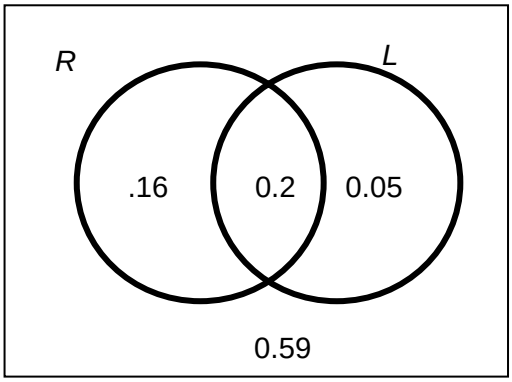
4766 Statistics 1

Section A

Q1 (i)	(With $\sum fx = 7500$ and $\sum f = 10000$ then arriving at the mean) (i) £0.75 scores (B1, B1) (ii) 75p scores (B1, B1) (iii) 0.75p scores (B1, B0) (incorrect units) (iv) £75 scores (B1, B0) (incorrect units) **After B0, B0** then sight of $\frac{7500}{10000}$ scores SC1. SC1or an answer in the range £0.74 - £0.76 or 74p – 76p (both inclusive) scores SC1 (units essential to gain this mark)	B1 for numerical mean (0.75 or 75 seen) B1dep for correct units attached	
	Standard Deviation: (CARE NEEDED here with close proximity of answers) • 50.2(0) using divisor 9999 scores B2 (50.20148921) • 50.198 (= 50.2) using divisor 10000 scores B1(*rmsd*) • If divisor is not shown (or calc used) and only an answer of 50.2 (i.e. not coming from 50.198) is seen then award B2 on b.o.d. (default)	B2 correct s.d. (B1) correct rmsd (B2) default	
	After B0 scored then an attempt at S_{xx} as evident by either $S_{xx} = (5000 + 200000 + 25000000) - \frac{7500^2}{10000}$ (= 25199375) or $S_{xx} = (5000 + 200000 + 25000000) - 10000(0.75)^2$ **scores (M1) or M1ft 'their 7500^2' or 'their 0.75^2'** NB The structure must be correct in both above cases with a max of 1 slip only after applying the f.t.	$\sum fx^2$ **= 25,205,000** **Beware** $\sum x^2$ **=25, 010, 100** **After B0** scored then (M1) or M1f.t. for attempt at S_{xx} *NB full marks for correct results from recommended method which is use of calculator functions*	**4**

(ii)	P(Two £10 or two £100) $= \frac{50}{10000}\times\frac{49}{9999}+\frac{20}{10000}\times\frac{19}{9999}$ = 0.0000245 + 0.0000038 = (0.00002450245 + 0.00000380038) = 0.000028(3) o.e. = (0.00002830283) After M0, M0 then $\frac{\mathbf{50}}{\mathbf{10000}}\times\frac{\mathbf{50}}{\mathbf{10000}}+\frac{\mathbf{20}}{\mathbf{10000}}\times\frac{\mathbf{20}}{\mathbf{10000}}$ o.e. Scores SC1 (ignore final answer but SC1 may be implied by sight of 2.9×10^{-5} o.e.) Similarly, $\frac{\mathbf{50}}{\mathbf{10000}}\times\frac{\mathbf{49}}{\mathbf{10000}}+\frac{\mathbf{20}}{\mathbf{10000}}\times\frac{\mathbf{19}}{\mathbf{10000}}$ scores SC1	M1 for either correct product seen (ignore any multipliers) M1 sum of both correct (ignore any multipliers) A1 CAO (as opposite with no rounding) (SC1 case #1) (SC1 case #2) **CARE** answer is also 2.83×10^{-5}	**3**
		TOTAL	**7**
Q2 (i)	Either P(all correct) $=\frac{1}{6}\times\frac{1}{5}\times\frac{1}{4}\times\frac{1}{3}\times\frac{1}{2}\times\frac{1}{1}=\frac{1}{720}$ or P(all correct) $=\frac{1}{6!}=\frac{1}{720}=0.00139$	M1 for 6! Or 720 (sioc) or product of fractions A1 CAO (accept 0.0014)	**2**
(ii)	Either P(picks T, O, M) $=\frac{3}{6}\times\frac{2}{5}\times\frac{1}{4}=\frac{1}{20}$ or P(picks T, O, M) $=\frac{1}{6}\times\frac{1}{5}\times\frac{1}{4}\times3!=\frac{1}{20}$ or P(picks T, O, M) $=\frac{1}{\binom{6}{3}}=\frac{1}{20}$	M1 for denominators M1 for numerators or 3! A1 CAO Or M1 for $\binom{6}{3}$ or 20 sioc M1 for $1/\binom{6}{3}$ A1 CAO	**3**
		TOTAL	**5**
Q3 (i)	p = 0.55	B1 cao	**1**
(ii)	E(X) = $0\times0.55+1\times0.1+2\times0.05+3\times0.05+4\times0.25=1.35$ E(X^2) = $0\times0.55+1\times0.1+4\times0.05+9\times0.05+16\times0.25$ = 0 + 0.1 + 0.2 + 0.45 + 4 = (4.75) Var(X) = 'their' $4.75 - 1.35^2$ = 2.9275 awfw (2.9275 – 2.93)	M1 for Σrp (at least 3 non zero terms correct) A1 CAO(no 'n' or 'n-1' divisors) M1 for Σr^2p (at least 3 non zero terms correct) M1dep for – their E(X)² provided Var(X) > 0 A1 cao (no 'n' or 'n-1' divisors)	**5**
(iii)	P(At least 2 both times) = $(0.05+0.05+0.25)^2$ = 0.1225 o.e.	M1 for $(0.05+0.05+0.25)^2$ or 0.35^2 seen A1cao: awfw (0.1225 - 0.123) or 49/400	**2**

