^3}{\theta^4} & 0 \leqslant l \leqslant \theta \\ 0 & \text{otherwise} \end{cases}$$

Therefore we calculate $\mathbb{E}\left(\frac{5L}{4}\right)$ as follows:

$$\begin{aligned} \mathbb{E}\left(\frac{5L}{4}\right) &= \frac{5}{4}\mathbb{E}(L) \\ &= \frac{5}{4}\int_0^\theta l \times \frac{4l^3}{\theta^4}\,dl \\ &= \frac{5}{4}\left[\frac{4l^5}{5\theta^4}\right]_0^\theta = \theta. \end{aligned}$$

Therefore $\frac{5L}{4}$ is an unbiased estimator of θ. I will leave it as an exercise for the reader to demonstrate that $\text{Var}\left(\frac{5L}{4}\right) = \frac{\theta^2}{24}$.

Discrete Bivariate Distributions

- The discrete random variables you have met thus far have been in one variable only. For example

x	2	3	5	7
$\mathbb{P}(X = x)$	$\frac{1}{2}$	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{1}{18}$

However we can have discrete *bivariate* distributions. For example

		X		
		2	3	5
	4	0	$\frac{1}{2}$	$\frac{1}{10}$
Y	5	$\frac{1}{5}$	$\frac{3}{20}$	$\frac{1}{20}$

From this we can see, say, $\mathbb{P}(X = 3, Y = 5) = \frac{3}{20}$.

- The marginal distribution is what one obtains if one of the variables is 'ignored'. In the above example the marginal distribution of X can be written

x	2	3	5
$\mathbb{P}(X = x)$	$\frac{1}{5}$	$\frac{13}{20}$	$\frac{3}{20}$

This can be added to the bivariate distribution thus:

		X		
		2	3	5
	4	0	$\frac{1}{2}$	$\frac{1}{10}$
Y	5	$\frac{1}{5}$	$\frac{3}{20}$	$\frac{1}{20}$
		$\frac{1}{5}$	$\frac{13}{20}$	$\frac{3}{20}$

$\mathbb{E}(X)$ and $\text{Var}(X)$ can be calculated in the usual way obtaining $\mathbb{E}(X) = \frac{31}{10}$ and $\text{Var}(X) = \frac{79}{100}$ (do it!). Similarly you can work out the marginal distribution of Y if you are so inclined.

- The *conditional* distribution of a bivariate distribution can be calculated *given that* one of the variables (X or Y) has taken a specific value. For the above example the "distribution of X conditional on $Y = 4$" is calculated by rewriting the 4 row with all the values divided by $\mathbb{P}(Y = 4) = \frac{3}{5}$.

x	2	3	5
$\mathbb{P}(X = x \mid Y = 4)$	0	$\frac{5}{6}$	$\frac{1}{6}$

  This is all from our friend $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$.

- A way to check whether X and Y are *independent* of each other in a bivariate distribution is to check whether every entry in the distribution is the product of the two relevant marginal probabilities. For example

		X			
		1	2	3	
	1	$\frac{1}{3}$	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{2}{3}$
Y	2	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$	$\frac{1}{3}$
		$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$	

  Here we see $\mathbb{P}(X = 2, Y = 1) = \frac{2}{9}$ is the same as $\mathbb{P}(X = 2) \times \mathbb{P}(Y = 1) = \frac{1}{3} \times \frac{2}{3} = \frac{2}{9}$. The same is true for *every* entry in the table, so X and Y are independent. It only takes one entry not to satisfy this to ensure X and Y are *not* independent.

- The *covariance* of a discrete bivariate distribution is defined

$$\text{Cov}(X, Y) \equiv \mathbb{E}((X - \mu_X)(Y - \mu_Y)).$$

  However this tends to be cumbersome to calculate so we use the equivalent formula

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mu_X \mu_Y.$$

  The covariance can be thought of as the correlation coefficient (r from Stats 1) for two probability distributions (sort of). The covariance can be both positive or negative (like the correlation coefficient).

