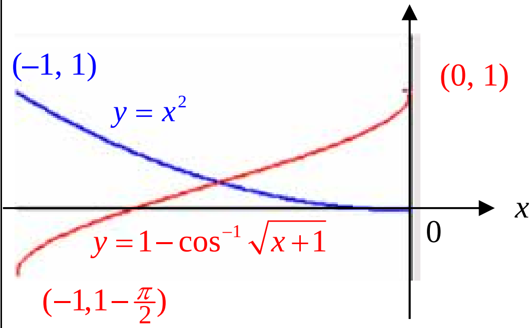
Paper 2 Solutions

1(i)	$\text{LHS} = u_{4n+4}$ $= u_{4n+3} + 2u_{4n+2}$ $= \left(u_{4n+2} + 2u_{4n+1}\right) + 2\left[u_{4n+1} + 2u_{4n}\right]$ $= u_{4n+2} + 2u_{4n+1} + 2u_{4n+1} + 4u_{4n}$ $= \left(u_{4n+1} + 2u_{4n}\right) + 4u_{4n+1} + 4u_{4n}$ $= 5u_{4n+1} + 6u_{4n}$ =RHS (Shown)
(ii)	Let P_n be the statement $u_{4n} = 5q$, $n \in \mathbb{Z}^+$ for some $q \in \mathbb{Z}$. When n = 1, $u_4 = u_3 + 2u_2$ $= u_2 + 2u_1 + 2u_2$ $= 3(1) + 2$ $= 5$ $= 5 \times 1$ Hence u_4 is divisible by 5. $\therefore$ P_1 is true. Assume that P_k is true for some $k \in \mathbb{Z}^+$, i.e. $u_{4k} = 5m$ for some $m \in \mathbb{Z}$. To show that P_{k+1} is also true, i.e. $u_{4k+4} = 5t$ for some $t \in \mathbb{Z}$. $\text{LHS} = u_{4k+4}$ $= 5u_{4k+1} + 6u_{4k}$ $= 5u_{4k+1} + 6\left(5m\right)$ $= 5\left(u_{4k+1} + 6\right)$ $= 5t,$ where $t \in \mathbb{Z}$ Hence u_{4k+4} is divisible by 5. $\therefore \text{P}_k \text{ is true} \Rightarrow \text{P}_{k+1} \text{ is true}$ Since P_1 is true and P_k is true $\Rightarrow \text{P}_{k+1}$ is true, by Mathematical induction, u_{4n} is divisible by 5 for all $n \in \mathbb{Z}^+$.

2(i)	(Closure) $(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2, y_1 \times y_2) \in V$ $(\because y_1 \times y_2 \neq 0)$ (Associativity) $\left[(x_1, y_1) \oplus (x_2, y_2)\right] \oplus (x_3, y_3) = (x_1 + x_2, y_1 y_2) \oplus (x_3, y_3)$ $= ((x_1 + x_2) + x_3, (y_1 y_2) y_3)$ $= (x_1 + (x_2 + x_3), y_1 (y_2 y_3))$ $= (x_1, y_1) \oplus \left[(x_2, y_2) \oplus (x_3, y_3)\right]$ (Identity) $(x_1, y_1) \oplus (0,1) = (x_1 + 0, y_1 \times 1)$ $= (x_1, y_1)$ (Inverse) $(x_1, y_1) \oplus (-x_1, \frac{1}{y_1}) = (x_1 + (-x_1), y_1 \times \frac{1}{y_1})$ $= (0,1)$ (Commutativity) $(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2, y_1 \times y_2)$ $= (x_2 + x_1, y_2 \times y_1)$ $= (x_2, y_2) \oplus (x_1, y_1)$	**B1** **B1** **B1** **B1** **B1**
(ii)	$0 \in \mathbb{R}$ but $0(x, y) = (0, 0) \notin P$, violating the closure axiom.	**B1**
(iii)	$k \otimes (x, y) = (kx, y)$	**B1** (accept any other definition that satisfies the axioms)