		TOTAL	**8**

Q4 (i)	$X \sim B(50, 0.03)$ **(*A*)** $P(X = 1) = \binom{50}{1} \times 0.03 \times 0.97^{49} = 0.3372$ **(*B*)** $P(X = 0) = 0.97^{50} = 0.2181$ $P(X > 1) = 1 - 0.2181 - 0.3372 = 0.4447$	M1 0.03×0.97^{49} or 0.0067(4)…. M1 $\binom{50}{1} \times pq^{49}$ (p+q =1) A1 CAO **(awfw 0. 337 to 0. 3372) or 0.34(2s.f.) or 0.34(2d.p.) but not just 0.34** B1 for 0.97^{50} or 0.2181 (awfw 0.218 to 0.2181) M1 for 1 – ('their' p (X = 0) + 'their' p(X = 1)) must have both probabilities A1 CAO (awfw 0.4447 to 0.445)	**3** **3**
(ii)	Expected number = np = 240 × 0.3372 = 80.88 – 80.93 = (81) *Condone 240 × 0.34 = 81.6 = (82) but for M1 A1f.t.*	M1 for 240 × prob (A) A1FT	**2**
		TOTAL	**8**
Q5 (i)	P(R) × P(*L*) = 0.36 × 0.25 = 0.09 ≠ P(*R*∩*L*) Not equal so not independent. (Allow 0.36 × 0.25 ≠ 0.2or 0.09 ≠ 0.2 or ≠ p(R ∩ L) so not independent)	M1 for 0.36 × 0.25 or 0.09 seen A1 (numerical justification needed)	**2**
(ii)		G1 for two overlapping circles labelled G1 for 0.2 and either 0.16 *or* 0.05 in the **correct places** G1 for all 4 **correct** probs in the **correct** places (including the 0.59) The last two G marks are independent of the labels	**3**
(iii)	$P(L \mid R) = \frac{P(L \cap R)}{P(R)} = \frac{0.2}{0.36} = \frac{5}{9} = 0.556$ (awrt 0.56) This is the probability that Anna is late given that it is raining. (must be in context) Condone 'if ' or 'when' or 'on a rainy day' for 'given that' but **not** the words 'and' or 'because' or 'due to'	M1 for 0.2/0.36 o.e. A1 cao E1 (indep of M1A1) **Order/structure must be correct i.e. no reverse statement**	**3**
		TOTAL	**8**

