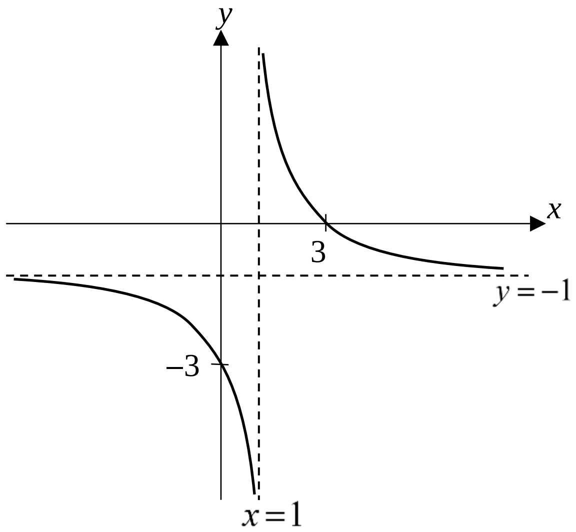
2024 TJC Preliminary Exam H2 Mathematics Paper 2 (Suggested solutions)

Section A: Pure Mathematics [40 marks]

1 A curve C has equation $y = \dfrac{3-x}{x-1}$.

(a) Sketch C, stating the equations of the asymptotes. [2]

(b) Find the exact volume of the solid generated when the region bounded by C, the x-axis and the line $x = 9$ is rotated through 2π radians about the x-axis. Leave your answer in the form $\pi(a + b\ln 2)$, where a and b are constants to be determined. [5]

	[Solutions]	Remarks
1(a)		$y = \dfrac{3-x}{x-1} = -1 + \dfrac{2}{x-1}$ Use dashed lines for asymptotes. When sketching additional lines to solve other parts of the problem, it is recommended that a side sketch be redrawn to avoid damaging the solution.
1(b)	Required volume $= \pi \int_3^9 \left(\dfrac{3-x}{x-1}\right)^2 \mathrm{d}x$ $= \pi \int_3^9 \left(-1 + \dfrac{2}{x-1}\right)^2 \mathrm{d}x$ $= \pi \int_3^9 1 - \dfrac{4}{x-1} + \dfrac{4}{(x-1)^2} \mathrm{d}x$ $= \pi \left[x - 4\ln\lvert x-1\rvert - 4(x-1)^{-1} \right]_3^9$ $= \pi\left[\left(9 - 4\ln\lvert 9-1\rvert - 4(9-1)^{-1}\right) - \left(3 - 4\ln\lvert 2\rvert - 4(2)^{-1}\right)\right]$ $= \pi\left(\dfrac{15}{2} - 4\ln\left(2^3\right) + 4\ln 2\right)$ $= \pi\left(\dfrac{15}{2} - 8\ln 2\right)$	An additional diagram drawn by the side is useful in helping identify the solid of revolution. It is more efficient to perform a long division first then squaring. Squaring first will result in a more complicated algebraic fraction that would still have to be dealt with by long division (or further splitting), which would increase the complexity of the working.

2 **(a)** Find the series expansion of $\dfrac{(8+x)^{\frac{1}{3}}}{\cos 2x}$ in ascending powers of x up to and including the term in x^2. [5]

(b) Find the range of validity of x for the expansion to be valid. [2]

	[Solutions]	**Remarks**
(a)	$\dfrac{(8+x)^{\frac{1}{3}}}{\cos 2x}=\dfrac{8^{\frac{1}{3}}\left(1+\frac{x}{8}\right)^{\frac{1}{3}}}{\cos 2x}$ $=\dfrac{2\left(1+\frac{1}{3}\left(\frac{x}{8}\right)+\frac{\frac{1}{3}\left(-\frac{2}{3}\right)}{2!}\left(\frac{x}{8}\right)^2+\cdots\right)}{1-\frac{(2x)^2}{2}+\cdots}$ $=2\left(1+\frac{x}{24}-\frac{x^2}{576}+\cdots\right)\left(1-2x^2+\cdots\right)^{-1}$ $=2\left(1+\frac{x}{24}-\frac{x^2}{576}+\cdots\right)\left(1+2x^2+\cdots\right)$ $=2\left(1+\frac{x}{24}+2x^2-\frac{x^2}{576}+\cdots\right)$ $=2+\frac{1}{12}x+\frac{1151}{288}x^2+\cdots$	• Do not differentiate and use Maclaurin's theorem. • Use binomial expansion/standard series in MF26: $(1+x)^n=1+nx+\frac{n(n-1)}{2!}x^2+\ldots, \lvert x\rvert<1$ $\cos x=1-\frac{1}{2}x^2+\ldots$, all x • Expand until x^2 term. Do not waste time expanding more than what is required. • Rewrite $\frac{1}{\left(1-2x^2\right)}$ as $\left(1-2x^2\right)^{-1}$ so that we can further expand it using binomial expansion again
(b)	For $\left(1-2x^2\right)^{-1}$: $\lvert 2x^2\rvert<1$ $\lvert x\rvert<\frac{1}{\sqrt{2}}$ $-\frac{1}{\sqrt{2}}<x<\frac{1}{\sqrt{2}}$ For $\left(1-\frac{x}{8}\right)^{\frac{1}{3}}$: $\left\lvert\frac{x}{8}\right\rvert<1 \Rightarrow -8<x<8$ Taking intersection, the range of validity is $-\frac{1}{\sqrt{2}}<x<\frac{1}{\sqrt{2}}$	Note that there are two main binomial expansions, $\left(1+\frac{x}{8}\right)^{\frac{1}{3}}$ and $\left(1-2x^2\right)^{-1}$. We need to find the validity range for each expansion and then find the range of values which satisfy both validity range by taking intersection.

3 A sequence $\{a_n\}$ is defined by $a_0 = 2$ and $a_n = a_{n-1} - \frac{2}{3}\left(\frac{1}{3}\right)^{n-2}$ where $n \geq 1$.

