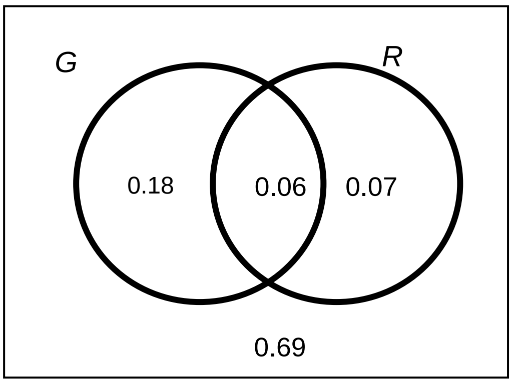
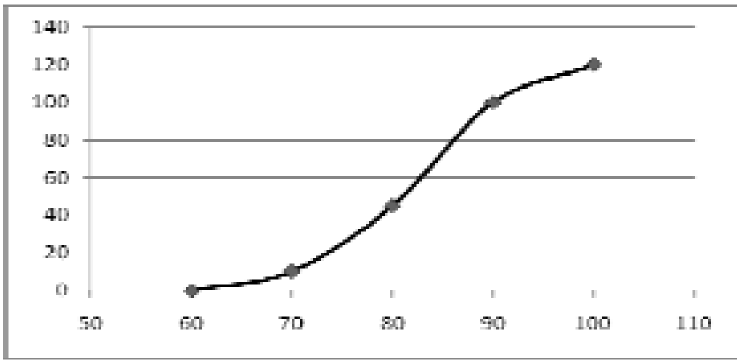
# 4766 Statistics 1

| | | | | |
|---|---|---|---|---|
| **1** | **(i)** | 5 \| 2<br>6 \| 3 4 7 8<br>7 \| 1 2 2 3 4 5 5 7 9<br>8 \| 1<br>Key 6 \| 3 represents 63 mph | G1 stem<br>G1 leaves CAO<br>G1 sorted<br>G1 key | **[4]** |
| | **(ii)** | Median = 72<br>Midrange = 66.5 | B1 FT<br>B1 CAO | **[2]** |
| | **(iii)** | *EITHER:* Median since midrange is affected by outlier (52)<br>*OR:* Median since the lack of symmetry renders the midrange less representative | E1 for median<br>E1 for explanation<br>**TOTAL** | **[2]**<br>**[8]** |
| **2** | **(i)** | ($A$) $P(X = 10) = P(5 \text{ then } 5) = 0.4\times0.25 = 0.1$ | B1 **ANSWER GIVEN** | **[1]** |
| | | ($B$) $P(X = 30) = P(10 \text{ and } 20) = 0.4\times0.25 + 0.2\times0.5 = 0.2$ | M1 for full calculation<br>A1 **ANSWER GIVEN** | **[2]** |
| | **(ii)** | $E(X) = 10\times0.1+15\times0.4+20\times0.1+25\times0.2+30\times0.2=20$<br><br>$E(X^2) =$<br>$100\times0.1+225\times0.4+400\times0.1+625\times0.2+900\times0.2=445$<br>$\text{Var}(X) = 445 - 20^2 = 45$ | M1 for $\Sigma rp$ (at least 3 terms correct)<br>A1 CAO<br>M1 for $\Sigma r^2p$ (at least 3 terms correct)<br>M1 dep for – their $E(X)^2$<br>A1 FT their E(X) provided Var( X ) > 0<br>**TOTAL** | **[5]**<br>**[8]** |
| **3** | **(i)** |  | G1 for two labelled intersecting circles<br>G1 for at least 2 correct probabilities<br>G1 for remaining probabilities | **[3]** |
| | **(ii)** | $P(G) \times P(R) = 0.24 \times 0.13 = 0.0312 \neq P(G\cap R)$ or $\neq 0.06$<br>So not independent. | M1 for $0.24 \times 0.13$<br>A1 | **[2]** |

