

Solutions and Comments for H2 Mathematics P2 2026

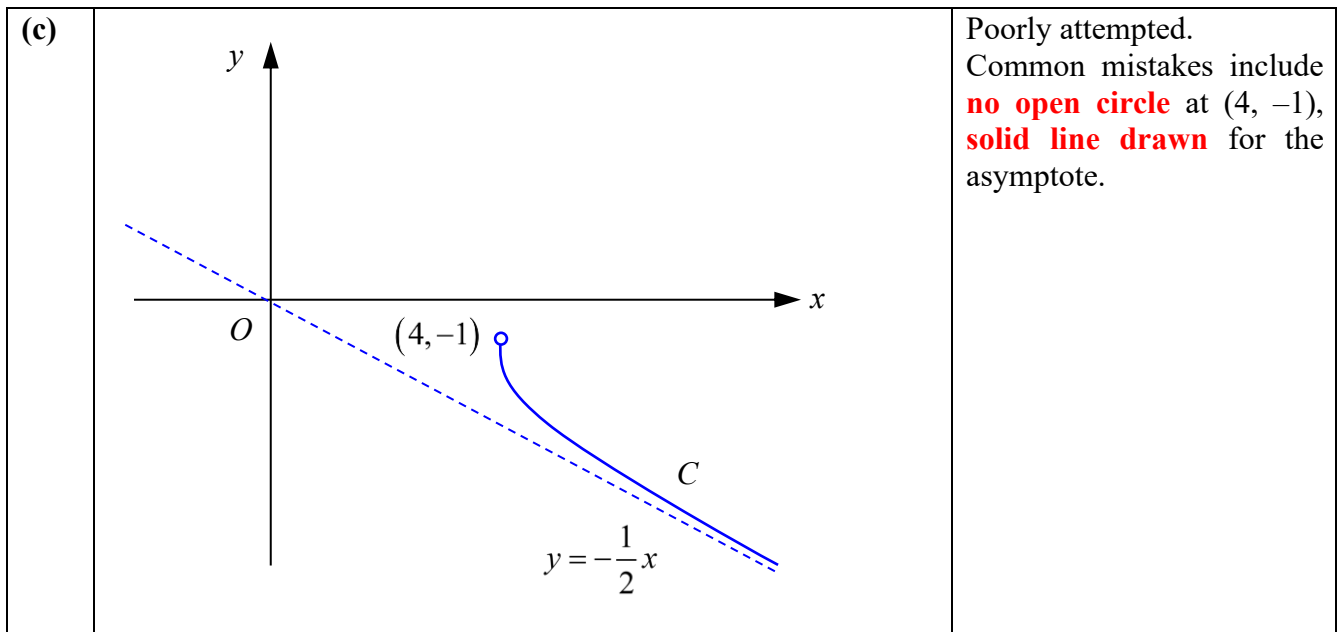
Section A: Pure Mathematics [40 marks]

1 The curve C has parametric equations

$$x = t + \frac{4}{t}, \quad y = -\frac{2}{t}, \quad 0 < t < 2.$$

- (a) Show $\frac{dy}{dx} = \frac{2}{t^2 - 4}$. Hence find the value of $\frac{dy}{dx}$ as $t \rightarrow 0$. [3]
- (b) The equation of an asymptote of C is of the form $y = kx$, where k is a real constant. Find the value of k . [1]
- (c) Sketch C , indicating clearly the coordinates of the endpoint of the curve and the equations of any asymptotes. [2]

1	Solution [6] Parametric curve	
(a)	$x = t + \frac{4}{t}, \quad y = -\frac{2}{t}$ $\frac{dx}{dt} = 1 - \frac{4}{t^2}, \quad \frac{dy}{dt} = \frac{2}{t^2}$ $\frac{dy}{dx} = \frac{dy}{dt} \div \frac{dx}{dt}$ $= \frac{\frac{2}{t^2}}{1 - \frac{4}{t^2}}$ $= \frac{\frac{2}{t^2}}{\frac{t^2 - 4}{t^2}}$ $= \frac{2}{t^2 - 4}$ As $t \rightarrow 0$, $\frac{dy}{dx} \rightarrow \frac{2}{0 - 4} = -\frac{1}{2}$	Generally well done. More working should be presented in trying to arrive at $\frac{2}{t^2 - 4}$ since it is a 'shown question'.
(b)	From (a), since $\frac{dy}{dx} \rightarrow -\frac{1}{2}$ (when $t \rightarrow 0, x \rightarrow \infty$), then the asymptote of C should be given by $y = -\frac{1}{2}x, k \in \mathbb{R}$. Hence, $k = -\frac{1}{2}$.	Majority of the students who obtained $k = -\frac{1}{2}$ did not provide the correct reason.



- 2 A start up company is launching a project to capture airborne dust particles using two different filtration models A and B , with its efficiency measured over time.

Model A is programmed such that the amount of dust particles captured in its first month of operation is 8 units and increases by a fixed amount of d units each month. In its first 12 months of operation, Model A captured a total of 360 units of dust particles.

- (a) Find the amount of dust particles captured by Model A in its 20th month of operation. [3]

Model B uses a depleting chemical catalyst to capture dust particles. The amount of dust particles captured in the first month is a units. For each subsequent month, the amount of dust particles captured is r times that of the previous month. Over its lifespan, Model B can capture a theoretical maximum of 1000 units of dust particles. In the first two months, Model B captured a total of 640 units of dust particles.

- (b) Find the values of a and r . [4]