<table>
<tr><td>3(i)</td><td>
</td><td>B1 correct shapes and domain

B1 coordinates of end points. Allow non-exact answer $(-1, -0.571)$</td></tr>
<tr><td>(ii)(a)</td><td>This procedure only yields non-negative values, whereas the required root is negative.</td><td>B1</td></tr>
<tr><td>(ii)(b)</td><td>Using $\mathrm{f}(x) = -\sqrt{1-\cos^{-1}\sqrt{x+1}}$ and $x_0 = 0$,

$x_1 = -1$

$\mathrm{f}(-1) = -\sqrt{1-\cos^{-1} 0} = -\sqrt{1-\frac{\pi}{2}}$ which is undefined as

$1-\frac{\pi}{2} < 0$.</td><td>B1

B1</td></tr>
<tr><td>(ii)(c)
(i)</td><td>$$x^2 = 1-\cos^{-1}\sqrt{x+1}$$
$$\cos^{-1}\sqrt{x+1} = 1-x^2$$
$$\sqrt{x+1} = \cos(1-x^2)$$
$$x = \cos^2(1-x^2) - 1$$
$$\therefore \mathrm{f}(x) = \cos^2(1-x^2) - 1$$</td><td>M1 making the other x the subject

A1</td></tr>
<tr><td>(ii)(c)
(ii)</td><td>$x_0 = 0$ $x_0 = -0.5$
$x_1 = -0.7081$ $x_1 = -0.4646$
$x_2 = -0.2287$ $x_2 = -0.4987$
$x_3 = -0.6595$ $x_3 = -0.4659$
$x_4 = -0.2868$ $x_4 = -0.4975$

Using $x_0 = -0.5$ results in a faster convergence to the root of (*) than using $x_0 = 0$.</td><td>B1 1st sequence

B1 2nd sequence

B1</td></tr>
</table>

4(a)(i)	By the directrix property: $\frac{OP}{Pd}=e$, where Pd is the distance from P to the directrix $x=p$. $OP=r$, $Pd=$ dist. from O to directrix $-$ horizontal component of OP $= p-r\cos\theta$ (This statement is true even if θ is not acute, although students do not have to justify it.) Hence, $\frac{r}{p-r\cos\theta}=e$ $r=ep-re\cos\theta$ $r\left(1+e\cos\theta\right)=ep$ $r=\frac{ep}{1+e\cos\theta}$ (shown)	**M1** for use of directrix property **M1** for finding distance Pd **MA1** for logical working to given answer
(a)(ii)	Area swept by OP is $=\frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{2\pi}{3}}\left(\frac{\frac{1}{4}p}{1+\frac{1}{4}\cos\theta}\right)^2\,\mathrm{d}\theta$ $=\frac{p^2}{32}\left(2.9712\right)$ $=0.092849p^2=0.0928p^2$ (3sf)	**M1** for correct use of formula **A1** for answer
(a) (iii)	Area of ellipse is $=\frac{1}{2}\int_{-\pi}^{\pi}\left(\frac{\frac{1}{4}p}{1+\frac{1}{4}\cos\theta}\right)^2\,\mathrm{d}\theta$ $=\frac{p^2}{32}\left(6.9219\right)$ $=0.21631p^2$ Let T_1 be time taken for P to travel from P_1 to P_2. By Kepler's Second Law, $\frac{0.092849p^2}{0.21631p^2}=\frac{T_1}{T}$ $T_1=0.429T$	**M1** for area of ellipse **A1** for answer

