Question 1

(i)	$\bar{t} = 112.8,\ \bar{v} = 0.6$ $b = \frac{Svt}{Svv} = \frac{405.2 - 3\times 564/5}{2.20 - 3^2/5} = \frac{66.8}{0.4} = 167$ OR $b = \frac{405.2/5 - 0.6\times 112.8}{2.20/5 - 0.6^2} = \frac{13.36}{0.08} = 167$ hence least squares regression line is: $t - \bar{t} = b(v - \bar{v})$ $\Rightarrow\ t - 112.8 = 167(v - 0.6)$ $\Rightarrow\ t = 167v + 12.6$	B1 for $\bar{t}$ and $\bar{v}$ used (SOI) M1 for attempt at gradient (*b*) A1 for 167 CAO M1 for equation of line A1 FT	**5**
(ii)	(*A*) For 0.5 litres, predicted time = $= 167 \times 0.5 + 12.6 = 96.1$ seconds (*B*) For 1.5 litres, predicted time = $= 167 \times 1.5 + 12.6 = 263.1$ seconds Any valid relevant comment relating to each prediction such as eg: 'First prediction is fairly reliable as it is interpolation and the data is a good fit' 'Second prediction is less certain as it is an extrapolation'	M1 for at least one prediction attempted A1 for both answers (FT their equation if $b>0$) NB for reading predictions off the graph only award A1 if accurate to nearest whole number E1 (first prediction) E1 (second prediction)	**4**
(iii)	The *v*-coefficient is the number of additional seconds required for each extra litre of water	E1 for indication of rate wrt *v* E1 *dep* for specifying ito units	**2**
(iv)	$v = 0.8 \Rightarrow$ predicted $t = 167 \times 0.8 + 12.6 = 146.2$ Residual = 156 – 146.2 = 9.8 $v = 1.0 \Rightarrow$ predicted $t = 167 \times 1.0 + 12.6 = 179.6$ Residual = 172 – 179.6 = –7.6	M1 for either prediction M1 for either subtraction A1 *CAO* for absolute value of both residuals B1 for both signs correct.	**4**
(v)	The residuals can be measured by finding the vertical distance between the plotted point and the regression line. The sign will be negative if the point is below the regression line (and positive if above).	E1 for distance E1 for vertical E1 for sign	**3**
			18

Question 2

(a) (i)	$X \sim \text{N}(28,16)$ $\text{P}(24 < X < 33) = \text{P}\left(\frac{24-28}{4} < Z < \frac{33-28}{4}\right)$ $= \text{P}(-1 < Z < 1.25)$ $= \Phi(1.25) - (1 - \Phi(1))$ $= 0.8944 - (1 - 0.8413)$ $= 0.8944 - 0.1587$ $= 0.7357$ (4 s.f.) *or* 0.736 (to 3 s.f.)	M1 for standardizing A1 for 1. 25 and -1 M1 for prob. with tables and correct structure A1 CAO (min 3 s.f., to include use of difference column)	**4**
(ii)	$25000 \times 0.7357 \times 0.1 = £1839$ $25000 \times 0.1587 \times 0.05 = £198$ Total = £1839 + £198 = £2037	M1 for either product, (with or without price) M1 for sum of both products with price A1 CAO awrt £2040	**3**
(iii)	$X \sim \text{N}(k, 16)$ From tables $\Phi^{-1}(0.95) = 1.645$ $\frac{33-k}{4} = 1.645$ $33 - k = 1.645 \times 4$ $k = 33 - 6.58$ $k = 26.42$ (4 s.f.) *or* 26.4 (to 3 s.f.)	B1 for ±1.645 seen M1 for correct equation in *k* with positive z-value A1 *CAO*	**3**
(b) (i)	H_0: $\mu = 0.155$; H_1: $\mu > 0.155$ Where μ denotes the mean weight in kilograms of the population of onions of the new variety	B1 for both correct & ito μ B1 for definition of μ	**2**
(ii)	Mean weight = 4.77/25 = 0.1908 Test statistic = $\frac{0.1908-0.155}{\sqrt{0.005}/\sqrt{25}} = \frac{0.0358}{0.01414}$ $= 2.531$ 1% level 1-tailed critical value of z = 2.326 2.531 > 2.236 so significant. There is sufficient evidence to reject H_0 It is reasonable to conclude that the new variety has a higher mean weight.	B1 M1 must include √25 A1FT B1 for 2.326 M1 For sensible comparison leading to a conclusion A1 for correct, consistent conclusion in words and in context	**6**
			18