- To calculate the covariance, first create the marginal distributions:

		X		
		1	3	4
	2	$\frac{1}{3}$	$\frac{1}{4}$	0
Y	5	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{8}$

$\Rightarrow$

		X			
		1	3	4	
	2	$\frac{1}{3}$	$\frac{1}{4}$	0	$\frac{7}{12}$
Y	5	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{5}{12}$
		$\frac{1}{2}$	$\frac{3}{8}$	$\frac{1}{8}$	

  Then use the marginal distributions to calculate μ_X and μ_Y.
  $\mu_X = \mathbb{E}(X) = \sum xp = 1 \times \frac{1}{2} + 3 \times \frac{3}{8} + 4 \times \frac{1}{8} = \frac{17}{8}$.
  $\mu_Y = \mathbb{E}(Y) = \sum yp = 2 \times \frac{7}{12} + 5 \times \frac{5}{12} = \frac{13}{4}$.
  Now we use this to calculate the covariance thus:

$$\begin{aligned}\text{Cov}(X, Y) &= \mathbb{E}(XY) - \mu_X \mu_Y \\ &= (1 \times 2 \times \tfrac{1}{3}) + (1 \times 5 \times \tfrac{1}{6}) + (3 \times 2 \times \tfrac{1}{4}) + (3 \times 5 \times \tfrac{1}{8}) + (4 \times 2 \times 0) + (4 \times 5 \times \tfrac{1}{8}) - \tfrac{17}{8} \times \tfrac{13}{4} \\ &= \frac{15}{32}.\end{aligned}$$

- If X and Y are independent then $\text{Cov}(X, Y) = 0$. However, if $\text{Cov}(X, Y) = 0$ this does not *necessarily* mean that X and Y are independent. But if $\text{Cov}(X, Y) \neq 0$ then X and Y cannot be independent.

- With an understanding of covariance we can write the relationship for $\text{Var}(aX \pm bY)$ when X and Y are *not* independent:

$$\text{Var}(aX \pm bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) \pm 2ab\text{Cov}(X, Y).$$

  Notice the extra term at the end of the formula we are used to from S3 for *independent* X and Y.

OCR Decision 1 Module Revision Sheet

The D1 exam is 1 hour 30 minutes long. You are allowed a graphics calculator.

Before you go into the exam make sure you are fully aware of the contents of the formula booklet you receive. Also be sure not to panic; it is not uncommon to get stuck on a question (I've been there!). Just continue with what you can do and return at the end to the question(s) you have found hard. If you have time check all your work, especially the first question you attempted... always an area prone to error.

J.M.S.

Algorithms

- An **algorithm** is defined to be a *finite* sequence of instructions for solving a problem.
- Flow diagrams
- The **Bubble sort** algorithm orders numerical data in descending order.
  1. If there is only one number in the list then stop.
  2. Make one pass down the list, comparing numbers in pairs and swapping as necessary.
  3. If no swaps have occurred then stop. Otherwise, ignore the last element of the list and return to 1.

  Note that the first pass will guarantee that the largest number is at the bottom, which is why we can then ignore it.
- The **Shuttle sort** algorithm also orders numerical data in descending order. It is rather better than bubble sort on average.
  1. Compare the 1st and 2nd numbers in the list and swap if necessary.
  2. Compare the 2nd and 3rd numbers in the list and swap if necessary. If a swap has occurred, compare the 1st and 2nd numbers and swap if necessary.
  3. Compare the 3rd and 4th numbers in the list and swap if necessary. If a swap has occurred, compare the 2nd and 3rd numbers, and so on up the list.

  And so on through the entire list.
- First fit
  1. Place each object in turn in the first available space in which it will fit.