(b)(i)	When $u=\tan\frac{\theta}{2}$, $\quad \frac{\mathrm{d}u}{\mathrm{d}\theta}=\frac{1}{2}\sec^2\frac{\theta}{2}$ $\int\frac{1}{(1+\cos\theta)^2}\,\mathrm{d}\theta=\int\frac{1}{(1+\cos\theta)^2}\frac{\mathrm{d}\theta}{\mathrm{d}u}\mathrm{d}u$ $=\int\frac{1}{\left(2\cos^2\frac{\theta}{2}\right)^2}\cdot 2\cos^2\frac{\theta}{2}\mathrm{d}u$ $=\int\frac{1}{2\cos^2\frac{\theta}{2}}\mathrm{d}u$ $=\frac{1}{2}\int\sec^2\frac{\theta}{2}\mathrm{d}u$ $=\frac{1}{2}\int 1+\tan^2\frac{\theta}{2}\mathrm{d}u$ $=\frac{1}{2}\int 1+u^2\ \mathrm{d}u$ $=\frac{1}{2}\left(u+\frac{1}{3}u^3\right)+C$ where C is an arbitrary constant $=\frac{1}{2}\left(\tan\frac{\theta}{2}+\frac{1}{3}\tan^3\frac{\theta}{2}\right)+C$	**M1** for correct change of $\mathrm{d}\theta$ term to $\mathrm{d}u$ term **M1** for use of double angle formula to simplify integrand **M1** for use of Pythagorean identity to simplify integrand **M1** for simplifying integrand to quadratic polynomial **MA1** for working leading to given answer
	<u>Alternative solution</u> When $u=\tan\frac{\theta}{2}$, $\frac{\mathrm{d}u}{\mathrm{d}\theta}=\frac{1}{2}\sec^2\frac{\theta}{2}$ $=\frac{1}{2}\left(1+\tan^2\frac{\theta}{2}\right)=\frac{1}{2}\left(1+u^2\right)$ $\tan\theta=\frac{2\tan\frac{\theta}{2}}{1-\tan^2\frac{\theta}{2}}=\frac{2u}{1-u^2}$ By triangle method, $\cos\theta=\frac{1-u^2}{\sqrt{(2u)^2+\left(1-u^2\right)^2}}=\frac{1-u^2}{\sqrt{\left(1+u^2\right)^2}}=\frac{1-u^2}{1+u^2}$	**M1** for any evidence of use of triangle method (or other similar methods to obtain relevant trigo expressions) **M1** for correct change of $\mathrm{d}\theta$ term to $\mathrm{d}u$ term **M1** for changing $\cos\theta$ to $\frac{1-u^2}{1+u^2}$ **M1** for simplifying integrand to quadratic polynomial **MA1** for working leading to given answer

	Hence, $\int \frac{1}{(1+\cos\theta)^2}\,\mathrm{d}\theta = \int \frac{1}{(1+\cos\theta)^2}\frac{\mathrm{d}\theta}{\mathrm{d}u}\,\mathrm{d}u$ $= \int \frac{1}{\left(1+\frac{1-u^2}{1+u^2}\right)^2}\frac{2}{1+u^2}\,\mathrm{d}u$ $= \int \frac{1}{\left(\frac{1+u^2+1-u^2}{1+u^2}\right)^2}\frac{2}{1+u^2}\,\mathrm{d}u$ $= \int \frac{1+u^2}{2}\,\mathrm{d}u$ $= \frac{1}{2}\left(u+\frac{1}{3}u^3\right)+C$ where C is an arbitrary constant $= \frac{1}{2}\left(\tan\frac{\theta}{2}+\frac{1}{3}\tan^3\frac{\theta}{2}\right)+C$	
(b)(ii)	Area is $\int_{-\frac{\pi}{2}}^{\frac{2\pi}{3}} \frac{1}{2}\cdot\frac{p^2}{(1+\cos\theta)^2}\,\mathrm{d}\theta$ $= \frac{1}{4}p^2\left[\tan\frac{\theta}{2}+\frac{1}{3}\tan^3\frac{\theta}{2}\right]_{-\frac{\pi}{2}}^{\frac{2\pi}{3}}$ $= \frac{1}{4}p^2\left[\left(\sqrt{3}+\frac{1}{3}\left(3\sqrt{3}\right)\right)-\left(-1-\frac{1}{3}\right)\right]$ $= \frac{1}{4}p^2\left(2\sqrt{3}+\frac{4}{3}\right)$ $= \frac{p^2}{6}\left(3\sqrt{3}+2\right)$	**M1** for correct application of formula in part (i) **A1** for final answer in exact form

