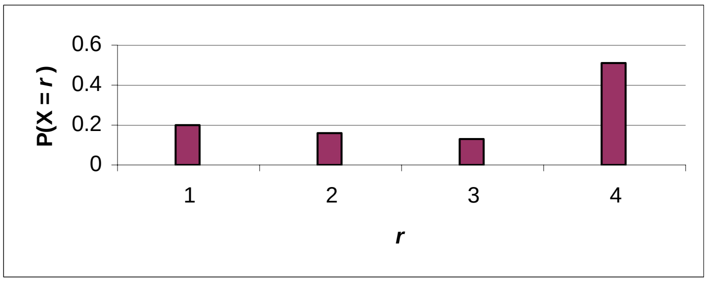
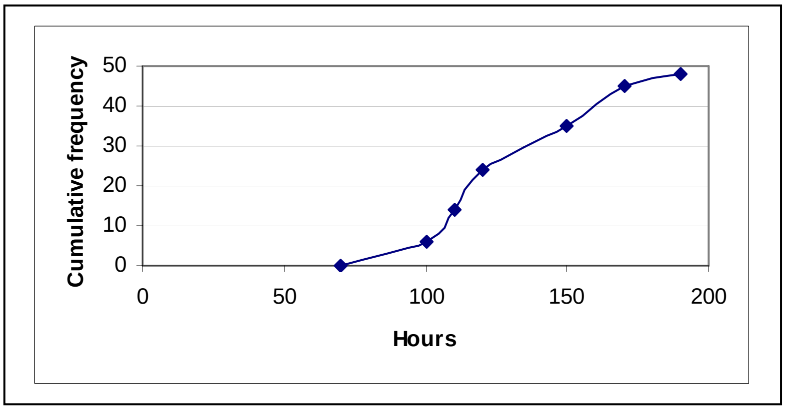
4766 Statistics 1

Q1 (i)	Mode = 7 Median = 12.5	B1 **cao** B1 **cao**	**2**
(ii)	Positive or positively skewed	E1	**1**
(iii)	(*A*) Median (*B*) There is a large outlier or possible outlier of 58 / figure of 58. Just 'outlier' on its own without reference to either 58 or large scores E0 Accept the large outlier affects the mean (more) E1	E1 cao E1indep	**2**
(iv)	There are $14.75 \times 28 = 413$ messages So total cost = 413×10 pence = £41.30	M1 for 14.75×28 but 413 can also imply the mark A1**cao**	**2**
		TOTAL	**7**
Q2 (i)	$\binom{4}{3} \times 3! = 4 \times 6 = 24$ codes or ${}^4P_3 = 24$ (M2 for ${}^4P_3$) *Or* $4 \times 3 \times 2 = 24$	M1 for 4 M1 for $\times 6$ A1	**3**
(ii)	$4^3 = 64$ codes	M1 for 4^3 A1 **cao**	**2**
		TOTAL	**5**
Q3 (i)	Probability = $0.3 \times 0.8 = 0.24$	M1 for 0.8 from (1-0.2) A1	**2**
(ii)	*Either:* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ $= 0.3 + 0.2 - 0.3 \times 0.2$ $= 0.5 - 0.06 = 0.44$ *Or:* $P(A \cup B) = 0.7 \times 0.2 + 0.3 \times 0.8 + 0.3 \times 0.2$ $= 0.14 + 0.24 + 0.06 = 0.44$ *Or:* $P(A \cup B) = 1 - P(A' \cap B')$ $= 1 - 0.7 \times 0.8 = 1 - 0.56 = 0.44$	M1 for adding 0.3 and 0.2 M1 for **subtraction** of (0.3×0.2) A1 **cao** M1 either of first terms M1 for last term A1 M1 for 0.7×0.8 or 0.56 M1 for complete method as seen A1	**3**
(iii)	$P(A\|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.06}{0.44} = \frac{6}{44} = 0.136$	M1 for numerator of their 0.06 only M1 for 'their 0.44' in denominator A1 FT (must be valid p)	**3**
		TOTAL	**8**

Q4 (i)	E(X) = $1\times0.2+2\times0.16+3\times0.128+4\times0.512=2.952$ Division by 4 or other spurious value at end loses A mark E(X^2) = $1\times0.2+4\times0.16+9\times0.128+16\times0.512=10.184$ Var(X) = $10.184 - 2.952^2 = 1.47$ (to 3 s.f.)	M1 for $\Sigma\, rp$ (at least 3 terms correct) A1 cao M1 for $\Sigma\, x^2p$ *at least 3 terms correct* M1 for E(X^2) – E(X) ^{2} Provided ans > 0 A1 FT their E(X) but not a wrong E(X^2)	**5**
(ii)	Expected cost = 2.952 × £45000 = £133000 (3sf)	B1 FT (no extra multiples / divisors introduced at this stage)	**1**
(iii)		G1 labelled linear scales G1 height of lines	**2**
		TOTAL	**8**
Q5 (i)	Impossible because the competition would have finished as soon as Sophie had won the first 2 matches	E1	**1**
(ii)	**SS, JSS, JSJSS**	B1, B1, B1 (-1 each error or omission)	**3**
(iii)	$0.7^2 + 0.3\times0.7^2 + 0.7\times0.3\times0.7^2 = 0.7399$ or 0.74(0) { 0.49 + 0.147 + 0.1029 = 0.7399}	M1 for any correct term M1 for any other correct term M1 for sum of all three correct terms A1 cao	**4**
		TOTAL	**8**