(a) By considering $\sum_{n=1}^{N}(a_n - a_{n-1})$, find an expression for a_N. [4]

(b) Hence explain whether the sequence is convergent. [1]

	[Solutions]	**Remarks**
3(a)	$a_n - a_{n-1} = -\frac{2}{3}\left(\frac{1}{3}\right)^{n-2}$ where $n \geq 1$ $\sum_{n=1}^{N}(a_n - a_{n-1}) = \left(-\frac{2}{3}\right)\sum_{n=1}^{N}\left(\frac{1}{3}\right)^{n-2}$ LHS $= \sum_{n=1}^{N}(a_n - a_{n-1})$ $= \cancel{a_1} - a_0$ $+ \cancel{a_2} - \cancel{a_1}$ $+ \cancel{a_3} - \cancel{a_2}$ $\vdots$ $\vdots$ $+ \cancel{a_{N-2}} - \cancel{a_{N-1}}$ $+ \cancel{a_{N-1}} - \cancel{a_{N-2}}$ $+ a_N - \cancel{a_{N-1}}$ $= a_N - a_0$ $= a_N - 2$ RHS $= \left(-\frac{2}{3}\right)\sum_{n=1}^{N}\left(\frac{1}{3}\right)^{n-2}$ $\left(= \left(-\frac{2}{3}\right)\left(\frac{1}{3}^{-1} + \frac{1}{3}^{0} + \frac{1}{3}^{1} + \dots + \frac{1}{3}^{N-2}\right)\right)$ $= \left(-\frac{2}{3}\right)\frac{\left(\frac{1}{3}\right)^{-1}\left[1 - \left(\frac{1}{3}\right)^{N}\right]}{1 - \frac{1}{3}}$ $= -3\left[1 - \left(\frac{1}{3}\right)^{N}\right]$ Thus $a_N - 2 = 3\left(\frac{1}{3}\right)^{N} - 3$ $\Rightarrow \quad a_N = 3\left(\frac{1}{3}\right)^{N} - 1$	- This is an equation. So when we take sum, we take sum **ON BOTH SIDES** of the equation. - For summation, if you cannot observe what sum is that, it is always good to write out few terms to see. - LHS: observe is sum of **DIFFRENCE of 2 SIMILAR TERMS**, so use MOD. - In MOD presentation, it is a **MUST to show the cancellation pattern.** - Pay attention to the **SUBSCRIPT** use. - RHS: Can write out a few terms to see as well to observe is sum of GP to **FINITE term**. - Sum of GP, learn to count the number of terms correctly. For summation **number of terms** can be obtained by "**upper** − **lower limit + 1**" if you did not remove any of the terms. **Pay attention to the use of capital letter vs small letter again.** - Simplify all answers.

3(b)	As $N \to \infty$, $\left(\frac{1}{3}\right)^N \to 0$. Hence $a_N \to -1$, a **finite value**. Thus the sequence is converges to -1.	- Check the concept of converge and watch necessary presentation. - Note the **difference of SEQUENCE converge vs SERIES converge.**

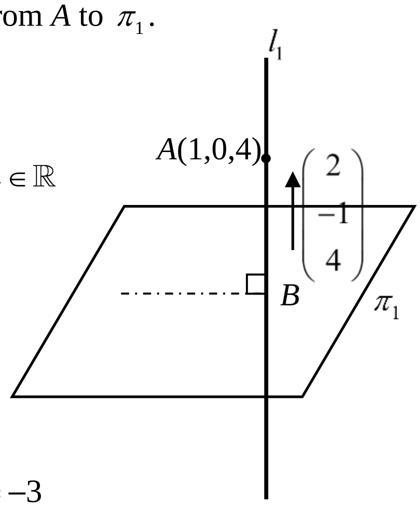
4 The line l_1 passes through the point A with coordinates $(1,0,4)$ and is perpendicular to the plane π_1 with equation $2x-y+4z=-3$.

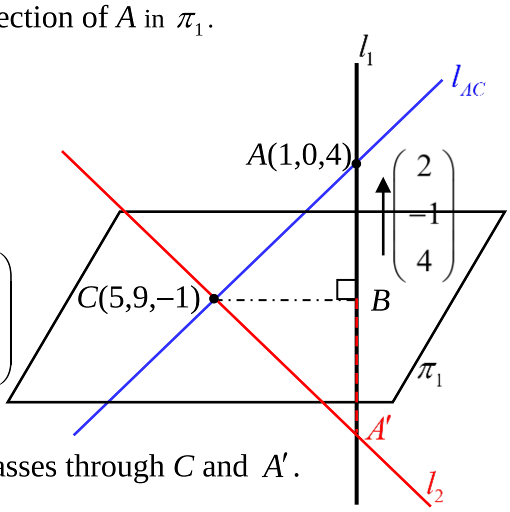
(a) Find the coordinates of the point B where l_1 meets π_1. [4]

(b) Verify that the point C with coordinates $(5,9,-1)$ lies on π_1. [1]

(c) Find a vector equation of the line l_2 which is a reflection of the line AC in π_1. [3]

(d) Find a vector equation, in scalar product form, of the plane π_2 which contains l_1 and l_2. [3]

	[Solutions]	**Remarks**
4(a)	B is the foot of perpendicular from A to π_1. B lies on l_1: $\overrightarrow{OB}=\begin{pmatrix}1\\0\\4\end{pmatrix}+\lambda\begin{pmatrix}2\\-1\\4\end{pmatrix}$ for some $\lambda\in\mathbb{R}$  B lies on π_1: $\overrightarrow{OB}\cdot\begin{pmatrix}2\\-1\\4\end{pmatrix}=-3$ Thus $\left[\begin{pmatrix}1\\0\\4\end{pmatrix}+\lambda\begin{pmatrix}2\\-1\\4\end{pmatrix}\right]\cdot\begin{pmatrix}2\\-1\\4\end{pmatrix}=-3$ $\Rightarrow\quad(2+16)+\lambda(4+1+16)=-3$ $\Rightarrow\quad 21\lambda=-21\quad\Rightarrow\quad\lambda=-1$ Thus $\overrightarrow{OB}=\begin{pmatrix}1\\0\\4\end{pmatrix}-\begin{pmatrix}2\\-1\\4\end{pmatrix}=\begin{pmatrix}-1\\1\\0\end{pmatrix}$ i.e. **coordinates** of B are $(-1,1,0)$	Take note the the direction of l_1 is indirectly given in the normal to π_1. **Notations**: Remember to mention that $\lambda\in\mathbb{R}$, and write "$\underset{\sim}{r}=$" when writing vector equations. Pay attention to details: **Coordinates** are required.
4(b)	$\overrightarrow{OC}\cdot\begin{pmatrix}2\\-1\\4\end{pmatrix}=\begin{pmatrix}5\\9\\-1\end{pmatrix}\cdot\begin{pmatrix}2\\-1\\4\end{pmatrix}=10-9-4=-3$ Thus C lies on π_1. (Verified)	Do not just write the calculation; include some statements.