5(i)	$\frac{\mathrm{d}P}{\mathrm{d}t}=k\left[I_0-I(t)\right]$ $\frac{\mathrm{d}^2P}{\mathrm{d}t^2}=-k\frac{\mathrm{d}I}{\mathrm{d}t}$ where $I=I(t)$ $=-k\left[R(t)-S(t)\right]$ $=-k\left[(250-5P)-\left(500-40P-10\frac{\mathrm{d}P}{\mathrm{d}t}\right)\right]$ $=-k\left(-250+35P+10\frac{\mathrm{d}P}{\mathrm{d}t}\right)$ $\frac{\mathrm{d}^2P}{\mathrm{d}t^2}+10k\frac{\mathrm{d}P}{\mathrm{d}t}+35kP=250k$ (shown!)	**M1** **A1** AG, so detailed working needed
(ii)	$k=1\Rightarrow\frac{\mathrm{d}^2P}{\mathrm{d}t^2}+10\frac{\mathrm{d}P}{\mathrm{d}t}+35P=250$ ---(1) This D.E. has characteristic equation $m^2+10m+35=0$ $(m+5)^2+10=0$ $m=-5\pm\sqrt{10}\mathrm{i}$ The complementary function is $P_c=\mathrm{e}^{-5t}\left(A\cos\sqrt{10}t+B\sin\sqrt{10}t\right)$, where A and B are arbitrary constants By observation, the particular integral is a constant D. Substitute $P=D$ into (1): $35D=250$ $D=\frac{50}{7}$ The G.S. is $P=\mathrm{e}^{-5t}\left(A\cos\sqrt{10}t+B\sin\sqrt{10}t\right)+\frac{50}{7}$ As $t\to\infty$, $\mathrm{e}^{-5t}\to 0$ (Note that $\left\|A\cos\sqrt{10}t+B\sin\sqrt{10}t\right\|\le\sqrt{A^2+B^2}$, a finite value) The price will stabilise at \$7.14 in the long run.	**M1** **A1** **A1** **B1** **A1** (accept an incorrect GS and/or PS here) **B1** accept $\$\frac{50}{7}$
(iii)	When t = 0, P = 5.50: $5.5=\mathrm{e}^0\left(A\cos 0+B\sin 0\right)+\frac{50}{7}$ $=A+\frac{50}{7}$ $A=-\frac{23}{14}$ When t = 0.5, P = 6.50:	**M1** substitute into G.S. **A1**

	$6.5 = \mathrm{e}^{-2.5}\left(-\frac{23}{14}\cos\frac{\sqrt{10}}{2} + B\sin\frac{\sqrt{10}}{2}\right) + \frac{50}{7}$	
	$B = -7.8490$	**A1**
	In 10 years' time, $t = 1$:	
	$P = \mathrm{e}^{-5}\left(A\cos\sqrt{10} + B\sin\sqrt{10}\right) + \frac{50}{7}$	
	$= 7.155 = 7.16$ (3 s.f.)	
	The price of the goods will be \$7.16 in 10 years' time.	**A1**

