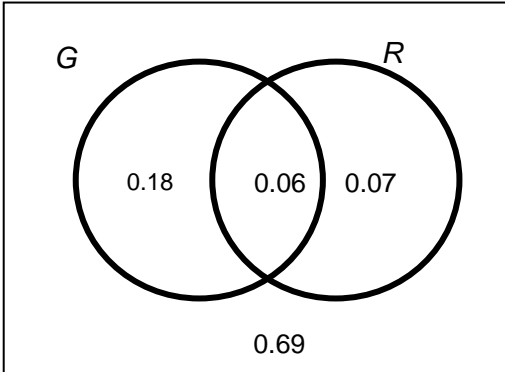


4766 Statistics 1

1	(i)	$ \begin{array}{c} 5 \quad \quad 2 \\ 6 \quad \quad 3 \quad 4 \quad 7 \quad 8 \\ 7 \quad \quad 1 \quad 2 \quad 2 \quad 3 \quad 4 \quad 5 \quad 5 \quad 7 \quad 9 \\ 8 \quad \quad 1 \\ \text{Key } 6 \quad \quad 3 \text{ represents } 63 \text{ mph} \end{array} $	G1 stem G1 leaves CAO G1 sorted G1 key	[4]
	(ii)	Median = 72 Midrange = 66.5	B1 FT B1 CAO	[2]
	(iii)	<i>EITHER:</i> Median since midrange is affected by outlier (52) <i>OR:</i> Median since the lack of symmetry renders the midrange less representative	E1 for median E1 for explanation	[2]
		TOTAL		[8]
2	(i)	(A) $P(X = 10) = P(5 \text{ then } 5) = 0.4 \times 0.25 = 0.1$ (B) $P(X = 30) = P(10 \text{ and } 20) = 0.4 \times 0.25 + 0.2 \times 0.5 = 0.2$	B1 ANSWER GIVEN	[1]
	(ii)	$E(X) = 10 \times 0.1 + 15 \times 0.4 + 20 \times 0.1 + 25 \times 0.2 + 30 \times 0.2 = 20$ $E(X^2) = 100 \times 0.1 + 225 \times 0.4 + 400 \times 0.1 + 625 \times 0.2 + 900 \times 0.2 = 445$ $\text{Var}(X) = 445 - 20^2 = 45$	M1 for $\sum rp$ (at least 3 terms correct) A1 CAO M1 for $\sum r^2p$ (at least 3 terms correct) M1 dep for – their $E(X)^2$ A1 FT their $E(X)$ provided $\text{Var}(X) > 0$	[5]
		TOTAL		[8]
3	(i)		G1 for two labelled intersecting circles G1 for at least 2 correct probabilities G1 for remaining probabilities	[3]
	(ii)	$P(G) \times P(R) = 0.24 \times 0.13 = 0.0312 \neq P(G \cap R) \text{ or } \neq 0.06$ So not independent.	M1 for 0.24×0.13 A1	[2]

	(iii)	$P(R G) = \frac{P(R \cap G)}{P(G)} = \frac{0.06}{0.24} = \frac{1}{4} = 0.25$	M1 for numerator M1 for denominator A1 CAO	[3]
			TOTAL	[8]
4	(i)	$P(20 \text{ correct}) = \binom{30}{20} \times 0.6^{20} \times 0.4^{10} = 0.1152$	M1 $0.6^{20} \times 0.4^{10}$ M1 $\binom{30}{20} \times p^{20} q^{10}$ A1 CAO	[3]
	(ii)	Expected number = $100 \times 0.1152 = 11.52$	M1 A1 FT (Must not round to whole number)	[2]
			TOTAL	[5]
5	(i)	$P(\text{Guess correctly}) = 0.1^4 = 0.0001$	B1 CAO	[1]
	(ii)	$P(\text{Guess correctly}) = \frac{1}{4!} = \frac{1}{24}$	M1 A1 CAO	[2]
			TOTAL	[3]
6	(i)	$20 \times 19 \times 18 = 6840$	M1 A1	[2]
	(ii)	$20^3 - 20 = 7980$	M1 for figures – 20 A1	[2]
			TOTAL	[4]