4(c)	Let A' be the point of reflection of A in π_1. Using ratio theorem, $\overrightarrow{OB}=\dfrac{\overrightarrow{OA}+\overrightarrow{OA'}}{2}$ $\Rightarrow \overrightarrow{OA'}=2\overrightarrow{OB}-\overrightarrow{OA}$ $=2\begin{pmatrix}-1\\1\\0\end{pmatrix}-\begin{pmatrix}1\\0\\4\end{pmatrix}=\begin{pmatrix}-3\\2\\-4\end{pmatrix}$  The line of reflection l_2 passes through C and A'. $\overrightarrow{A'C}=\begin{pmatrix}-3\\2\\-4\end{pmatrix}-\begin{pmatrix}5\\9\\-1\end{pmatrix}=\begin{pmatrix}-8\\-7\\-3\end{pmatrix}$ Thus equation of l_2 is $\underset{\sim}{r}=\begin{pmatrix}5\\9\\-1\end{pmatrix}+\mu\begin{pmatrix}8\\7\\3\end{pmatrix},\ \mu\in\mathbb{R}$.	**Draw a good diagram**. Think carefully about the information provided; ***B*** is the midpoint of A and A', ***not C.*** Other variations include using $\overrightarrow{CB}=\dfrac{\overrightarrow{CA}+\overrightarrow{CA'}}{2}$ and so on, but these require more calculation.
4(d)	π_2 contains l_1 and l_2 A normal to π_2 is $\begin{pmatrix}8\\7\\3\end{pmatrix}\times\begin{pmatrix}2\\-1\\4\end{pmatrix}=\begin{pmatrix}31\\-26\\-22\end{pmatrix}$ Equation of π_2 is $\underset{\sim}{r}\cdot\begin{pmatrix}31\\-26\\-22\end{pmatrix}=\overrightarrow{OA}\cdot\begin{pmatrix}31\\-26\\-22\end{pmatrix}=\begin{pmatrix}1\\0\\4\end{pmatrix}\cdot\begin{pmatrix}31\\-26\\-22\end{pmatrix}$ i.e., $\underset{\sim}{r}\cdot\begin{pmatrix}31\\-26\\-22\end{pmatrix}=-57$ <u>**Alternative solution**</u> π_2 contains l_1 and l_2 $\Rightarrow$ π_2 also contains line AC A normal to π_2 is $\overrightarrow{AC}\times\begin{pmatrix}2\\-1\\4\end{pmatrix}=\begin{pmatrix}4\\9\\-5\end{pmatrix}\times\begin{pmatrix}2\\-1\\4\end{pmatrix}=\begin{pmatrix}31\\-26\\-22\end{pmatrix}$ Equation of π_2 is $\underset{\sim}{r}\cdot\begin{pmatrix}31\\-26\\-22\end{pmatrix}=\overrightarrow{OC}\cdot\begin{pmatrix}31\\-26\\-22\end{pmatrix}=\begin{pmatrix}5\\9\\-1\end{pmatrix}\cdot\begin{pmatrix}31\\-26\\-22\end{pmatrix}$ i.e., $\underset{\sim}{r}\cdot\begin{pmatrix}31\\-26\\-22\end{pmatrix}=-57$	**Refer to your diagram.** There are multiple ways to obtain this answer, even without solving part (c), as $\overrightarrow{CA},\overrightarrow{CB}$ and $\overrightarrow{AB}$ are all vectors on the required plane.

5 **(a)** The complex number $z = x + \mathrm{i}y$ is represented by the point $P(x, y)$ in an Argand diagram and satisfies the equation

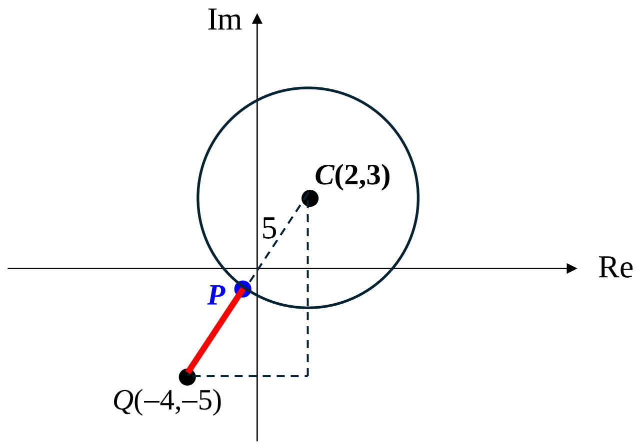
$$zz^* = (2+3\mathrm{i})z^* + (2-3\mathrm{i})z + 12 .$$

(i) Show that P is a point on a circle, and state the centre and radius of the circle. [3]

(ii) The point Q represents the complex number $-4-5\mathrm{i}$. Find the smallest possible length PQ. [2]

(b) **(i)** Show that $\dfrac{1+\cos\theta+\mathrm{i}\sin\theta}{1-\cos\theta-\mathrm{i}\sin\theta} = \mathrm{i}\cot\dfrac{\theta}{2}$. [3]

(ii) It is given that $z = \mathrm{e}^{\mathrm{i}\theta}$. Find the set of integer values of n such that $\left(\dfrac{1+z}{1-z}\right)^n$ is always real. [2]

	[Solutions]	**Remarks**
5(a)(i)	Subst $z = x + y\mathrm{i}$ into $zz^* = (2+3\mathrm{i})z^* + (2-3\mathrm{i})z + 12$ $(x+\mathrm{i}y)(x-\mathrm{i}y) = (2+3\mathrm{i})(x-\mathrm{i}y) + (2-3\mathrm{i})(x+\mathrm{i}y) + 12$ $x^2 + y^2 = 2x - 2y\mathrm{i} + 3x\mathrm{i} + 3y + 2x + 2y\mathrm{i} - 3x\mathrm{i} + 3y + 12$ $= 4x + 6y + 12$ $x^2 - 4x + y^2 - 6y = 12$ $(x-2)^2 + (y-3)^2 = 12 + 4 + 9 = 25 = 5^2$ Thus P lies on a circle with centre (2, 3) and radius 5.	This is a "**Show**" qn. Need to give the equation of a circle: $(x-h)^2 + (y-k)^2 = r^2$ Use completing squares
5(a)(ii)	 Let C be the centre of the circle. Use Pythagoras Theorem, $CQ = \sqrt{(2-(-4))^2 + (3-(-5))^2} = \sqrt{6^2 + 8^2} = 10$ Shortest length PQ = ***CQ* – radius of circle** $= 10 - 5 = 5$	**Sketch a diagram** to help you see the position of P for length PQ to be shortest. P and Q are points! *Poor presentation:* $Q \neq -4-5\mathrm{i}$ point ≠ complex no. $\lvert Q \rvert \neq \sqrt{4^2 + 5^2}$ Should be length OQ