	Section B		
Q6 (i)	Mean = $\frac{180.6}{12}$ = 15.05 or 15.1 S_{xx} = $3107.56 - \frac{180.6^2}{12}$ or 3107.56 – 12(their 15.05)^2 = (389.53) $s = \sqrt{\frac{389.53}{11}}$ = 5.95 or better NB Accept answers seen without working (from calculator)	B1 for mean M1 for attempt at S_{xx} A1 cao	3
(ii)	$\bar{x}$ + 2s = 15.05 + 2 × 5.95 = 26.95 $\bar{x}$ – 2s = 15.05 – 2 × 5.95 = 3.15 So no outliers	M1 for attempt at either M1 for both A1 for limits and conclusion FT their mean and sd	3
(iii)	New mean = 1.8 × 15.05 + 32 = 59.1 New *s* = 1.8 × 5.95 = 10.7	B1FT M1 A1FT	3
(iv)	New York has a higher mean or ' is on average' higher (oe) New York has greater spread /range /variation or SD (oe)	E1FT using ^{0}F ($\bar{x}$ dep) E1FT using ^{0}F (σ dep)	2
(v)	Upper bound: (70), 100, 110, 120, 150, 170, 190 Cumulative frequency: (0), 6, 14, 24, 35, 45, 48  NB all G marks dep on attempt at cumulative frequencies.	B1 for all **correct** cumulative frequencies (may be implied from graph). **<u>Ignore cf of 0 at this stage</u>** G1 for linear scales (linear from 70 to 190) ignore x < 70 vertical: 0 to 50 but not beyond 100 (no inequality scales) G1 for labels G1 for points plotted as (UCB, their cf). <u>Ignore (70,0) at this stage</u>. **No** mid – point or LCB plots**.** G1 for joining all of 'their points'(line or smooth curve) **AND now including (70,0)**	5
(vi)	NB All G marks dep on attempt at cumulative frequencies Line on graph at cf = 43.2(soi) or used 90th percentile = 166	M1 for use of 43.2 A1FT but **dep on 3rd G mark earned**	2
		TOTAL	**18**

Q7 (i)	X ~ B(12, 0.05) $(\boldsymbol{A})$ $\mathrm{P}(\boldsymbol{X}=1) = \binom{12}{1} \times 0.05 \times 0.95^{11} = 0.3413$ OR from tables $0.8816 - 0.5404 = 0.3412$ $(\boldsymbol{B})$ $\mathrm{P}(\boldsymbol{X} \geq 2) = 1 - 0.8816 = 0.1184$ $(\boldsymbol{C})$ Expected number $\mathrm{E}(\boldsymbol{X}) = \boldsymbol{np} = 12 \times 0.05 = 0.6$	M1 0.05×0.95^{11} M1 $\binom{12}{1} \times pq^{11}$ (p+q) = 1 A1 cao OR: M1 for 0.8816 seen and M1 for subtraction of 0.5404 A1 cao M1 for 1 – P(X ≤ 1) A1 cao M1 for 12×0.05 A1 cao (= 0.6 seen)	3 2 2
(ii)	*Either*: $1 - 0.95^n \leq \frac{1}{3}$ $0.95^n \geq \frac{2}{3}$ $n \leq \log \frac{2}{3} / \log 0.95$, so $n \leq 7.90$ Maximum $n = 7$ *Or:* (using tables with $p = 0.05$): $n = 7$ leads to P($X \geq 1$) = 1 - P($X = 0$) = 1 – 0.6983 = 0.3017 (< ⅓) or 0.6983 (> 2/3) $n = 8$ leads to P($X \geq 1$) = 1 - P($X = 0$) = 1 – 0.6634 = 0.3366 (> ⅓) or 0.6634 (< 2/3) Maximum $n = 7$ (total accuracy needed for tables) *Or:* (using trial and improvement): $1 - 0.95^7 = 0.3017$ (< ⅓) or $0.95^7 = 0.6983$ (> 2/3) $1 - 0.95^8 = 0.3366$ (> ⅓) or $0.96^8 = 0.6634$ (< 2/3) Maximum $n = 7$ (3 sf accuracy for calculations) NOTE: $n = 7$ unsupported scores SC1 only	M1 for equation in n M1 for use of logs A1 cao M1indep M1indep A1 cao dep on both M's M1indep (as above) M1indep (as above) A1 cao dep on both M's	 3
(iii)	Let X ~ B(60, p) Let p = probability of a bag being faulty H_0: $p = 0.05$ H_1: $p < 0.05$ P($X \leq 1$) = $0.95^{60} + 60 \times 0.05 \times 0.95^{59} = 0.1916 > 10\%$ So not enough evidence to reject H_0 Conclude that there is not enough evidence to indicate that the new process reduces the failure rate or scientist incorrect/ wrong.	B1 for definition of p B1 for H_0 B1 for H_1 M1 A1 for probability M1 for comparison A1 E1	 8
		TOTAL	**18**