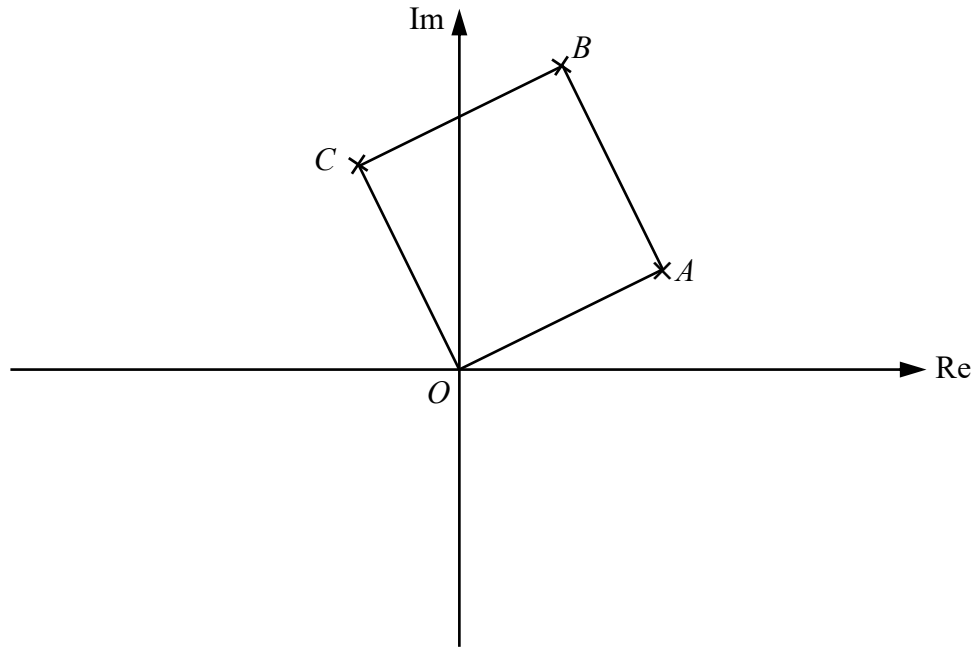
The startup company will decommission Model B once its total amount of dust particles captured reaches at least 99% of its maximum capacity over its lifespan.

- (c) Find the minimum number of months that Model B will be deployed before it is decommissioned. [3]

2	Solution [10] AP/GP	
(a)	$S_{12} = \frac{n}{2}(2a + (n-1)d)$ $360 = \frac{12}{2}(2(8) + 11d)$ $360 = 96 + 66d$ $66d = 264$ $d = 4$ $u_{20} = a + 19d$ $= 8 + 19(4)$ $= 84$ Model A captured 84 units of dust particles in its 20th month of operation.	Generally, well done. Though there were occasional mistakes. Common mistakes: - Wrong formula usage: like $S_{12} = \frac{12}{2}(8 + (11)d)$. - Expanding incorrectly like: $6(16 + 11d) = 96 + 11d$. - Interpreting that we require the total particles captured, rather than just the amount in the 29th month. - Interpreting 360 as the amount collected in the 12th month only.
(b)	$S_{\infty} = \frac{a}{1-r} = 1000$ $a = 1000(1-r) \quad \text{---(1)}$ $S_2 = a + ar = a(1+r) = 640 \quad \text{---(2)}$ Substitute (1) into (2):	Generally well-done. There were occasional mistakes too. Common mistakes: - Unable to interpret the “<i>theoretical maximum</i>” information, resulting in incorrect equations like $\frac{a(r^n - 1)}{r - 1} = 1000$.

	$a(1+r) = 640$ $1000(1-r)(1+r) = 640$ $1-r^2 = \frac{640}{1000}$ $r^2 = 0.36$ $r = \pm 0.6$ $r = 0.6$ (rej $r < 0$, amount of dust particles cannot be -ve) $a = 1000(1-0.6)$ $a = 400$	- Writing infinite geometric sum incorrectly, like $S_{\infty} = \frac{1}{a-r}$. - Writing $a + ar = \frac{a(r-1)^2}{r-1}$ instead of $\frac{a(r^2-1)}{r-1}$. - Accepting $r = -\frac{3}{5}$ because it is ‘decreasing’! (But this means there is negative amount of dust! In fact, $0 < r < 1$ would implying the amount of dust is decreasing in this context.)	
(c)	Let n be the number of months that Model B will be deployed before it is decommissioned. $S_n \geq 0.99S_{\infty}$ $\frac{a(1-0.6^n)}{1-0.6} \geq 0.99\left(\frac{a}{1-0.6}\right)$ $1-0.6^n \geq 0.99$	Though this question was well attempted. Many students did not arrive at the correct final answer.	
	Method 1: $0.6^n \leq 0.01$ $n \ln 0.6 \leq \ln 0.01$ $n \geq \frac{\ln 0.01}{\ln 0.6}$ $n \geq 9.015$	Method 2: By GC, when $n = 9$, $1-0.6^n = 0.9899 < 0.99$ when $n = 10$, $1-0.6^n = 0.994 > 0.99$	It was heartening to see that students realise $\ln 0.6 < 0$ and changed the inequality sign. Reminder: Students cannot simply state ‘By GC’ without showing further working! Please reminder to state the table or sketch the graph from the GC!
	Model B will be deployed for a minimum of <u>10 months</u> before it is decommissioned.	There were a significant number of students who arrived at 9 months instead due to incorrect rounding or misreading from their table.	

3



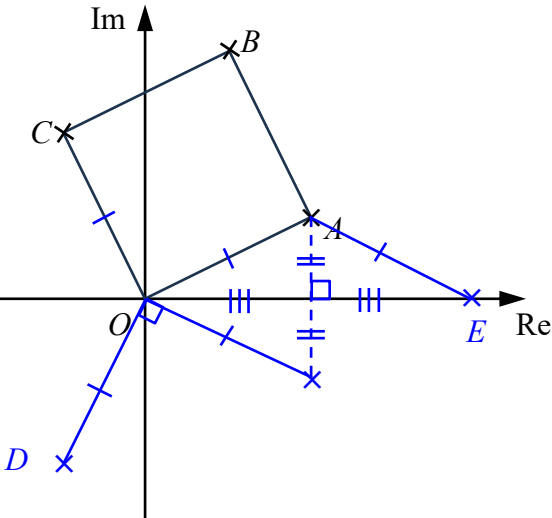
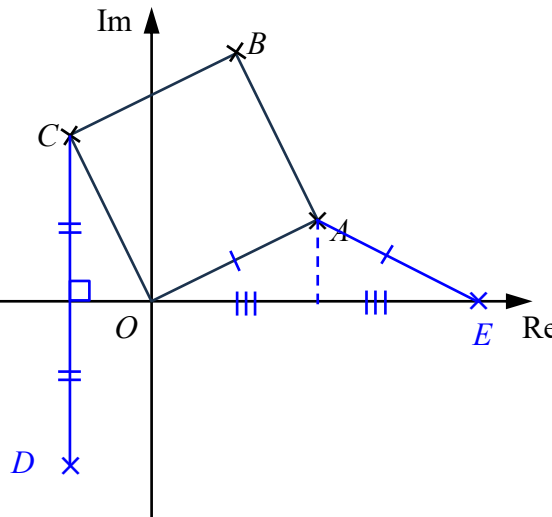
The point A on the Argand diagram represents the complex number z , where $|z| = 3$ and $\arg(z) = \frac{\pi}{7}$. Points B and C are on the diagram such that $OABC$ is a square.