5(b)(i)	**Proof 1**	
	$\frac{1+\cos\theta+\mathrm{i}\sin\theta}{1-\cos\theta-\mathrm{i}\sin\theta}=\frac{1+\left(2\cos^2\frac{\theta}{2}-1\right)+\mathrm{i}\left(2\sin\frac{\theta}{2}\cos\frac{\theta}{2}\right)}{1-\left(1-2\sin^2\frac{\theta}{2}\right)-\mathrm{i}\left(2\sin\frac{\theta}{2}\cos\frac{\theta}{2}\right)}$	Use double angle formula $\cos\theta=2\cos^2\frac{\theta}{2}-1$ $\cos\theta=1-2\sin^2\frac{\theta}{2}$ $\sin\theta=2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$
	$=\frac{2\cos\frac{\theta}{2}\left(\cos\frac{\theta}{2}+\mathrm{i}\sin\frac{\theta}{2}\right)}{2\sin\frac{\theta}{2}\left(\sin\frac{\theta}{2}-\mathrm{i}\cos\frac{\theta}{2}\right)}$	
	$=\cot\frac{\theta}{2}\cdot\frac{\mathrm{i}\left(\sin\frac{\theta}{2}-\mathrm{i}\cos\frac{\theta}{2}\right)}{\left(\sin\frac{\theta}{2}-\mathrm{i}\cos\frac{\theta}{2}\right)}$	*Factor out* i in the numerator (***see given answer***)
	$=\mathrm{i}\cot\frac{\theta}{2}$ (shown)	
	Proof 2	
	$\frac{1+\cos\theta+\mathrm{i}\sin\theta}{1-\cos\theta-\mathrm{i}\sin\theta}=\frac{1+\mathrm{e}^{\mathrm{i}\theta}}{1-\mathrm{e}^{\mathrm{i}\theta}}$	$\mathrm{e}^{\mathrm{i}\theta}=\cos\theta+\mathrm{i}\sin\theta$
	$=\frac{\mathrm{e}^{\mathrm{i}\frac{\theta}{2}}\left(\mathrm{e}^{-\mathrm{i}\frac{\theta}{2}}+\mathrm{e}^{\mathrm{i}\frac{\theta}{2}}\right)}{\mathrm{e}^{\mathrm{i}\frac{\theta}{2}}\left(\mathrm{e}^{-\mathrm{i}\frac{\theta}{2}}-\mathrm{e}^{\mathrm{i}\frac{\theta}{2}}\right)}$	Note that $\mathrm{e}^{\mathrm{i}\theta}+\mathrm{e}^{-\mathrm{i}\theta}=2\cos\theta$ $\mathrm{e}^{\mathrm{i}\theta}-\mathrm{e}^{-\mathrm{i}\theta}=2\sin\theta\,\mathrm{i}$
	$=\frac{\cos\left(-\frac{\theta}{2}\right)+\mathrm{i}\sin\left(-\frac{\theta}{2}\right)+\left(\cos\frac{\theta}{2}+\mathrm{i}\sin\frac{\theta}{2}\right)}{\cos\left(-\frac{\theta}{2}\right)+\mathrm{i}\sin\left(-\frac{\theta}{2}\right)-\left(\cos\frac{\theta}{2}+\mathrm{i}\sin\frac{\theta}{2}\right)}$	
	$=\frac{2\cos\frac{\theta}{2}}{-2\mathrm{i}\sin\frac{\theta}{2}}\times\frac{\mathrm{i}}{\mathrm{i}}$	All steps should be clearly shown as the final answer is given
	$=\mathrm{i}\cot\frac{\theta}{2}$ (shown)	
	Proof 3	
	$\frac{1+\cos\theta+\mathrm{i}\sin\theta}{(1-\cos\theta)-\mathrm{i}\sin\theta}\times\frac{(1-\cos\theta)+\mathrm{i}\sin\theta}{(1-\cos\theta)+\mathrm{i}\sin\theta}$	The conjugate of $(1-\cos\theta)-\mathrm{i}\sin\theta$ is $(1-\cos\theta)+\mathrm{i}\sin\theta$
	$=\frac{1-\cos^2\theta+(1+\cos\theta)\mathrm{i}\sin\theta+(1-\cos\theta)\mathrm{i}\sin\theta-\sin^2\theta}{(1-\cos\theta)^2-\sin^2\theta}$	
	$=\frac{1-\left(\cos^2\theta+\sin^2\theta\right)+2\mathrm{i}\sin\theta}{1-2\cos\theta+\left(\cos^2\theta+\sin^2\theta\right)}$	
	$=\frac{2\mathrm{i}\sin\theta}{2(1-\cos\theta)}$	
	$=\frac{\mathrm{i}\sin\frac{\theta}{2}\cos\frac{\theta}{2}}{\left(1-\left(1-2\sin^2\frac{\theta}{2}\right)\right)}$	Use double angle formula
	$=\mathrm{i}\cot\frac{\theta}{2}$ (shown)	

5(b)(ii)

Method 1 (Preferred)

$z = e^{i\theta} = \cos\theta + i\sin\theta$

Using (i), $\left(\frac{1+z}{1-z}\right)^n = \left(i\cot\frac{\theta}{2}\right)^n = (i)^n \cot^n\frac{\theta}{2}$

Link to earlier result!

See that $\cot^n\frac{\theta}{2}$ is real for all n.

To be real, $(i)^n = 1$ or -1, n is an even integer

Set of integer values of n = $\{n : n = 2k,\ k \in \mathbb{Z}\}$

Distinguish the 2 sets

$\mathbb{Z}$ = set of all **integers**

$\mathbb{R}$ = set of all **real nos**

Give answer in set notation.