Question 3

(i)	Mean = $\dfrac{\Sigma xf}{n} = \dfrac{0+20+12+3}{80} = \dfrac{35}{80}$ (= 0.4375)	B1 for mean ***NB answer given***	**1**
(ii)	Variance = $0.6907^2 = 0.4771$ So Poisson distribution may be appropriate, since mean is close to variance	B1 for variance E1*dep on squaring s*	**2**
(iii)	P(X = 1) = $e^{-0.4375}\dfrac{0.4375^1}{1!}$ = 0.282 (3 s.f.) *Either:* Thus the expected number of 1's is 22.6 which is reasonably close to the observed value of 20. *Or*: This probability compares reasonably well with the relative frequency 0.25	M1 for probability calc. M0 for tables unless interpolated (0.2813) A1 B1 for expectation of 22.6 or r.f. of 0.25 E1 for comparison	**4**
(iv)	$\lambda = 8\times0.4375 = 3.5$ Using tables: P($X \geq 12$) = 1 – P($X \leq 11$) = 1 – 0.9997 = 0.0003	B1 for mean (SOI) M1 for using tables to find 1 – P($X \leq 11$) A1 FT	**3**
(v)	The probability of at least 12 free repairs is very low, so the model is not appropriate. This is probably because the mean number of free repairs in the launderette will be much higher since the machines will get much more use than usual.	E1 for 'at least 12' E1 for very low E1	**3**
(vi)	*(A)* $\lambda = 0.4375 + 0.15 = 0.5875$ P(X = 3) = $e^{-0.5875}\dfrac{0.5875^3}{3!}$ = 0.0188 (3 s.f.) *(B)* P(Drier needs 1) = $e^{-0.15}\dfrac{0.15^1}{1!}$ = 0.129 P(Each needs just 1) = 0.282 $\times$0.129 = 0.036	B1 for mean (SOI) M1 A1 B1 for 0.129 (SOI) B1FT for 0.036	**3** **2**
			18

Question 4

<table>
<tr>
<td>(i)</td>
<td>H_0: no association between ambition and home location;
H_1: some association between ambition and home location;
<table>
<tr><td rowspan="2" colspan="2">Observed</td><td colspan="2">Home location</td></tr>
<tr><td>City</td><td>Non-city</td></tr>
<tr><td rowspan="2">Ambition</td><td>Good results</td><td>102</td><td>147</td></tr>
<tr><td>Other</td><td>75</td><td>156</td></tr>
</table>
<table>
<tr><td rowspan="2" colspan="2">Expected</td><td colspan="2">Home location</td></tr>
<tr><td>City</td><td>Non-city</td></tr>
<tr><td rowspan="2">Ambition</td><td>Good results</td><td>91.82</td><td>157.18</td></tr>
<tr><td>Other</td><td>85.18</td><td>145.82</td></tr>
</table>
<table>
<tr><td rowspan="2" colspan="2">Contribution to the test statistic</td><td colspan="2">Home location</td></tr>
<tr><td>City</td><td>Non-city</td></tr>
<tr><td rowspan="2">Ambition</td><td>Good results</td><td>1.129</td><td>0.659</td></tr>
<tr><td>Other</td><td>1.217</td><td>0.711</td></tr>
</table>
$X^2 = 3.716$
Refer to χ_1^2
Critical value at 5% level = 3.841
Result is not significant
There is insufficient evidence to conclude that there is any association between home location and ambition.
NB if H_0 H_1 reversed, or ‘correlation’ mentioned, do not award first B1or final B1 or final E1</td>
<td>B1 in context

M1 A1 for attempt at expected values

M1 for valid attempt at $(O-E)^2/E$

A1CAO for X^2

B1 for 1 dof SOI
B1 CAO for cv
B1 dep on attempt at cv
E1 conclusion in context</td>
<td>1

4

4</td>
</tr>
<tr>
<td>(ii) (A)</td>
<td>Expected Country, Results = 249 * 156 / 480 = 80.93
Expected Country, Other = 231 * 156 / 480 = 75.08</td>
<td>B1
B1</td>
<td>2</td>
</tr>
<tr>
<td>(B)</td>
<td>Refer to χ_2^2
Critical value at 5% level = 5.991
Result is significant
There is evidence to conclude that there is association between home location and ambition.</td>
<td>B1 for 2 dof SOI
B1 CAO for cv
E1 for conclusion in context</td>
<td>3</td>
</tr>
<tr>
<td>(C)</td>
<td>‘Country’ students are much less likely than city or town to have ‘Results’ as their main ambition. Low contributions show that city and town students do not appear to differ markedly in their ambitions.</td>
<td>E1 for correct obs^n for ‘Country’
E1 for additional correct observation (must refer to contributions)</td>
<td>2</td>
</tr>
<tr>
<td>(iii)</td>
<td>Conclusion in (i) is valid if only categorizing home location into city and non-city. However if non-city is subdivided into town and country this additional subdivision gives the data more precision and allows the relationship in part (ii) (C) to be revealed.</td>
<td>E1

E1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18</td>
</tr>
</table>