- (a) (i) The points B and C represent complex numbers v and w respectively. Find v and w in terms of z . [2]
 (ii) State the exact values of $|v|$ and $\arg(v)$. [2]

Points D and E represent the complex numbers $-iz^*$ and $2\operatorname{Re}(z)$ respectively.

- (b) (i) On the same Argand diagram in the Printed Answer Booklet, plot the points D and E . Show clearly the geometrical relationship between the points C and D , and between points A and E . [2]
 (ii) Show that the complex number representing $\overline{DB}$ is $z + iz + iz^*$. [1]
 (iii) Show that BD is perpendicular to CE . [2]

3	Solutions [9] Complex Numbers	
(a)(i)	iz represents the point C , hence $w = iz$. $z + iz = (1 + i)z$ represents the point B , hence $v = (1 + i)z$.	When attempted, this was well done. Students are able to identify w more successfully as compared to v .
(a)(ii)	$ v = OB = \sqrt{OA^2 + OC^2} = \sqrt{18} = 3\sqrt{2}$. $\arg(v) = \angle BOA + \arg(z) = \frac{\pi}{4} + \frac{\pi}{7} = \frac{11\pi}{28}$.	Students who saw most success appreciated the geometry of the problem. For example, they realise $ v $ was just the length of the diagonal of the square, i.e. OB . A minority of students tried to use modulus

		properties, but were unable to apply argument properties. These are not in syllabus anymore, so students are discouraged to apply them as it might blind them from another solution. A common mistake was deriving w and $\arg(w)$ instead. Several students also had misconceptions about the modulus brackets, usually splitting the sum incorrectly like $z + iz = z + i z = 3 + 3i$.
(b)(i)	$-iz^* = (-i)(z^*)$. Consider the point representing z^*, then rotate it 90° clockwise about O. 	This question proved challenging to students. Although some students marked the points correctly, many attempts did not show the required geometrical relations between the points. This meant indicating equal lengths, angles, special angles, or parallel lines, etc ... Students had far more success in labelling E as compared to D.
(b)(i)	Alternative $(-iz^*)^* = (-i)^*(z^*)^* = iz$, so $-iz^*$ and iz are conjugates. 	Students did notice that D was the reflection of C. However, they did not managed to capture the full geometrical description of the reflection. They labelled lengths OC and OD are equal, but that is not sufficient to indicate a reflection. It is important that CD is perpendicular to the Re-axis as well.

(b)(ii)	Complex number that represents the vector $\overrightarrow{DB}$ $= z_B - z_D$ $= (1+i)z - (-iz^*)$ $= z + iz + iz^*.$	If attempted, it was well done!
(b)(iii)	Complex number that represents the vector $\overrightarrow{EC}$ $= z_C - z_E$ $= iz - 2\operatorname{Re}(z)$ $i(z + iz + iz^*) = iz - z - z^*$ $= iz - [z + z^*]$ $= iz - 2\operatorname{Re}(z).$ This implies that rotating $\overrightarrow{DB}$ 90° anticlockwise about D will produce the vector $\overrightarrow{EC}$. Hence, BD is perpendicular to EC.	The question proved challenging to students. Students who attempted had the right intuition to first deduce complex number representing vector $\overrightarrow{CE}$. There was good intuition of defining $z = a + ib$, however the rigor at which this is done is sloppy. It is essential to mention that $a, b \in \mathbb{R}$, if not a, b could be complex numbers, i.e. could $a = 2 + 3i$? That would break most of the workings that followed. Students tried to use the dot product. However many assumed the dot product was equivalent to matrix multiplication, which is false. Some wanted to use 2D vectors. But incorrectly equated the 2D or 3D vectors to the complex numbers. The vectors represent them, but are not equal to them. The methods which saw most success realise that $z + z^* = 2\operatorname{Re}(z)$ and showed that the complex numbers involved are multiplications of i of one another.

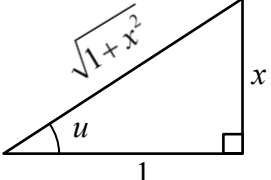
- 4 (a) Find $\int \frac{1-x^2}{(1+x^2)^2} dx$ using the substitution $x = \tan u$. [6]

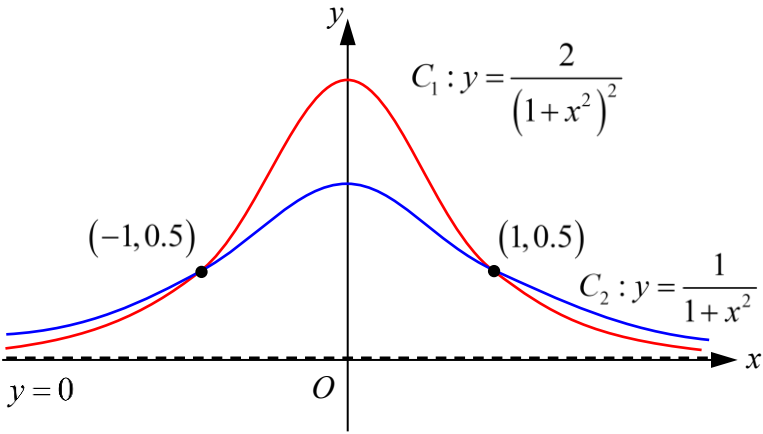
The curves C_1 and C_2 have equations $y = \frac{2}{(1+x^2)^2}$ and $y = \frac{1}{1+x^2}$ respectively.

- (b) Sketch C_1 and C_2 on the same graph, giving the coordinates of any points where C_1 and C_2 intersect and the equations of any asymptotes. [3]

The finite region R is bounded by C_1 and C_2 .