Method 2 (Not preferred)

Using (i), $\left(\frac{1+z}{1-z}\right)^n = \left(i\cot\frac{\theta}{2}\right)^n$

$$\arg\left(\frac{1+z}{1-z}\right)^n = n\arg\left(i\cot\frac{\theta}{2}\right)$$
$$= n\left(\arg i + \arg\left(\cot\frac{\theta}{2}\right)\right)$$
$$= n\left(\frac{\pi}{2} + m\pi\right),\quad m \in \mathbb{Z}$$

Note that

Since $\cot\frac{\theta}{2}$ is real,

$\arg\left(\cot\frac{\theta}{2}\right) = m\pi,\ m \in \mathbb{Z}$

For $\left(\frac{1+z}{1-z}\right)^n$ to be always real, argument = $k\pi,\ k \in \mathbb{Z}$

z is real and positive

$\Leftrightarrow \arg z = 2k\pi,\ k \in \mathbb{Z}$

For $n\left(\frac{\pi}{2} + m\pi\right) = k\pi \Rightarrow \frac{n\pi}{2} + mn\pi = k\pi$

n is an even integer

Set of integer values of n = $\{n : n = 2k,\ k \in \mathbb{Z}\}$

$$\arg\left(\frac{1+z}{1-z}\right)^n = n\left(\arg(1+z) - \arg(1-z)\right)$$

Note that $\arg(1+z) \neq \arg 1 + \arg z$ ✘

Should be $\arg(zw) = \arg z + \arg w$

Section B: Probability and Statistics [60 marks]

6 **(a)** The 11 letters of the word REFRESHMENT are arranged in a row.

(i) Find the number of different arrangements that can be made. [2]

(ii) Find the number of different arrangements that can be made such that all the E's are together and all the R's are together but the E's and the R's are not together. [2]

(b) A 4-letter codeword is formed using the letters in the word REFRESHMENT. Find the number of different codewords that can be formed. [4]

	[Solutions]	**Remarks**
6(a)(i)	Required no. of arrangements = $\dfrac{11!}{2!3!} = 3\ 326\ 400$	R E F S H M N T R E E
6(a)(ii)	**<u>M1: Slot in method</u>** F S H M N T <u>RR</u> <u>EEE</u> Required no. of arrangements = $6! \times {}^{7}C_{2} \times 2! = 30\ 240$ **<u>M2: complement method</u>** EEE RR _ _ _ _ _ _ $8! - 7! \times 2! = 30\ 240$ EEE RR _ _ _ _ _ _ 8! : Group Es as one unit, Rs as one unit. Total = 6 others + Es unit + Rs unit = 8 units arrange them. 7!: Group Es as one unit, Rs as one unit. Group the Es unit and Rs unit together. Total = 6 others + (E unit and R unit together)unit = 7 units arrange them 2! : Arrange the Es unit and the Rs unit	For objects not together means **to separate.** **Use slotting in method** Extra careful use of complement. Note the use of complement must have the no restriction with Es and Rs grouped respectively into unit first! It is **not (i) answer.**

6(b)	**Case 1: All distinct letters** No. of ways $= {}^{8}C_{4} \times 4! = 1680$ **Case 2: 1 pair of identical and 2 other different letters** No. of ways $= {}^{2}C_{1} \times {}^{7}C_{2} \times \frac{4!}{2!} = 504$ **Case 3: 2 pairs of identical letters (RR & EE)** No. of ways $= \frac{4!}{2!2!} = 6$ **Case 4: 3 E's and 1 other different letter** No. of ways $= {}^{7}C_{1} \times \frac{4!}{3!} = 28$ Total no. of arrangements = 1680 + 504 + 6 + 28 = 2218	- Do not over split your cases. - Choosing number of items from identical objects is always 1 way

7 Two boys, Joseph and Elliot, play a game by each tossing a coin. Joseph tosses a 20-cent coin and Elliot tosses a 50-cent coin. The probability that the 20-cent coin and the 50-cent coin shows a head are $\frac{3}{5}$ and p respectively.

If both coins show heads, Joseph gets to keep Elliot's coin.

If both coins show tails, Joseph gives his coin to Elliot.

If one coin shows a head and the other coin shows a tail, both get to keep their own coins. Let W, in cents, be the amount of money Joseph wins in a game.

(a) Find, in terms of p, the probability distribution of W and E(W). [4]

(b) Find the value of p for the game to be fair. [1]

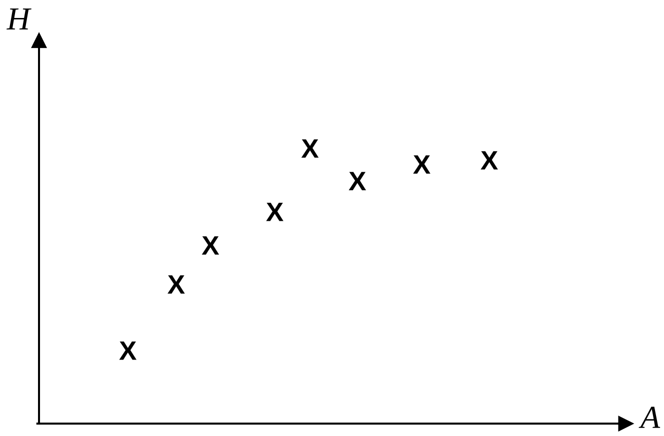
(c) Suppose Elliot's coin is fair i.e. $p = \frac{1}{2}$ and the boys played 40 games. Find the probability that Joseph wins an average of more than 15 cents per game. [3]

	[Solutions]	Remarks
7(a)	<table><tr><td>outcome</td><td>HH</td><td>HT or TH</td><td>TT</td></tr><tr><td>w</td><td>50</td><td>0</td><td>-20</td></tr><tr><td>$\mathrm{P}(W=w)$</td><td>$\frac{3}{5}p$</td><td>$\frac{3}{5}(1-p)+\frac{2}{5}p$ $=\frac{3}{5}-\frac{1}{5}p$</td><td>$\frac{2}{5}(1-p)$</td></tr></table> $\mathrm{E}(W)=50\times\left(\frac{3}{5}p\right)-20\times\left(\frac{2}{5}(1-p)\right)$ $=38p-8$	**Read the question carefully**: W represents Joseph's winnings, so his initial 20-cent coin should not be included in the value of W.
7(b)	For the game to be fair, $\mathrm{E}(W)=0$ $38p-8=0 \Rightarrow p=\frac{4}{19}$	**Important Understanding** A "fair game" is one in which the expected winnings is 0. It is **not** one where P(J wins) = P(E wins)
7(c)	Given $p=\frac{1}{2}$, $\mathrm{E}(W)=38\left(\frac{1}{2}\right)-8=11$ Using GC, Var(W) = 709 $\overline{W}=\frac{W_1+W_2+W_3+...+W_{40}}{40}$ Since n = 40 is large, by Central Limit Theorem, $\overline{W}\sim \mathrm{N}\left(11, \frac{709}{40}\right)$ approximately Required probability = $\mathrm{P}\left(\overline{W}>15\right)=0.171$ (3 s.f.)	The question asks about Joseph's **average** winnings. This is a big hint for the application of CLT. Remember to justify the use of CLT, and that CLT **does not** let the original r.v. (i.e. W) become normal. It only applies to $\overline{W}$.