6(i)	H_0: The two factors are independent of each other H_1: The two factors are not independent of each other Under H_0, i.e. if the factors are independent of each other, the expected frequency of observations where both factors A and B are present is $(\text{grand total})\times \mathrm{P}(\text{A present } \cap \text{B present})$ $=(\text{grand total})\times \mathrm{P}(\text{A present})\times \mathrm{P}(\text{B present})$ (since A and B are independent factors) $=(\text{grand total})\times \dfrac{\text{row total}}{\text{grand total}}\times\dfrac{\text{column total}}{\text{grand total}}$ $=\dfrac{\text{row total}\times\text{column total}}{\text{grand total}}$ $=\dfrac{(2a)\times(a+b)}{2(a+b)}=a$	**B1** for use of formula $\dfrac{\text{row total}\times\text{column total}}{\text{grand total}}$ to obtain answer **B1** for explanation using independence as to why the formula is true
(ii)	Under H_0, Expected freq | Factor A present | Factor A absent Factor B present | a | a Factor B absent | b | b We assume that $a, b \geq 5$. Test statistic is $\dfrac{((a+x)-a)^2}{a}+\dfrac{((a-x)-a)^2}{a}$ $+\dfrac{((b+x)-b)^2}{b}+\dfrac{((b-x)-b)^2}{b}$ $=\dfrac{x^2}{a}+\dfrac{x^2}{a}+\dfrac{x^2}{b}+\dfrac{x^2}{b}$ $=2x^2\left(\dfrac{1}{a}+\dfrac{1}{b}\right)$ To conclude at 5% level of significance that factors are not independent of each other, $2x^2\left(\dfrac{1}{a}+\dfrac{1}{b}\right)\geq 3.841$.	**M1** for all expected frequencies **B1** for assumption **M1** for formula to calculate contributions to test statistic **A1** for answer **B1** for inequality (based on previous answer)

7(i)	Let X be the number of calls that the call center receives in t minutes. $X \sim \text{Po}(3t)$ $\text{P}(X \geq 1) = 1 - \text{P}(X = 0) = 1 - \text{e}^{-3t}$	**M1** for correct Poisson variable defined and used **A1** for answer
(ii)	If time between the two consecutive calls is less than t minutes, then the number of calls received in t minutes is at least 1. $\text{P}(T < t) = \text{P}(X \geq 1) = 1 - \text{e}^{-3t}$ Cdf of T, $\text{F}(t) = 1 - \text{e}^{-3t}, t \geq 0$ By differentiation, pdf of T, $\text{f}(t) = 3\text{e}^{-3t}, t \geq 0$ This is the pdf of an exponential distribution with parameter 3. (shown)	**M1** for identifying $\text{P}(T < t) = \text{P}(X \geq 1)$ **MA1** for correct pdf of T through differentiation
(iii)	Let u be the upper quartile of T, l be the lower quartile of T. $\text{F}(u) = 0.75$ $1 - \text{e}^{-3u} = 0.75$ $u = -\frac{1}{3}\ln 0.25$ $= \frac{1}{3}\ln 4$ $\text{F}(l) = 0.25$ $1 - \text{e}^{-3l} = 0.25$ $l = -\frac{1}{3}\ln 0.75$ $= \frac{1}{3}\ln\frac{4}{3}$ Interquartile range $= u - l$ $= \frac{1}{3}\ln 4 - \frac{1}{3}\ln\frac{4}{3}$ $= \frac{1}{3}\ln\left(\frac{4}{\frac{4}{3}}\right)$ $= \frac{1}{3}\ln 3$	**M1** for method of find either the upper quartile or the lower quartile **M1** for taking the difference of the upper and lower quartiles **A1** for answer in a single term

<table>
<tr>
<td>8(i)</td>
<td>The Wilcoxon matched pair signed rank test takes the magnitudes of the differences into consideration, as opposed to the sign test where the magnitudes are not considered at all.</td>
<td>B1 for answer</td>
</tr>
<tr>
<td>(ii)</td>
<td>Let X and Y be the students results for BT1 and BT2 respectively, Let $D = Y - X$ and m be the median of d
$H_0:\ m = 0$
$H_1:\ m > 0$

Under H_0,

<table>
<tr><td>D</td><td>-3</td><td>-9</td><td>8.5</td><td>11</td><td>2</td><td>4</td><td>24</td><td>20.5</td><td>5.5</td><td>9.5</td><td>0</td><td>-1.5</td></tr>
<tr><td>Rank</td><td>3</td><td>7</td><td>6</td><td>9</td><td>2</td><td>4</td><td>11</td><td>10</td><td>5</td><td>8</td><td>-</td><td>1</td></tr>
<tr><td>+</td><td></td><td></td><td>6</td><td>9</td><td>2</td><td>4</td><td>11</td><td>10</td><td>5</td><td>8</td><td></td><td></td></tr>
<tr><td>-</td><td>3</td><td>7</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>1</td></tr>
</table>
From the table: $T_+ = 55$, $T_- = 3 + 7 + 1 = 11$.
Test statistic: $T = \min\{T_+, T_-\} = 11$

Discarding the entry (K) with zero difference, $n = 11$.
From table, $T_{crit} = 13$, so reject H_0 if $T \le 13$.