- (c) Using the result in part (a), find the area of the region R . [3]
 (d) Find the volume of the solid generated when the region R is rotated 2π radians about the x -axis. [3]

4	Solution [15] Integration by Substitution + Applications of Integration	
(a)	$x = \tan u \Rightarrow \frac{dx}{du} = \sec^2 u.$ $\cos u = \frac{1}{\sqrt{1+x^2}}, \sin u = \frac{x}{\sqrt{1+x^2}}$  $\int \frac{1-x^2}{(1+x^2)^2} dx = \int \frac{1-\tan^2 u}{(1+\tan^2 u)^2} (\sec^2 u) du$ $= \int \frac{1-\tan^2 u}{1+\tan^2 u} du$ $= \int \frac{1-\tan^2 u}{\sec^2 u} du$ $= \int \cos^2 u - \sin^2 u du$ $= \int \cos 2u du$ $= \frac{1}{2} \sin 2u + c$ $= \frac{1}{2} (2) \sin u \cos u + c$ $= \left(\frac{x}{\sqrt{x^2+1}} \right) \left(\frac{1}{\sqrt{x^2+1}} \right) + c$ $= \frac{x}{x^2+1} + c.$	Students are unable to arrive at $\int \cos 2u du$ efficiently. Majority of the students who arrive at the answer $\frac{1}{2} \sin 2u$ leave their answer as $\frac{1}{2} \sin 2(\tan^{-1} x)$. A number of students differentiate $\tan u$ by applying quotient rule on $\frac{\sin u}{\cos u}$. Wrong derivative $\frac{dx}{du} = -\sec^2 u$ or $\frac{dx}{du} = \operatorname{cosec}^2 u$ were also seen.

(b)		Asymptotic behaviour is badly drawn with overlapping tail end graphs or tail end graphs touching the axes.
(c)	$\int_{-1}^1 \frac{2}{(1+x^2)^2} - \frac{1}{1+x^2} dx$ $= \int_{-1}^1 \frac{2 - (1+x^2)}{(1+x^2)^2} dx$ $= \int_{-1}^1 \frac{1-x^2}{(1+x^2)^2} dx$ $= \left[\frac{x}{1+x^2} \right]_{-1}^1$ $= \left(\frac{1}{2} \right) - \left(\frac{-1}{2} \right)$ $= 1 \text{ units}^2.$	Quite a number of students: 1. Use GC to calculate the area without using (a) result. 2. Attempt to find the area by $2 \int_0^1 \frac{2}{(1+x^2)^2} - \frac{1}{1+x^2} dx$ instead of $\int_{-1}^1 \frac{2}{(1+x^2)^2} - \frac{1}{1+x^2} dx$ 3. Introduce modulus sign in their working which is unnecessary. Common mistakes include wrong limits in the integration. $\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{2}{(1+x^2)^2} - \frac{1}{1+x^2} dx.$
(d)	Volume $= \pi \int_{-1}^1 \left(\frac{2}{(1+x^2)^2} \right)^2 dx - \pi \int_{-1}^1 \left(\frac{1}{1+x^2} \right)^2 dx = 7.89 \text{ units}^3.$	Quite a number of students 1. Consider $\pi \int_{-1}^1 \left(\frac{2}{(1+x^2)^2} - \frac{1}{1+x^2} \right)^2 dx$ instead. 2. Did not realise that GC can be used to solve the question.

Section B: Statistics [60 marks]

- 5** The probability distribution of a discrete random variable X is given as follows:

x	-2	0	2	4
$P(X=x)$	p	0.2	q	0.3

Given $E(X) = 1.2$,

- (a) Find the values of p and q .

[3]

- (b) Find the value of $\text{Var}(X)$.

[1]

5	Solution [4] DRV	
(a)	$p + 0.2 + q + 0.3 = 1$ $p + q = 0.5$ ----- (1) $E(X) = -2p + 2q + 4(0.3) = 1.2$ $-2p + 2q = 0$ $-p + q = 0$ ----- (2) Solving (1) and (2), $p = q = 0.25$	Very well done. Students are advised to write down the 2 equations clearly before attempting to solve them.
(b)	$E(X^2) = (-2)^2(0.25) + 2^2(0.25) + 4^2(0.3) = 6.8$ $\text{Var}(X) = E(X^2) - (E(X))^2 = 5.36$	Well done except for a few students who computed $E(X^2) =$ $(-2)(0.25)^2 + \dots + 4(0.3)^2$ instead.

- 6 A consumer wants to investigate the average delivery timing by recording 50 timings of food delivery to his house for dinner over 6 months.

(a) In context of the question, give a reason why this method of sampling is not suitable. [1]

A delivery platform's auditor collects the timings from a random sample of 50 food deliveries. The food delivery timings, denoted by x minutes, of the random sample are summarised below.

$$\sum (x - 30) = -66.5, \quad \sum (x - 30)^2 = 586.0763$$

(b) Find the unbiased estimates for the population mean and variance. [2]

The platform's auditor chooses another random sample of 45 timings.

(c) Use your estimates from part (b) to find the probability that in a randomly chosen sample of 45 delivery timings, the mean delivery timing is between 28 minutes and 35 minutes. [3]

6	Solution [6] Sampling	
(a)	It is not suitable as the sampling method introduces bias. This is because the timings recorded (i) are only for dinner timings and not other timings like breakfast or lunch; (ii) are only made by the consumer to his house but not other people in other locations.	Common mistakes include reasons being "6 months duration is too long or too short" or "weather conditions" or "traffic conditions"
(b)	Unbiased estimate of population mean, $\bar{x}$ $= 30 + \frac{\sum (x - 30)}{50}$ $= 30 + \frac{-66.5}{50}$ $= 28.67 \text{ min.}$ Unbiased estimate of population variance, s^2 $= \frac{1}{49} \left[\sum (x - 30)^2 - \frac{(\sum (x - 30))^2}{50} \right]$ $= \frac{1}{49} \left[586.0763 - \frac{(-66.5)^2}{50} \right]$ $= 10.1557 \text{ min}^2$ $= 10.2 \text{ min}^2 \text{ (3 s.f.)}$	Well done. Students are reminded that 28.67 is an exact answer and should be left as it is .
(c)	As $n = 45 \geq 30$ is large, by CLT, $\bar{X} \sim N\left(28.67, \frac{10.1557}{45}\right) \text{ approx.}$ $P(28 \leq \bar{X} \leq 35) = 0.921 \text{ (3 s.f.)}$	Quite a number of students make the mistake of "By CLT, $X \sim N$ and $\bar{X} \sim N$ " Wrong answer of 0.560 is obtained because students keyed in the wrong standard deviation into the GC.