8 A trainee nurse Angeline is investigating how the head circumferences of young children vary with age. The age, A months, and the head circumference, H cm, of a random sample of 8 young children are given in the table.

A	2	5	7	11	13	16	20	24
H	34.5	38.5	41	43	47	45	46.3	46.5

(a) The value of the product moment correlation coefficient between H and A is 0.880, correct to 3 decimal places, and a scatter diagram for the data is shown below.



(i) Explain whether a linear model is a good model for the relationship between H and A. [1]

(ii) Identify one of the data that Angeline may have recorded wrongly and justify your answer. [1]

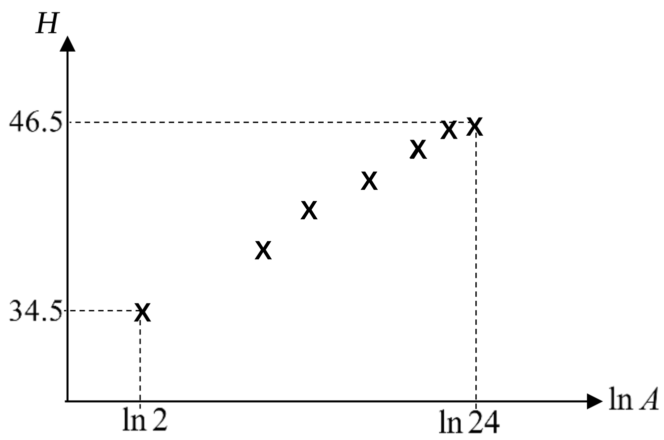
(b) **For the rest of this question, you should not include the wrongly recorded data.**

Based on the scatter diagram in **(a)**, Angeline thinks that a model with equation $H = a + b\ln A$ is an appropriate model.

(i) Sketch a scatter diagram for H against $\ln A$. [1]

(ii) Use your calculator to find the equation of the least square regression line of H on $\ln A$ and the value of the corresponding product moment correlation coefficient. [2]

(iii) Use your equation to estimate the head circumference of a 13-month-old child. Give two reasons why you would expect this estimate to be reliable. [3]

	[Solutions]
8(a) (i)	Although $r_{A,H} = 0.880$ is relatively close to 1, the scatter **diagram shows** the general trend that H increases at a decreasing rate **rather than a constant rate**. Hence a linear model is **not** appropriate. **Comments:** - r-value of 0.88 is not low and is not indicating weak linear. - **Scatter diagram is always a much accurate way than *r*-value** to see the suitability of a linear model. - **Question has given all data and is not asking to explain in context. So students need to learn to analyse questions when to use which.** - To explain WHY LINEAR cannot, hence the need to explain what about LINEAR model ie "constant rate" that cause it to be not suitable. **MUST do some comparative.**
8(a) (ii)	(13, 47) as it does not follow the increase at a decreasing rate **trend set by the rest of the data.** Or (13, 47) as it is **FAR away** from the increase at a decreasing rate graph formed **by the rest of the data.** **Comments:** - Important concept (***) **DATA points can be above or below the trend set by the rest of the data.** - Use of keywords "rest of the data", "follow", "far away", "trend/graph" - **CANNOT** use instantaneous point i.e. relative position of points to argue why point is outlier. (see ***) **MUST** argue is about the point **with respect to the trend/ graph.** - Must state the trend as one could pick another point eg A =2 to say that it does not follow the linear trend set by the rest of the points!
8(b) (i)	 4 things to note when drawing scatter diagram 1) Max and min of respective axis 2) Relative position of points/ Shape form 3) Total number of points 4) Axis

8(b) (ii)	From GC, regression line of H on $\ln A$ is $H = 5.0349\ln A + 30.899$ $H = 5.03\ln A + 30.9$ (3 s.f.) and $r = 0.997$ (3 s.f.)
8(b) (iii)	When $A = 13$, $H = 5.0349\ln 13 + 30.899$ = 43.8 (3 s.f.) Thus the head circumference of a 13-month-old child is 43.8 cm. It is a reliable estimate as $A = 13$ is within the data range [2, 24] **and** r = 0.997 being **close to 1** indicates a **strong linear** correlation **between ln *A* and *H***. <u>Reliability must remember 2 'R's</u> 1) Range. In particular is the **range of sub in value** and NOT range of estimate obtained. 2) R-value. Pay attention to **all the keywords needed.** -Be clear in your answer, DO NOT USE 'it' as too ambiguous to know what you are referring to. - Final answer is non- exact, please remember to give 3 s.f.

9 In this question, you should state the parameters of any normal distributions you use.

In golf, a player's driving distance refers to the distance a ball travels when it is hit from the tee using a golf club known as a driver.

Records from past competitions show the following statistics.

- The proportion of male players with driving distance greater than 261 metres is equal to the proportion of male players with driving distance less than 170 metres.
- The proportion of male players with driving distance greater than 261 metres is equal to the proportion of female players with driving distance greater than 181 metres.
- The mean driving distance of a female player is 147.5 metres.

It may be assumed that the driving distances of male and female players follow normal distributions. The standard deviations of the driving distances of male and female players are denoted by σ_m and σ_f respectively.

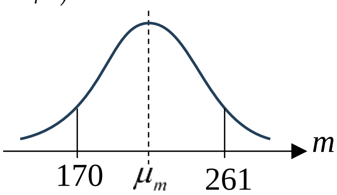
(a) State the mean driving distance of a male player. [1]

(b) Show that $\sigma_m = k\sigma_f$, where k is a constant to be determined. [2]

It is given that $\sigma_m = 24.14$ and $\sigma_f = 17.77$.