Section B

Q6 (i)	Median = 4.06 – 4.075 (inclusive) Q_1 = 3.8 Q_3 = 4.3 Inter-quartile range = 4.3 – 3.8 = 0.5	B1cao B1 for Q_1 (cao) B1 for Q_3 (cao) B1 ft for IQR must be using t-values not locations to earn this mark	**4**
(ii)	Lower limit ' their 3.8' – 1.5 × 'their 0.5' = (3.05) Upper limit ' their 4.3' + 1.5 × 'their 0.5' = (5.05) Very few if any temperatures <u>below 3.05 (but not zero)</u> None <u>above 5.05</u> 'So few, if any outliers' scores SC1	B1ft: must have -1.5 B1ft: must have +1.5 E1ft dep on -1.5 and Q_1 E1ft dep on+1.5 and Q_3 Again, must be using t-values NOT locations to earn these 4 marks	**4**
(iii)	Valid argument such as 'Probably not, because there is nothing to suggest that they are not genuine data items; (they do not appear to form a separate pool of data.') Accept: exclude outlier – 'measuring equipment was wrong' or 'there was a power cut' or ref to hot / cold day [Allow suitable valid alternative arguments]	E1	**1**
(iv)	Missing frequencies 25, 125, 50	B1, B1, B1 (all cao)	**3**
(v)	Mean = (3.2×25 + 3.6×125 + 4.0×243 + 4.4×157 + 4.8×50)/600 = 2432.8/600 = 4.05(47)	M1 for at least 4 midpoints correct and being used in attempt to find $\sum ft$ A1cao: awfw (4.05 – 4.055) ISW or rounding	**2**
(vi)	New mean = 1.8 × 'their 4.05(47)' + 32 = 39.29(84) to 39.3 New s = 1.8 × 0.379 = 0.682	B1 FT M1 for 1.8 × 0.379 A1 CAO awfw (0.68 – 0.6822)	**3**
		TOTAL	**17**

Q7 (i)	$X \sim B(10, 0.8)$ **(*A*)** *Either* $P(X = 8) = \binom{10}{8} \times 0.8^8 \times 0.2^2 = 0.3020$ (awrt) *or* $P(X = 8) = P(X \leq 8) - P(X \leq 7)$ $= 0.6242 - 0.3222 = 0.3020$ **(*B*)** *Either* $P(X \geq 8) = 1 - P(X \leq 7)$ $= 1 - 0.3222 = 0.6778$ *or* $P(X \geq 8) = P(X = 8) + P(X = 9) + P(X = 10)$ $= 0.3020 + 0.2684 + 0.1074 = 0.6778$	M1 $0.8^8 \times 0.2^2$ or 0.00671… M1 $\binom{10}{8} \times p^8 q^2$; (p +q =1) Or $45 \times p^8 q^2$; (p +q =1) A1 CAO **(0.302) not 0.3** OR: M2 for 0.6242 – 0.3222 A1 CAO M1 for 1 – 0.3222 (s.o.i.) A1 CAO awfw 0.677 – 0.678 or M1 for sum of 'their' p(X=8) plus correct expressions for p(x=9) and p(X=10) A1 CAO awfw 0.677 – 0.678	**3** **2**
(ii)	Let $X \sim B(18, p)$ Let p = probability of delivery (within 24 hours) (for population) H_0: $p = 0.8$ H_1: $p < 0.8$ $P(X \leq 12) = 0.1329 > 5\%$ ref: [pp =0.0816] So not enough evidence to reject H_0 Conclude that there is not enough evidence to indicate that less than 80% of orders will be delivered within 24 hours Note: use of critical region method scores M1 for region {0,1,2,…,9, 10} M1dep for 12 does not lie in critical region then A1dep E1dep as per scheme	B1 for definition of p B1 for H_0 B1 for H_1 M1 for probability 0.1329 M1dep strictly for comparison of 0.1329 with 5% (seen or clearly implied) A1dep on both M's E1dep on M1,M1,A1 for conclusion in context	**7**

(iii)	Let $X \sim B(18, 0.8)$ H_1: $p \neq 0.8$ LOWER TAIL $P(X \leq 10) = 0.0163 < 2.5\%$ $P(X \leq 11) = 0.0513 > 2.5\%$ UPPER TAIL $P(X \geq 17) = 1 - P(X \leq 16) = 1 - 0.9009 = 0.0991 > 2.5\%$ $P(X \geq 18) = 1 - P(X \leq 17) = 1 - 0.9820 = 0.0180 < 2.5\%$ So critical region is {0,1,2,3,4,5,6,7,8,9,10,18} o.e. Condone $X \leq 10$ and $X \geq 18$ or $X = 18$ but **not** $p(X \leq 10)$ and $p(X \geq 18)$ Correct CR without supportive working scores SC2 max after the 1st B1 (SC1 for each fully correct tail of CR)	B1 for H_1 B1 for 0.0163 or 0.0513 seen M1dep for either correct comparison with **2.5% (not 5%)** (seen or clearly implied) A1dep for correct lower tail CR (must have zero) B1 for 0.0991 or 0.0180 seen M1dep for either correct comparison with **2.5% (not 5%)** (seen or clearly implied) A1dep for correct upper tail CR	**7**
		TOTAL	**19**