- 7 The following table shows the number of charge cycles, n , undergone by a battery, and its capacity, as a percentage of its original capacity, p %.

n	100	150	200	250	300	350	400	450
p	96	90	85	83	80	79	77	76

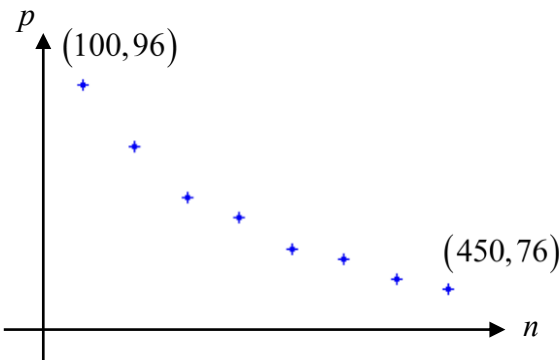
- (a) Sketch a scatter diagram of the data, labelling the axes. Explain why the scatter diagram is consistent with a model of the form $p = a + \frac{b}{n}$, where a and b are constants. [2]
- (b) For the model of the form $p = a + \frac{b}{n}$, calculate the product moment correlation coefficient and the constants a and b . [2]
- (c) Using the model in part (b), estimate the capacity, as a percentage of its original capacity, when the battery undergoes 365 charge cycles. Comment on the reliability of the estimate. [2]

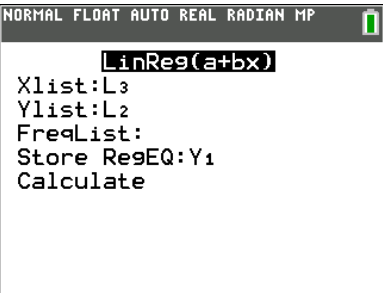
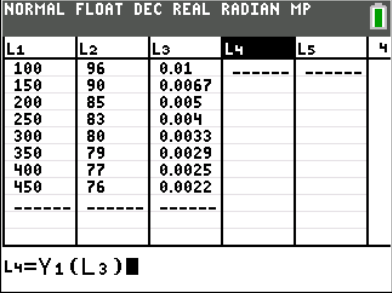
For the model $p = f(n)$, the residual for a point (c, d) is $d - f(c)$.

Let S be the sum of the squares of residuals for the model $p = f(n)$.

For the values of a and b found in part (b),

- (d) (i) find S when $f(n) = a + \frac{b}{n}$, and [1]
- (ii) state the range of S when $f(n) = a + \frac{b}{dn}$, where d is non-zero real constant. [1]

7	Solutions [8] CnR	
(a)	 From the scatter diagram, as n increases, p decreases at a decreasing rate. This is consistent with the model of $p = a + \frac{b}{n}$.	Generally students are able to sketch the scatter diagram but there was a key common error: Scatter's shape looks more linear than curvilinear. There are still a few who <u>did not label the boundary points and axes</u>, swapped the axes erroneously and/or had the incorrect number of points. Regards the explanation, the key phrase underlined is crucial for candidates to obtain the mark. Candidates who attempted to explain in their own words were usually lacking.

(b)	 $r = 0.989$ (3 s.f.) $p = 71.4514 + \frac{2580.3804}{n}$ $a = 71.5, b = 2580$ (3 s.f.)	Generally well done except for some with calculation errors.
(c)	When $n = 365$, $p = 71.4514 + \frac{2580.3804}{365} = 78.5 \text{ (3 s.f.)}.$ When the battery undergoes 365 charge cycles, its capacity, as a percentage of its original capacity is 78.5% (3 s.f.). As $n = 365$ is within the data range of $100 \leq n \leq 450$, the estimate is obtained by an interpolation. Moreover, $r = 0.989$ is close 1, p and $\frac{1}{n}$ has a strong positive linear relationship. Thus, the estimate is reliable.	Generally well done except for some with calculation errors. There were some candidates who used the 3sf values to calculate p and were penalized due to an inaccurate answer obtained. With regards to the explanations, many candidates only mentioned the data range, which is insufficient. (Also even though not penalized here, there were some inaccuracies in the phrasing for this part: p and $\frac{1}{n}$ (not n) have a strong positive linear (some omitted this word) relationship)
(d)(i)	 	

	NORMAL FLOAT DEC REAL RADIAN MP L1 L2 L3 L4 L5 4 100 96 0.01 97.255 ----- 150 90 0.0067 88.654 ----- 200 85 0.005 84.353 ----- 250 83 0.004 81.773 ----- 300 80 0.0033 80.053 ----- 350 79 0.0029 78.824 ----- 400 77 0.0025 77.902 ----- 450 76 0.0022 77.186 ----- ----- ----- ----- ----- ----- L4(1)=97.255219615581 NORMAL FLOAT DEC REAL RADIAN MP L1 L2 L3 L4 L5 5 100 96 0.01 97.255 1.5756 150 90 0.0067 88.654 1.8118 200 85 0.005 84.353 0.4182 250 83 0.004 81.773 1.5057 300 80 0.0033 80.053 0.0028 350 79 0.0029 78.824 0.031 400 77 0.0025 77.902 0.8143 450 76 0.0022 77.186 1.4056 ----- ----- ----- ----- ----- L5(1)=1.5755762833393 NORMAL FLOAT DEC REAL RADIAN MP L1 L2 L3 L4 L5 5 100 96 0.01 97.255 ----- 150 90 0.0067 88.654 ----- 200 85 0.005 84.353 ----- 250 83 0.004 81.773 ----- 300 80 0.0033 80.053 ----- 350 79 0.0029 78.824 ----- 400 77 0.0025 77.902 ----- 450 76 0.0022 77.186 ----- ----- ----- ----- ----- ----- L5=(L2-L4)² 1-Var Stats $\bar{x}$ =0.9456223589 Σx =7.564978871 Σx^2 =10.8470024 Sx =0.7263794003 σx =0.6794657122 n =8 $\min X$ =0.0027755371 Q_1 =0.2245993593																												
	$f(n) = 71.4514 + \frac{2580.3804}{n}.$ <table><tr><td>n</td><td>100</td><td>150</td><td>200</td><td>250</td><td>300</td><td>350</td><td>400</td><td>450</td></tr><tr><td>p</td><td>96</td><td>90</td><td>85</td><td>83</td><td>80</td><td>79</td><td>77</td><td>76</td></tr><tr><td>$f(n)$</td><td>97.3</td><td>88.7</td><td>84.3</td><td>81.8</td><td>80.1</td><td>78.8</td><td>77.9</td><td>77.2</td></tr></table> For $f(n) = g(n)$, $S = (96 - g(100))^2 + (90 - g(150))^2 + \dots + (76 - g(450))^2$$= (96 - 97.255)^2 + (90 - 88.654)^2 + \dots + (76 - 77.186)^2$$= 1.576 + 1.812 + \dots + 1.406$$= 7.56(3 \text{ s.f.})$	n	100	150	200	250	300	350	400	450	p	96	90	85	83	80	79	77	76	$f(n)$	97.3	88.7	84.3	81.8	80.1	78.8	77.9	77.2	Mixed performance. Some were successful some left blank.
n	100	150	200	250	300	350	400	450																					
p	96	90	85	83	80	79	77	76																					
$f(n)$	97.3	88.7	84.3	81.8	80.1	78.8	77.9	77.2																					
(d)(ii)	If $f(n) = g(dn)$, it will still be a model of the form $p = \alpha + \frac{\beta}{n}$ but $f(n) \neq g(n)$ for $d \neq 1$ and $f(n) = g(n)$ for $d = 1$. So, the sum of squares of residuals of $f(n)$ will be more than or equal to that of $g(n)$. $S \geq 7.56(3 \text{ s.f.})$	Poorly done. Most candidates left this blank or got this incorrect. Common errors: $0 < S < \text{xxx}$ $S > \text{xxx}$																											