(c) Find the probability that the driving distance of a randomly chosen male player is more than 1.5 times the driving distance of a randomly chosen female player. [3]

(d) Find the probability that the difference in driving distances of two randomly chosen female players is less than 15 metres. [3]

	[Solutions]	**Remarks**
9(a)	Let M and F denote the driving distances (in metres) of **a** male and **a** female player respectively. $M \sim \mathrm{N}\left(\mu_m, \sigma_m^{\ 2}\right)$ and $F \sim \mathrm{N}\left(147.5, \sigma_f^{\ 2}\right)$. Given: $\mathrm{P}(M > 261) = \mathrm{P}(M < 170)$  **Then, by symmetry**, $\mu_m = \frac{170+261}{2} = 215.5$ Thus the mean driving distance of a male player is 215.5 metres.	Note that $\mu_m = \mathrm{E}(M)$ (population mean)
9(b)	Given: $\mathrm{P}(M > 261) = \mathrm{P}(F > 181)$ $\mathrm{P}\left(Z > \frac{261-215.5}{\sigma_m}\right) = \mathrm{P}\left(Z > \frac{181-147.5}{\sigma_f}\right)$ $\frac{45.5}{\sigma_m} = \frac{33.5}{\sigma_f}$ $\frac{\sigma_m}{\sigma_f} = \frac{45.5}{33.5}$ $\sigma_m = \frac{91}{67}\sigma_f$ or $1.36\,\sigma_f$ (3 s.f.)	**Need to show standardisation**

9(c)	Given: $M \sim \mathrm{N}\left(215.5\,,24.14^2\right)$ and $F \sim \mathrm{N}\left(147.5\,,17.77^2\right)$ $\mathrm{E}\left(M-\frac{3}{2}F\right)=215.5-\frac{3}{2}(147.5)=-5.75$ $\mathrm{Var}\left(M-\frac{3}{2}F\right)=24.14^2+\left(\frac{3}{2}\right)^2(17.77)^2=1293.228625\ \text{(exact)}$ Hence $M-\frac{3}{2}F \sim \mathrm{N}\left(-5.75\,,1293.228625\right)$ Required probability $=\mathrm{P}\left(M>\frac{3}{2}F\right)=\mathrm{P}\left(M-\frac{3}{2}F>0\right)=0.436\ \text{(3 s.f.)}$	Give the final **values** of the mean and variance
9(d)	$\mathrm{E}\left(F_1-F_2\right)=\mathrm{E}\left(F\right)-\mathrm{E}\left(F\right)=0$ $\mathrm{Var}\left(F_1-F_2\right)=2\mathrm{Var}\left(F\right)=2\left(17.77\right)^2=631.5458\ \text{(exact)}$ Hence $F_1-F_2 \sim \mathrm{N}\left(0,\,631.5458\right)$ Required probability $=\mathrm{P}\left(\lvert F_1-F_2\rvert<15\right)=\mathrm{P}\left(-15<F_1-F_2<15\right)$ $=0.449\ \text{(3 s.f.)}$	$\mathrm{P}\left(\lvert X\rvert<a\right)=\mathrm{P}\left(-a<X<a\right)$ $\mathrm{P}\left(\lvert X\rvert>a\right)=1-\mathrm{P}\left(-a\le X\le a\right)$

10 A ministry spokesman reported that students spend an average of 6.5 hours per week on co-curricular activities (CCA) in school. Mr Gru believes that the average time spend on CCA per week in his school is less than this average. To test his belief, he tasks his student Kevin to take a random sample of 50 students in his school. The times, x hours, spent on CCA per week are summarised below.

$$\sum x = 306.68 \qquad \sum x^2 = 1916.22$$

(a) State what it means for a sample to be random in this context. [1]

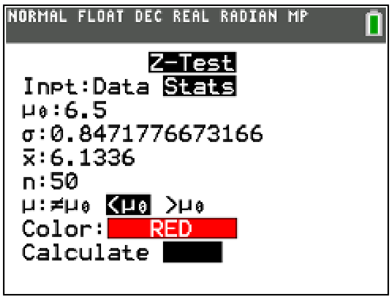
(b) Calculate unbiased estimates of the population mean and variance of the times spent on CCA per week. [2]

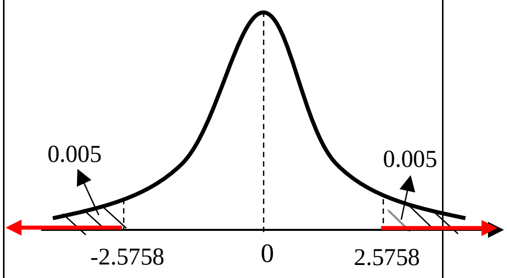
(c) Carry out a test and determine whether the p-value provides strong evidence to support Mr Gru's belief. [4]

(d) Kevin suggests to Mr Gru that it is necessary to assume that the times spent on CCA per week is normally distributed in order to carry out the test. Explain whether this assumption is necessary. [1]

(e) Student Bob takes another random sample of 50 students and finds that the mean and standard deviation of their times spent on CCA per week are m hours and 0.9 hours respectively. The result of a test at the 1% significance level is that the average time spent on CCA per week by students in his school differs from the average time reported by the ministry spokesman. Find the range of values of m. [5]

	[Solutions]	**Remarks**
10(a)	A random sample is obtained when the students **are selected independently** and ***every*** student in the school has an ***equal probability*** of being ***selected***.	To indicate random sample, need to state 2 factors: 1. equal chance/probability of being chosen 2. **selection** is independent Note that it is the selection that is independent and not the time spend on CCA is independent. Do not write probability of a student chosen is independent of another student chosen.
10(b)	X denotes the time spent (in hours) by a student in Mr Gru's school on CCA per week. Unbiased estimate of population mean, $\bar{x} = \frac{306.68}{50} = 6.1336$ **(exact)** Unbiased estimate of population variance, $s^2 = \frac{1}{49}\left[1916.22 - \frac{306.68^2}{50}\right] \approx 0.71771 = 0.718$ (3 s.f.)	DO NOT ROUND OFF EXACT DECIMALS TO 3 s.f.

10(c)	Let μ be the population mean time (in hours) spent on CCA by students in Mr Gru's school. $H_0 : \mu = 6.5$ $H_1 : \mu < 6.5$ Under H_0, since $n = 50$ is large, $\overline{X} \sim N\left(6.5, \frac{0.71771}{50}\right)$ approximately by **Central Limit Theorem**. Test statistic, $Z = \dfrac{\overline{X} - 6.5}{\sqrt{\dfrac{0.71771}{50}}} \sim N(0,1)$ approx Using GC, p-value = 0.00111 Since p-value of 0.00111 is very small, this indicates that H_0 will only be rejected if level of significance is at least 0.111% which is **very small**. This means that there is **very strong** evidence to support Mr Gru's belief.	Define μ Under H_0, μ=6.5 $\overline{X} \sim N\left(6.1336, \frac{0.71771}{50}\right)$ ✗ $Z = \dfrac{6.1336 - 6.5}{\sqrt{\dfrac{0.71771}{50}}} \sim N(0,1)$ ✗  Answer the question! Decide if p-value obtained provides strong evidence to support H_1. The smaller the p-value, the stronger the evidence!
10(d)	It is not necessary as the sample size of 50 is large and thus by Central Limit Theorem, the **sample mean** of the **times spent on CCA per week**, i.e. $\overline{X}$ would follow a normal distribution approximately.	State clearly and explicitly what exactly is normally distributed under CLT! Common wrong answers: 1. It is not necessary as the sample size of 50 is large and thus by CLT **it** is normally distributed. 2. It is not necessary as the sample size of 50 is large and thus by CLT, **the time** is normally distributed. 3. It is not necessary as the sample size of 50 is large and thus by CLT **the mean time** is normally distributed.