Since $T = 11 \le 13$, we reject H_0. There is sufficient evidence at 5% level of significance that the students performed better, on average, for BT2 as compared to BT1.</td>
<td>B1 for correct hypothesis with properly defined symbols

M1 for ranks of absolute differences

M1 for T_+, T_-, and final T statistic

M1 for T_{crit} from correct position in table and showing test statistic lies in rejection region

A1 for final statement</td>
</tr>
<tr>
<td>(iii)</td>
<td>A more appropriate test to perform is the paired sample t test.</td>
<td>B1 for test</td>
</tr>
</table>

9(i)	Sample mean $\bar{x} = \frac{\Sigma x}{n_x} = \frac{21.4}{10} = 2.14$ Unbiased estimate of population variance, s^2 $= \frac{1}{n_x - 1}\left(\Sigma x^2 - \frac{(\Sigma x)^2}{n_x}\right)$ $= \frac{1}{10-1}\left(46.3 - \frac{(21.4)^2}{10}\right)$ $= 0.056$ 95% confidence interval for μ_x is $\left(\bar{x} \pm t_{0.025}^{(10-1)}\sqrt{\frac{s^2}{n_x}}\right)$ $= \left(2.14 \pm 2.2622\sqrt{\frac{0.056}{10}}\right)$ $= (1.97, 2.31)$ (3sf)	**M1** for calculating s^2 **M1** for using correct formula for confidence interval **A1** for answer
(ii)	There is a 5% chance that the 95% confidence interval for μ from a random sample does not actually contain μ .	**B1** for answer
(iii)	There is sufficient evidence at 5% level of significance that the mean mass of the batch of chickens raised with Feed A differ from 2.35kg. Since $\mu_0 = 2.35$ does not lie within $(1.97, 2.31)$, the 95% confidence interval of μ , this implies that $\bar{x} = 2.14$ does not lie within the acceptance region of the two tail hypothesis test at 5% level of significance with $\mu_0 = 2.35$. Hence, we reject the null hypothesis and arrive at the conclusion as stated.	**B1** for correct conclusion **B1** for reasonable explanation (as long as they state the necessary implication)
(iv)	Let μ_x be the population mean mass of the batch of chickens raised with Feed A Let μ_y be the population mean mass of the batch of chickens raised with Feed B $H_0 : \mu_y - \mu_x = k$ $H_1 : \mu_y - \mu_x > k$	**B1** for defining the two different μ **B1** for both hypotheses correct

(v)	Pooled estimate of the population variance is $$\frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n_x} + \Sigma y^2 - \frac{(\Sigma y)^2}{n_y}}{n_x + n_y - 2}$$ $$= \frac{46.3 - \frac{(21.4)^2}{10} + 81.72 - \frac{(34.8)^2}{15}}{10+15-2}$$ $= 0.064696$ (5sf) (Also accept if formulas like $\frac{n_x s_x^2 + n_y s_y^2}{n_x + n_y - 2}$ are used)	**M1** for correct formula **A1** for correct answer
(vi)	Sample mean $\bar{y} = \frac{\Sigma y}{n_y} = \frac{34.8}{15} = 2.32$ Under H_0, $$T \sim \frac{(\bar{Y} - \bar{X}) - (\mu_y - \mu_x)}{s_{pooled}\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \sim t(n_x + n_y - 2) = t(23)$$ When Xin's belief is supported at 10% level of significance, reject H_0, $$t > t_{crit} = 1.3195$$ $$\frac{(2.32 - 2.14) - (k)}{\sqrt{0.064696}\sqrt{\frac{1}{10} + \frac{1}{15}}} > 1.3195$$ $$k < 0.0430 \text{ (3sf)}$$ Hence. $0 < k < 0.0430$.	**M1** for correct test statistic **M1** for use of correct rejection region **M1** for substituting correct numbers into the expression **A1** for answer