- 8 White truffles are harvested from the forest. According to harvest records, the distribution of the mass of a white truffle has mean 20 grams and standard deviation 3.5 grams.

It is known that climate change affects the mass of white truffles. To investigate these effects, Lucia collects a random sample of 32 white truffles with a mean mass of 19 grams. She decides to test, at the 5% level of significance, whether the mean mass of white truffles has decreased.

- (a) State hypotheses for Lucia's test, defining any parameter that you use. [2]
- (b) State one assumption Lucia makes to carry out her test. [1]
- (c) Carry out Lucia's test at the 5% level of significance. Show the values you use to carry out the test and give your conclusion in the context of the question. [3]

Jason also wishes to investigate whether the mean mass of white truffles has decreased. He collects another random sample of n white truffles, where $n \geq 30$, with a mean mass of k grams.

Past records revealed that across many harvest expeditions that collected n truffles, 98% of them harvested truffles with a mean mass of more than k grams.

- (d) State the meaning of p -value in context of Jason's test. [1]
- (e) Carry out Jason's test at the 2.5% level of significance. Show the values you use to carry out the test and give your conclusion in the context of the question. [2]

8	Solution [9] Hypothesis testing	
(a)	Let μ be the population mean mass of a white truffle. To test the null hypothesis, $H_0 : \mu = 20$, against the alternative hypothesis, $H_1 : \mu < 20$.	Not well done. Students need to note that μ is a population mean mass , not just the mean mass. Moreover, we are testing to check if the mean mass of the truffles is still 20g. Note that 19g is the sample's mean mass.
(b)	Lucia needs to assume the standard deviation of the mass of the truffles remains as 3.5 grams.	This question proved extremely challenging to students. Many students had the misconception that Central Limit Theorem that had to be assumed . It is a theorem that is always true when the sample size is large enough (i.e. when $n \geq 30$.) This would also mean that there is no need to assume the distributions are normal distribution, because the sample size is large enough, we can take it that the sample mean mass distribution is approximately normal.

(c)	Let X be the mass of a white truffle. Perform a one-tail test at 5% level of significance. Under H_0, as $n = 32 \geq 30$ is large, by CLT, $\bar{X} \sim N\left(20, \frac{3.5^2}{32}\right) \text{ approx.}$ $Z = \frac{\bar{X} - 20}{3.5/\sqrt{32}} \sim N(0,1) \text{ approx.}$ Method 1: $p\text{-value} = 0.05302 > 0.05$ We do not reject H_0. Method 2: Critical region: $z < -1.6449$ $z_{\text{cal}} = \frac{19 - 20}{3/\sqrt{32}} = -1.6162 > -1.6449$ We do not reject H_0. There is insufficient evidence at 5% level of significance to suggest that the mean mass of the truffle has decreased.	This was shockingly not well done. Students are recommended to understand and follow the structure of hypothesis testing. As the distribution of the mass is unknown, one must properly apply the Central Limit Theorem to approximate the mean mass distribution to be normal. This establishes the test statistic's distribution to conduct the test. Students wrote incorrect distributions like $\bar{X} \sim N\left(20, \frac{3.5}{32}\right), \bar{X} \sim N(20, 3.5^2)$ or $Z = \frac{19 - 20}{3.5/\sqrt{32}} \sim N(0,1)$. A strong minority of students had the right numerical calculation but concluded to reject H_0. Students will need to first establish if H_0 is rejected (or not) before writing the conclusion. We numerically use $p\text{-value}$ or z_{crit} make the decision to reject (not reject) H_0. This will then imply if there is sufficient (or insufficient) evidence to suggest if H_1 is reasonable. Writing “do not reject H_0 as there is insufficient evidence” is a logical flaw as we are saying the implication is the other way around. Several students incorrectly used strong terms like accept. As the test is non-assertive, one should reject or do not reject the hypothesis only. Some written conclusions were too assertive, i.e. claiming that we have proved, determined, confirm that mean mass has indeed decreased. Note that a hypothesis test does not assert things to be absolutely true. Instead, the test helps us indicate whether there is sufficient evidence to suggest the hypothesis is reasonable.
-----	---	--

(d)	It is the probability that the sample mean is at least as extreme as the given sample mean of k grams, assuming that the mean mass of the truffle is 20 grams. OR It is the probability that the sample mean is less than the given sample mean of k grams, assuming that the mean mass of the truffle is 20 grams.	This question actually proved challenging for students. Many students defined the level of significance instead of p-value.
(e)	Perform a one-tail test at 2.5% level of significance. Under H_0, as $n \geq 30$ is large, by CLT, $\bar{X} \sim N\left(20, \frac{3.5^2}{n}\right) \text{ approx.}$ $P(\bar{X} > k) = 0.98$ (given) $1 - P(\bar{X} \leq k) = 0.98$ $P(\bar{X} \leq k) = 0.02$ $p\text{-value} = P(\bar{X} \leq k) = 0.02 < 0.025$ We reject H_0. There is sufficient evidence at 2.5% level of significance to conclude that the mean mass of the truffle has decreased.	This question proved challenge for students. Many students did not appreciate the need for the test statistic distribution, i.e. $\bar{X} \sim N\left(20, \frac{3.5^2}{n}\right)$ approx. due to CLT. Only with this can they use the information provided in the question to deduce $P(\bar{X} > k) = 0.98$. Students should poor understanding of p-value, where many were unable to recognise that $P(\bar{X} \leq k) = 0.02$ was indeed the p-value and could use it directly to make the decision to reject/not reject H_0.