  Nice and simple...
- First fit decreasing

Graphs

- Rather a lot of highly tedious definitions to know I'm afraid. A **graph** is made up of **nodes/vertices** and connected by **arcs/edges**.
- (There can exist **loops** where an arc comes back to the same node it started at and **multiple arcs** where more than one arc connects two nodes, but usually you won't have to worry about these.) A **simple** graph is one without loops and multiple arcs.
- A **subgraph** is any 'subset' of a given graph. I don't know if a graph is its own subgraph. I also don't know if no graph is a subgraph of a given graph.
- A **complete graph** is one where every node is connected to every other node by exactly one arc. The notation K_n is used for the complete graph of n nodes. Every simple graph with $\leqslant n$ nodes is a subgraph of K_n. The number of arcs in K_n is $\frac{n(n-1)}{2}$; make sure you know why.
- A **bipartite graph** is one where a set of r nodes are connected to s nodes such that no nodes in r are connected to one another, similarly for s. A **complete bipartite graph** is one where every node in r is connected to every node in s and is denoted $K_{r,s}$. The number of nodes in $K_{r,s}$ is $r + s$ and the number of arcs is rs.
- A **trail/route** is a 'journey' taken from node-to-node in a graph. Repetitions of nodes are allowed.
- A **path** is a trail where no node is passed more than once.
- A **closed trail** is one where the start and end nodes are the same. Repetition of nodes allowed.
- A **cycle** is a closed trail where repetitions of nodes are not allowed (except the start and end nodes).
- The **order** of a node is the number of arcs that lead away from the node. They can be **odd** and **even** order.
- A **connected graph** is one where a path (not necessarily direct) can be found from any two nodes.
- An **Eulerian graph** is a connected graph with a closed trail containing every *arc* exactly once. This occurs if and only if *every node is of even order.*

  Informally this means you can draw the shape without taking your pen off the page, not going over previously drawn lines (arcs), with your pen ending up in the same place (node) you started.
- A **semi-Eulerian graph** is a connected graph which has a trail (crucially not closed) that contains every arc exactly once. This occurs if and only if *exactly two nodes have odd order.*

  Informally this means you can draw the shape without taking your pen off the page, not going over previously drawn lines (arcs), with your pen ending up in a different place (node) to where you started.
- A **planar graph** is one which can be drawn in a plane (2D Surface) such that arcs only meet at nodes; i.e. the arcs don't cross one another. It should be noted that some graphs can appear non-planar, but can be re-drawn to be planar.

Euler's relationship relates the number of of regions R, nodes N, and arcs A for any *connected, planar* graph. It states

$$R + N = A + 2.$$

Note that the region outside the graph is counted as a region.

- A **tree** is a connected graph with no cycles. Any *connected* graph contains at least one subgraph which is a tree. This is very important later on for Prim & Kruskal.

Networks

- Arcs in a graph can sometimes be given a value or **weight** (i.e. a 'cost' or 'time' or value to move along the arc). A graph with weights is called a **network**.
- A **digraph** is a directed graph where one can only move in specified directions around the network.
- **Matrices** can be useful in representing the weights to get between different nodes. For a non-directed graph there should always be reflectional symmetry around the leading diagonal of the matrix. For a digraph you must be careful to show the 'From' and 'To'.

Prim's Algorithm

- Prim's algorithm can be applied to a network to discover the minimum spanning tree T. For a graph follow:
  1. Select any node to be the first node of T.
  2. Consider the arcs which connect nodes in T to nodes outside T. Pick the one with minimum weight. Add this arc and the extra node to T. (If there are two or more arcs of minimum weight, choose any one of them.)
  3. Repeat 2 until T contains every node of the graph.
- For a matrix Prim's algorithm works like this:
  1. Select any node to be the first node of T.
  2. Circle the new node of T in the top row, and cross out the row corresponding to this new node.
  3. Find the smallest weight left in the columns corresponding to the nodes of T, and circle this weight. Then choose the node whose row the weight is in to join T. (If there are several possibilities for the weight, choose any one of them.)
  4. Repeat 2 and 3 until T contains every node.

Kruskal's Algorithm

- Like Prim, Kruskal's algorithm discovers the minimum spanning tree. Formally it states:
  1. Choose the arc of least weight.
  2. Choose from those arcs remaining the arc of least weight which does *not* form a cycle with already chosen arcs. (If there are several such arcs, choose one arbitrarily.)
  3. Repeat 2 until $n - 1$ arcs have been chosen.

Informally you look at the network and pick out the lowest weighted arc and add it to your set. Ensuring that you avoid any cycles you keep picking the lowest weighted arcs until you have a spanning tree.

There is no matrix method that I know of for Kruskal. If presented with a matrix then draw out the network.