7	(i)	$10 \times 2 = 20.$	M1 for 10×2 A1 CAO	[2]												
	(ii)	$\text{Mean} = \frac{10 \times 65 + 35 \times 75 + 55 \times 85 + 20 \times 95}{120} = \frac{9850}{120} = 82.08$ <p>It is an estimate because the data are grouped.</p>	M1 for midpoints M1 for double pairs A1 CAO E1 indep	[4]												
	(iii)	$10 \times 65^2 + 35 \times 75^2 + 55 \times 85^2 + 20 \times 95^2 (= 817000)$ $S_{xx} = 817000 - \frac{9850^2}{120} (= 8479.17)$ $s = \sqrt{\frac{8479.17}{119}} = 8.44$	M1 for Σfx^2 M1 for valid attempt at S_{xx} A1 CAO	[3]												
	(iv)	$\bar{x} - 2s = 82.08 - 2 \times 8.44 = 65.2$ $\bar{x} + 2s = 82.08 + 2 \times 8.44 = 98.96$ <p>So there are probably some outliers.</p>	M1 FT for $\bar{x} - 2s$ M1 FT for $\bar{x} + 2s$ A1 for both E1 dep on A1	[4]												
	(v)	Negative.	E1	[1]												
	(vi)	<table style="margin-left: auto; margin-right: auto;"> <tbody> <tr> <td>Upper bound</td> <td>60</td> <td>70</td> <td>80</td> <td>90</td> <td>100</td> </tr> <tr> <td>Cumulative frequency</td> <td>0</td> <td>10</td> <td>45</td> <td>100</td> <td>120</td> </tr> </tbody> </table> 	Upper bound	60	70	80	90	100	Cumulative frequency	0	10	45	100	120	C1 for cumulative frequencies S1 for scales L1 for labels 'Length and CF' P1 for points J1 for joining points dep on P1 All dep on attempt at cumulative frequency.	[5]
Upper bound	60	70	80	90	100											
Cumulative frequency	0	10	45	100	120											
TOTAL				[19]												

8	(i)	(A) $P(\text{Low on all 3 days}) = 0.5^3 = 0.125$ or $1/8$	M1 for 0.5^3 A1 CAO	[2]
		(B) $P(\text{Low on at least 1 day}) = 1 - 0.5^3 = 1 - 0.125 = 0.875$	M1 for $1 - 0.5^3$ A1 CAO	[2]
		(C) $P(\text{One low, one medium, one high})$ $= 6 \times 0.5 \times 0.35 \times 0.15 = 0.1575$	M1 for product of probabilities $0.5 \times 0.35 \times 0.15$ or $21/800$ M1 $\times 6$ or $\times 3!$ or 3P_3 A1 CAO	[3]
	(ii)	$X \sim B(10, 0.15)$ (A) $P(\text{No days}) = 0.85^{10} = 0.1969$ Or from tables $P(\text{No days}) = 0.1969$	M1 A1	[2]
		(B) <i>Either</i> $P(1 \text{ day}) = \binom{10}{1} \times 0.15^1 \times 0.85^9 = 0.3474$ <i>or</i> from tables $P(1 \text{ day}) = P(X \leq 1) - P(X \leq 0)$ $= 0.5443 - 0.1969 = 0.3474$	M1 $0.15^1 \times 0.85^9$ M1 $\binom{10}{1} \times p^1 q^9$ A1 CAO OR: M2 for $0.5443 - 0.1969$ A1 CAO	[3]
	(iii)	Let $X \sim B(20, 0.5)$ <i>Either:</i> $P(X \geq 15) = 1 - 0.9793 = 0.0207 < 5\%$ <i>Or:</i> Critical region is $\{15, 16, 17, 18, 19, 20\}$ 15 lies in the critical region. So there is sufficient evidence to reject H_0 Conclude that there is enough evidence to indicate that the probability of low pollution levels is higher on the new street. H_1 has this form as she believes that the probability of a low pollution level is greater in this street.	<i>Either:</i> B1 for correct probability of 0.0207 M1 for comparison <i>Or:</i> B1 for CR, M1 for comparison A1 CAO dep on B1M1 E1 for conclusion in context E1 indep	[5]
			TOTAL	[17]