- 9 (a) Events A and B are such that $P(B' | A') = \frac{4}{5}$ and $P(A) + 5P(B) = \frac{5}{2}$. Find $P(A \cap B)$. [4]
- (b) A school organises a camp for 6 seniors and 5 juniors.
- (i) For the opening ceremony, the 11 students are arranged in a straight line for a photograph. Find the number of possible arrangements if the students standing at each end of the line are not both seniors. [2]
- (ii) For a discussion session, the 11 students are seated at a round table. Find the number of possible arrangements if no 2 juniors are seated together. [2]
- (iii) For an activity, the 11 students are grouped into a group of 5 and a group of 6. Find the number of ways to group the students if each group has at least 2 juniors. [2]

9	Solutions [10] PnC + Prob	
(a)	$P(B' A') = \frac{4}{5}$ $\frac{P(A' \cap B')}{P(A')} = \frac{4}{5}$ $\frac{1 - P(A \cup B)}{1 - P(A)} = \frac{4}{5}$ $5 - 5P(A) - 5P(B) + 5P(A \cap B) = 4 - 4P(A)$ $5P(A \cap B) = P(A) + 5P(B) - 1$ $5P(A \cap B) = \frac{5}{2} - 1 \quad (\text{given})$ $P(A \cap B) = \frac{3}{10}.$	This was generally well-attempted. Students knew the conditional probability definition and applied it well. Students who saw most success managed used the facts: $P(A' \cap B') = 1 - P(A \cup B)$ and $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ to their advantage. A common mistake was that $P(A' \cap B') = 1 - P(A \cap B)$.
(bi)	No. of ways without restrictions – No. of ways with seniors at the end of each row $= 11! - {}^6C_2 \times 2! \times 9!$ $= 29030400$	Many interpreted that both ends were juniors only. But it was also possible to have one Junior and one Senior. A general tip is to realise that complement method might be in play if the word 'not' is used.
(ii)	No. of ways to arrange the seniors about the circle $= (6-1)!$ No. of ways to slot in the juniors $= {}^6C_5 \times 5!$ Thus, $\text{No. of ways required} = (6-1)! \times {}^6C_5 \times 5! = 86400$	This saw more success than (bi). Common mistakes were: - Not arranging the Juniors. - Arranging the Juniors in a circle again after the seniors were already arranged in a circle. (Arranging people in a circle once orientates the positions, this means the rotation symmetry is lost.

		So, there is no need to do $(n-1)!$ again in this circle arrangement anymore.)
(iii)	<u>Case 1:</u> 3 Juniors in group of 5 $\binom{5}{3}\binom{6}{2} = 150$ <u>Case 2:</u> 2 Juniors in group of 5 $\binom{5}{2}\binom{6}{3} = 200$ Total number of ways $150 + 200 = 350$.	This saw more success than (bi) as well. Common mistakes: - Incurring order (i.e. $\times n!$) to the chosen participants when there was no need to do so. - Multiplying $2!$ to consider select one group then the next. - Threating both scenarios as the same and just add the same amount together, i.e. $200 + 200$. - Did not consider all cases when using a complement method.

- 10** A robot is programmed to move along the x -axis on the x - y plane. It starts at the origin $O(0,0)$.

The robot receives a signal to either:

- move one unit to the right with a probability of 0.3; or
- move one unit to the left, with a probability of 0.7.

The signals are independent of each other.

The robot receives 10 signals. The random variable R denotes the number of times the robot moves to the right.

- (a) Find the probability the robot moves exactly 3 times to the right. [1]
 (b) Find the probability the robot moves more than 3 but less than 8 times to the right. [2]

The random variable X denotes the x -coordinate of the final position of the robot after the 10 signals.

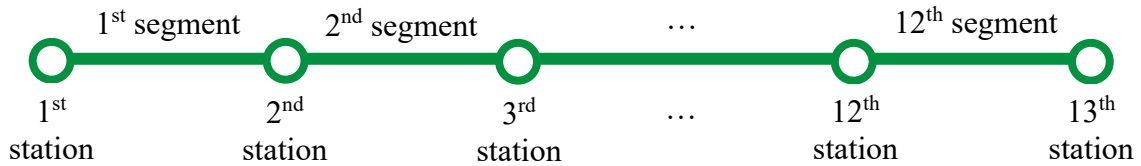
For example, if the robot moves exactly 8 times to the right (and so, 2 times to the left) it will end up at the final position with x -coordinate 6.

- (c) (i) If the robot moves exactly r times to the right, show that x -coordinate of the final position of the robot is $2r - 10$. [1]
 (ii) Find $P(X = -4)$. [1]
 (iii) Find the mean and variance of X . [3]
- (d) In a separate experiment involving n independent signals, the probability that the robot moves to the right more than 4 times exceeds 0.2. Find the least possible value of n . [2]

10	Solution [10] Binomial Distribution, DRV	
(a)	Let R be the number of right moves in 10 steps. $R \sim B(10, 0.3)$ $P(R = 3) = 0.267$ (3 s.f.)	Generally well done.
(b)	$P(3 < R < 8)$ $= P(R \leq 7) - P(R \leq 3)$ $= 0.349$ (3 s.f.)	Many candidates made mistakes for this part: e.g. $P(R \leq 8) - P(R \leq 3)$ $P(R \leq 8) - P(R \leq 2)$ $P(R \leq 7) - P(R \leq 3)$ etc.
(c)(i)	$x = \text{Number of right moves} - \text{number of left moves}$ $= r - (10 - r)$ $= 2r - 10.$ (Shown)	Generally well done.
(c)(ii)	$X = -4 \Rightarrow 2R - 10 = -4 \Rightarrow R = 3.$ $P(X = -4) = P(R = 3) = 0.267$ (3 s.f.)	Generally students are able to identify that the answer is the same as (a) which was good.