Dijkstra's Algorithm

- **Dijkstra's** algorithm finds the shortest path from a node to other nodes in a network. The moment you create a permanent label you know that the shortest route from the starting node to that node has been found. Formally it states:

  1. Label the start node with zero and box this label.
  2. Consider the node with the most recently boxed label. Suppose this node to be X and let D be its permanent label. Then, in turn, consider each node directly joined to X but not yet permanently boxed. For each such node, Y say, temporarily label it with the lesser of $D+$ (the weight of arc XY) and its existing label (if any).
  3. Choose the least of all temporary labels on the network. Make this label permanent by boxing it.
  4. Repeat 2 and 3 until the destination node has a permanent label.
  5. Go backwards through the network, retracing the path of shortest length from the destination node to the start node.

- Here you should make use of these boxes to keep track of your progress; an examiner will also use these to mark your exam.

8	4
~~6~~ ~~5~~	4

This means that this node was the eighth node to be made permanent and it's shortest distance from the starting node was 4. Whilst discovering that 4 was the best there were temporary distances of 6 and 5 that were improved upon.

Travelling Salesperson

- A **Hamiltonian cycle** is defined as a tour which contains every node precisely once. It is often useful to find the Hamiltonian cycle of least total weight. This is a hard problem of factorial order.

- The **nearest neighbour** algorithm finds a reasonably good Hamitonian cycle. It does *not* necessarily find the best.

  1. Choose any starting node.
  2. Consider the arcs which join the previous chosen node to not-yet-chosen nodes. From these arcs pick one that has minimum weight. Choose this arc, and the new node on the end of it, to join the cycle.
  3. Repeat 2 until all nodes have been chosen.
  4. Then add the arc that joins the last-chosen node to the first-chosen node.

Informally you start at a node and run to the next node along the arc of least weight avoiding nodes you've been to before and keep doing that until you run out of nodes. Then go back to the start.

- A lower bound can be found with the **Lower Bound** algorithm.

  1. Choose an arbitrary node, say X. Find the total of the two smallest weights of arcs incident at X.
  2. Consider the network obtained by ignoring X and all arcs incident to X. Find the total weight of the minimum connector for this network.
  3. The sum of the two totals is a lower bound.

- The **Tour Improvement** algorithm can be used to improve a discovered Hamiltonian tour.

  1. Let $i = 1$.
  2. If $d(V_i, V_{i+2}) + d(V_{i+1}, V_{i+3}) < d(V_i, V_{i+1}) + d(V_{i+2}, V_{i+3})$ then swap V_{i+1} and V_{i+2}.
  3. Replace i by $i + 1$.
  4. If $i \leqslant n$ then go back to 2.

Route Inspection

- The **Chinese Postman** algorithm is a method to find the least-weight closed trail containing every arc (note arc and *not* node). It is so called because a postman needs to travel along roads (arcs) when delivering letters. The algorithm states

  1. Find all nodes of odd order.
  2. For each pair of odd nodes find the connecting path of minimum weight.
  3. Pair up all the odd nodes so that the sum of the weights of the connecting paths from 2 is minimised.
  4. In the original graph, duplicate the minimum weight paths found in 3.
  5. Find a trail containing every arc for the new (Eulerian) graph.

- If you have four odd order nodes $\{A,B,C,D\}$ then there are $3 \times 1 = 3$ comparisons to make: $AB - CD$, $AC - BD$, $AD - BC$.

  If you have six odd order nodes $\{A,B,C,D,E,F\}$ then you have $5 \times 3 \times 1 = 15$ comparisons to make: $AB-CD-EF$, $AB-CE-DF$, $AB-CF-DE$, $AC-BD-EF$, $AC-BE-DF$, $AC-BF-DE$, $AD-BC-EF$, $AD-BE-CF$, $AD-BF-CE$, $AE-BC-DF$, $AE-BD-CF$, $AE-BF-CD$, $AF-BC-DE$, $AF-BD-CE$, $AF-BE-CD$.

  If you have $2n$ odd order nodes then you need to make $(2n-1) \times (2n-3) \times \ldots 3 \times 1$ comparisons. In an exam situation I can't believe they would make you do more than four odd ordered nodes.

Linear Programming

- The **Simplex** algorithm solves…