10(e)	Using Bob's sample, $s^2 = \frac{50}{49}\left(0.9^2\right)$ $H_0 : \mu = 6.5$ $H_1 : \mu \neq 6.5$ Under H_0, since n = 50 is large, by Central Limit Theorem, $\overline{X} \sim N\left(6.5\,,\frac{0.9^2}{49}\right)$ approximately. Test statistic, $Z = \dfrac{\overline{X} - 6.5}{\sqrt{\dfrac{0.9^2}{49}}} \sim N(0,1)$ Level of significance: 1% Critical region: $z \leq -2.5758$ or $z \geq 2.5758$ To conclude that the mean time spent on CCA per week by students in Mr Gru's school differs from the average time reported by the ministry spokesman, H_0 is rejected, i.e., $\dfrac{m-6.5}{\left(\frac{0.9}{7}\right)} \leq -2.5758$ or $\dfrac{m-6.5}{\left(\frac{0.9}{7}\right)} \geq 2.5758$ $0 \leq m \leq 6.17$ or $m \geq 6.83$ (3 s.f.)	m = sample mean 0.9 = sample standard deviation, NOT population s.d Since σ^2 is unknown, we need to estimate it using s^2. Recall: $s^2 = \frac{n}{n-1}\left(\text{sample variance}\right)$  invnorm(0.99, 0,1,CENTER)

11 On average, Alex sleeps less than 6 hours on 75% of nights. The probability that he wakes up late on a school day is 0.625. On days where he wakes up late for school, there is a 96% chance that he has slept less than 6 hours the night before.

(a) Find the probability that he wakes up late when he has slept less than 6 hours the night before. [3]

(b) Determine, with justification, whether the event that he wakes up late for school is independent of the event that he has slept less than 6 hours. [1]

A school week has 5 school days. The number of days he wakes up late for school in a school week is denoted by X.

(c) State, in context, 2 assumptions needed for X to be well-modelled by a binomial distribution. [2]

Assume now that X can be modelled by a binomial distribution.

(d) Find the probability that, in a randomly chosen week, Alex wakes up late for school on at most 3 days. [1]

A school term has 10 weeks.

(e) Find the probability that Alex wakes up late for school on at most 3 days in a week for more than 4 weeks in a randomly chosen school term. State the distribution that you use. [3]

(f) Find the probability that Alex wakes up late for school on 32 days in a randomly chosen school term. State the distribution that you use. [2]

(g) Find the probability that Alex wakes up late for school on at most 3 days in a week for 4 weeks and wakes up late for school on 4 days in a week for the other weeks in a randomly chosen school term. [2]

	[Solutions]	**Remarks**
11(a)	Let S be the event that Alex sleeps less than 6 hours. Let L be the event that Alex wakes up late for school. Given: $\mathrm{P}(S)=0.75$, $\mathrm{P}(L)=0.625$ and $\mathrm{P}(S\mid L)=0.96$ $\Rightarrow \dfrac{\mathrm{P}(L\cap S)}{\mathrm{P}(L)}=0.96$ $\Rightarrow \mathrm{P}(L\cap S)=0.96(0.625)=0.6$ Required probability $=\mathrm{P}(L\mid S)=\dfrac{\mathrm{P}(L\cap S)}{\mathrm{P}(S)}=\dfrac{0.6}{0.75}=0.8$	A logical question to ask yourself when trying to determine whether the given scenario is a $P(A\mid B)$ scenario or $P(A\cap B)$ scenario is whether A and B are occurring simultaneously or A is occurring under the constraint that B has already occurred. Need to remember the conditional probability “formula” correctly.

11(b)	Since $\mathrm{P}(S \mid L) = 0.96 \neq 0.75 = \mathrm{P}(S)$, S and L are **not independent** events.	More than often, determination of independence between two events is a mathematical question on whether either of the conditions $P(A \mid B) = P(A)$ or $P(A \cap B) = P(A) \times P(B)$ has been satisfied.
11(c)	The 2 assumptions are (1) Whether Alex wakes up late on **a** school day is **independent** of whether he wakes up late on **other** days. (2) The probability that Alex wakes up late is the **same constant** for <u>**all** school days</u>.	Always answer using the context of the questions. Do not generically specify the terms “trial” or “outcome”, without mention of the scenario. <u>Fixed no. of trials</u> and <u>two outcomes only</u> are often already implied in the question. Hence they are usually not assumptions. When phrasing the assumption on the independence of events, it is recommended to have the structure “whether … **instance of event A** happens is independent of the occurrences of **other instances of event A** i.e. repeat the event when comparing. This will avoid misphrasing.
11(d)	Given: X be the number of days Alex wakes up late for school in a school week of 5 days. $X \sim \mathrm{B}(5, 0.625)$ Required probability $= \mathrm{P}(X \leq 3) = 0.61853$ (5 s.f.) $= 0.619$ (3 s.f.)	Always give an intermediate value that is 2 more significant figures than the requirement in case a subsequent part requires the carry-over of the more accurate value.

11(e)	Let Y be the number of weeks (in a school term of 10 weeks) where he wakes up late for school on at most 3 days in a week. $Y \sim \mathrm{B}(10,\ 0.61853)$ Required probability $= \mathrm{P}(Y > 4)$ $= 1 - \mathrm{P}(Y \le 4) = 0.863$ (3 s.f.)	When defining the distribution, always use the more accurate value (with more significant figures) if a previous value is required.
11(f)	Let W be the number of days (in a school term of $10 \times 5 = 50$ days) where he wakes up late for school. $W \sim \mathrm{B}(50,\ 0.625)$ Required probability $= \mathrm{P}(W = 32) = 0.114$ (3 s.f.)	
11(g)	$X \sim \mathrm{B}(5,\ 0.625)$ Required probability $= \dfrac{10!}{4!\ 6!}\left[\mathrm{P}(X \le 3)\right]^4 \left[\mathrm{P}(X = 4)\right]^6$ = 0.0169 (3 s.f.)	Consider the following: You have 10 weeks. 4 of them needs Alex to wake up late for at most 3 times a week. 6 of them requires Alex to wake up 4 times a week. How many combinations of such a scenario are there?