(c)(iii)	$E(R) = (10)(0.3) = 3$ $\text{Var}(R) = (10)(0.3)(0.7) = 2.1$ $E(X) = E(2R - 10)$ $= 2E(R) - 10$ $= 2(3) - 10$ $= -4$ $\text{Var}(X) = \text{Var}(2R - 10)$ $= 2^2 \text{Var}(R)$ $= 4(2.1)$ $= 8.4$	Most students were not able to obtain the right answers for this part. Many found $E(R)$ and $\text{Var}(R)$ and concluded that $E(X)$ and $\text{Var}(X)$ had the same answers. Some applied the wrong properties for $\text{Var}(2R - 10)$ e.g. $2\text{Var}(R) - 10$. There were others who tried setting up a DRV table to find $E(X)$ and $\text{Var}(X)$ from scratch, it would prove to be rather tedious to do so and also, mostly students will not be able to obtain the exact answers. There were a few candidates who tried using $X = R - L$ to find $E(X)$ and $\text{Var}(X)$. However, this is an incorrect method because R and L are not independent variables and we should not be applying properties of expectation and variance to this expression.
(d)	Let Y be the number of right moves in n steps. $Y \sim B(n, 0.3)$ $P(Y > 4) > 0.2$ $1 - P(Y \leq 4) > 0.2$ $P(Y \leq 4) < 0.8$ By GC, When $n = 10$, $P(Y \leq 4) = 0.8497$ When $n = 11$, $P(Y \leq 4) = 0.7897$ Thus, least $n = 11$.	Poorly done. A number of candidates did not define a new variable nor state the distribution $B(n, 0.3)$ at all. Also, another very key error made was $1 - P(Y \leq 3)$ resulting in an incorrect result.

- 11 Mr Lim travels to work by train every weekday morning. His journey begins at the 1st station and ends at 13th station. The train passes through 12 segments, where each segment is the distance between two consecutive stations. The following diagram illustrates his journey.



Mr Lim leaves the 1st station at 8.25 am every weekday. If Mr Lim arrives at the final station at or before 9 am, he is considered to be early. Otherwise, he is considered to be late.

On fine days, the travel time for each train segment is independent and normally distributed with mean 2.5 minutes and standard deviation 0.8 minutes.

- (a) Show that, on a randomly chosen fine day, the probability that he arrives early for work is 0.9644, correct to 4 decimal places. [2]
 (b) Over 75 fine days, find the expected number of times Mr Lim arrives early for work. [1]

On rainy days, the train takes 20% longer to travel across each train segment.

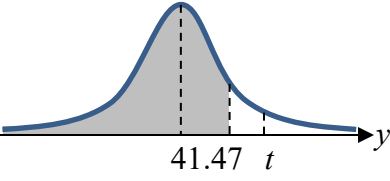
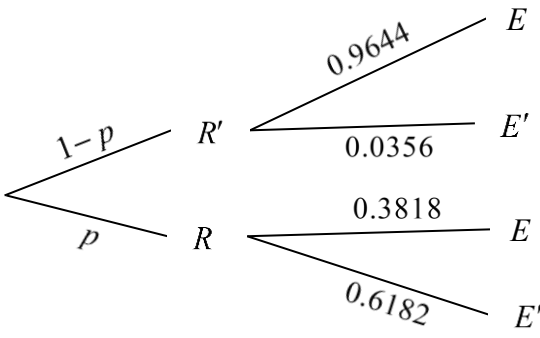
- (c) Show that, on a randomly chosen rainy day, the probability that he is late for work is 0.6182, correct to 4 decimal places. [3]
 (d) On rainy days, Mr Lim wishes to change his departure time so that he is early for work on at least 95% of the rainy days. Find the latest time, to the nearest minute, that he should leave the first station on a rainy day. [3]

It is known that there is a probability of p that it rains on a weekday morning.

Regardless of the weather, it is also known that when Mr Lim leaves the first station at 8.25 am, he is early for work 80% of the time.

- (e) One day, Mr Lim is late for work. Find the probability it was a rainy day. [4]

11	Solutions [13] Normal Distribution, Conditional Probability	
(a)	Let T be the travel time of a train segment. $T \sim N(2.5, 0.8^2)$ Let X be total travel time across 12 train segments on a fine day. $X = T_1 + T_2 + \dots + T_{12} \sim N(12(2.5), 12(0.8^2)) = N(30, 7.68)$ $P(X \leq 35) = 0.9644016 = 0.9644$ (4 d.p.) (Shown)	Generally well done
(b)	Expected number of times Mr Lim is early over 250 fine days $= 75 \times 0.9644$ $= 72.3$ (3 s.f.)	While many students were able to do this part, many left an incorrect response by rounding their answer to the nearest integer 72 instead of to the nearest 3sf . (Specifically mentioned in A Level 2024 Examiners Report).

(c)	Let Y be the total travel time across 12 train segments on a rainy day. Let $Y = 1.2X \sim N(1.2(30), 1.2^2(7.68)) = N(36, 11.0592)$ $P(Y > 35) = 0.6181795 = 0.6182$ (4 d.p.) (Shown)	Most students were able to obtain this part.
(d)	Let t be the time taken to travel across 12 train segments for Mr Lim to be early for work. $P(Y \leq t) \geq 0.95$ By GC, $t \geq 41.47$ $t \geq 42$ (to nearest min) Thus, he should leave at 8.18 am. 	Mixed performance. Some students did not use inequalities and also did not conclude with the inequality in mind. There were also a couple of candidates who stated $P(Y \leq 35) \geq 0.95$ and considered the unknown t to be an unknown mean which is incorrect interpretation of the question.
(e)	Let R be the event that it rains on a weekday. Let E be the event that Mr Lim is early for work. Given the following: $P(R) = p$ $P(E) = 0.8$ From (a), $P(E R') = 0.9644$ From (b)(i), $P(E' R) = 0.6182$ Raining or Fine Early or Late  $P(E \cap R) + P(E \cap R') = P(E)$ $P(R)P(E R) + P(R')P(E R') = P(E)$ $p(0.3818) + (1-p)(0.9644) = 0.8$ $0.5826p = 0.1644$ $p = \frac{0.1644}{0.5826}$ $= 0.28218$ $= 0.2822$ (4 d.p.)	Poorly done. Most left this blank. For those who attempted, most misinterpreted the question and stopped at finding p. Not many realized that this was a conditional probability question.

	$P(R E')$ $= \frac{P(R \cap E')}{P(E')}$ $= \frac{0.28218 \times 0.6182}{1 - 0.8}$ $= 0.872 \text{ (3 s.f.)}$	
--	---	--