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

| | | | | |
|---|---|---|---|---|
| 7 | (i) | $10 \times 2 = 20$. | M1 for $10 \times 2$<br>A1 CAO | [2] |
| | (ii) | Mean = $\dfrac{10\times65+35\times75+55\times85+20\times95}{120}=\dfrac{9850}{120}=82.08$<br>It is an estimate because the data are grouped. | M1 for midpoints<br>M1 for double pairs<br>A1 CAO<br>E1 indep | [4] |
| | (iii) | $10 \times 65^2 + 35 \times 75^2 + 55 \times 85^2 + 20 \times 95^2$ (= 817000)<br>$S_{xx} = 817000-\dfrac{9850^2}{120}$ (= 8479.17)<br>$s = \sqrt{\dfrac{8479.17}{119}} = 8.44$ | M1 for $\Sigma fx^2$<br>M1 for valid attempt at $S_{xx}$<br>A1 CAO | [3] |
| | (iv) | $\bar{x} - 2s = 82.08 - 2\times8.44 = 65.2$<br>$\bar{x} + 2s = 82.08 + 2\times8.44 = 98.96$<br>So there are probably some outliers. | M1 FT for $\bar{x} - 2s$<br>M1 FT for $\bar{x} + 2s$<br>A1 for both<br>E1 dep on A1 | [4] |
| | (v) | Negative. | E1 | [1] |
| | (vi) | Upper bound: 60, 70, 80, 90, 100<br>Cumulative frequency: 0, 10, 45, 100, 120<br> | C1 for cumulative frequencies<br>S1 for scales<br>L1 for labels 'Length and CF'<br>P1 for points<br>J1 for joining points dep on P1<br>All dep on attempt at cumulative frequency. | [5] |
| | | | TOTAL | [19] |

| | | | | |
|---|---|---|---|---|
| **8** | **(i)** | (*A*) P(Low on all 3 days) = $0.5^3 = 0.125$ or $^1/_8$ | M1 for $0.5^3$<br>A1 CAO | **[2]** |
| | | (*B*) P(Low on at least 1 day) = $1 – 0.5^3 = 1 – 0.125 = 0.875$ | M1 for $1 – 0.5^3$<br>A1 CAO | **[2]** |
| | | (*C*) P(One low, one medium, one high)<br>$= 6 \times 0.5 \times 0.35 \times 0.15 = 0.1575$ | M1 for product of probabilities $0.5 \times 0.35 \times 0.15$ or $^{21}/_{800}$<br>M1 $\times 6$ or $\times 3!$ or $^3P_3$<br>A1 CAO | **[3]** |
| | **(ii)** | X ~ B(10, 0.15)<br>(*A*) P(No days) = $0.85^{10} = 0.1969$<br>Or from tables P(No days) = 0.1969 | M1<br>A1 | **[2]** |
| | | (*B*) *Either* P(1 day) $= \binom{10}{1} \times 0.15^1 \times 0.85^9 = 0.3474$<br>*or* from tables P(1 day) $= P(X \le 1) - P(X \le 0)$<br>$= 0.5443 – 0.1969 = 0.3474$ | M1 $0.15^1 \times 0.85^9$<br>M1 $\binom{10}{1} \times p^1 q^9$<br>A1 CAO<br>OR: M2 for 0.5443 – 0.1969<br>A1 CAO | **[3]** |
| | **(iii)** | Let $X$ ~ B(20, 0.5)<br>*Either:* $P(X \ge 15) = 1 – 0.9793 = 0.0207 < 5\%$<br><br>*Or:* Critical region is {15,16,17,18,19,20}<br>15 lies in the critical region.<br><br>So there is sufficient evidence to reject $H_0$<br><br>Conclude that there is enough evidence to indicate that the probability of low pollution levels is higher on the new street.<br><br>$H_1$ has this form as she believes that the probability of a low pollution level is greater in this street. | *Either:*<br>B1 for correct probability of 0.0207<br>M1 for comparison<br>*Or:*<br>B1 for CR,<br>M1 for comparison<br><br>A1 CAO dep on B1M1<br><br>E1 for conclusion in context<br><br>E1 indep<br><br>**TOTAL** | **[5]**<br><br>**[17]** |