10(i)	$\mathrm{E}(X)=\sum_{r=1}^{\infty} rq^{r-1}p$ $=p\sum_{r=1}^{\infty} rq^{r-1}$ $=p\left(\dfrac{1}{(1-q)^2}\right)=\dfrac{p}{p^2}=\dfrac{1}{p}$	**M1** for expectation formula **MA1** for use of given summation to lead to given answer
(ii)	$\mathrm{E}(X(X-1))=\sum_{r=1}^{\infty} r(r-1)q^{r-1}p$ $=pq\sum_{r=1}^{\infty} r(r-1)q^{r-2}$ $=pq\left(\dfrac{2}{(1-q)^3}\right)$ $=\dfrac{2pq}{p^3}=\dfrac{2q}{p^2}$ $\mathrm{E}(X^2)-\mathrm{E}(X)=\dfrac{2q}{p^2}$ $\mathrm{E}(X^2)=\dfrac{2q}{p^2}+\dfrac{1}{p}$ $\mathrm{Var}(X)=\mathrm{E}(X^2)-(\mathrm{E}(X))^2$ $=\dfrac{2q}{p^2}+\dfrac{1}{p}-\dfrac{1}{p^2}$ $=\dfrac{2q+p-1}{p^2}$ $=\dfrac{2q-q}{p^2}=\dfrac{q}{p^2}$	**M1** for expectation formula for $\mathrm{E}(X(X-1))$ **M1** for use of given summation formula **M1** for use of variance formula **A1** AG
(iii)	Let X be "number of attempts a student take to solve a mathematical puzzle" $\mathrm{H}_0: X \sim \mathrm{Geo}(0.4)$ $\mathrm{H}_1: X$ does not follow $\mathrm{Geo}(0.4)$ Under H_0, x: 1, 2, 3, 4, 5, ≥ 6 O_i: 49, 19, 6, 3, 5, 18 E_i: 40, 24, 14.4, 8.64, 5.184, 7.776 Test distribution: $\chi^2(6-1)=\chi^2(5)$ At 0.1% level of significance, reject H_0 if $\chi^2 \geq 15.09$	**M1** for hypotheses **M1** for test distribution and critical value **M1** for E_i values **M1** for contributions to test statistic **M1** for observing test statistic lies within rejection region **A1** for conclusion

Table in (iii):

x	1	2	3	4	5	≥ 6
O_i	49	19	6	3	5	18
E_i	40	24	14.4	8.64	5.184	7.776

	<table><tr><td>x</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>≥6</td></tr><tr><td>O_i</td><td>49</td><td>19</td><td>6</td><td>3</td><td>5</td><td>18</td></tr><tr><td>E_i</td><td>40</td><td>24</td><td>14.4</td><td>8.64</td><td>5.184</td><td>7.776</td></tr><tr><td>$\frac{(O_i - E_i)^2}{E_i}$</td><td>2.025</td><td>1.0417</td><td>4.9</td><td>3.6817</td><td>0.0065 301</td><td>13.443</td></tr></table> From the table, $\chi^2 = 25.1 \geq 15.09$. Reject H_0. There is sufficient evidence at the 1% level of significance that the distribution does not follow $\text{Geo}(0.4)$.	
(iv)	By GC, mean of the given results is 2.5, variance of given results is 3.75. Using $p = 0.4$, we observe $\text{E}(X) = \frac{1}{p} = \frac{1}{0.4} = 2.5$ $\text{Var}(X) = \frac{q}{p^2} = \frac{0.6}{0.4^2} = 3.75$ Hence, Professor Vince can conclude that even if the data exhibits the same mean and variance as the hypothesized distribution, the distribution may still not be a good fit for the data.	**B1** for mean and variance of data **B1** for using parts (i) and (ii) to obtain the mean and variance of the distribution **B1** for the conclusion