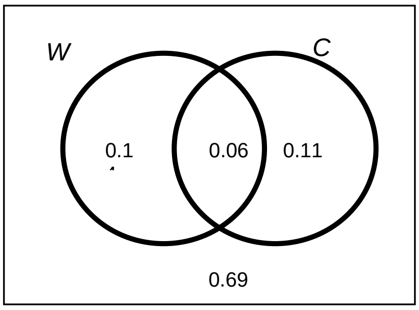
4766 Statistics 1

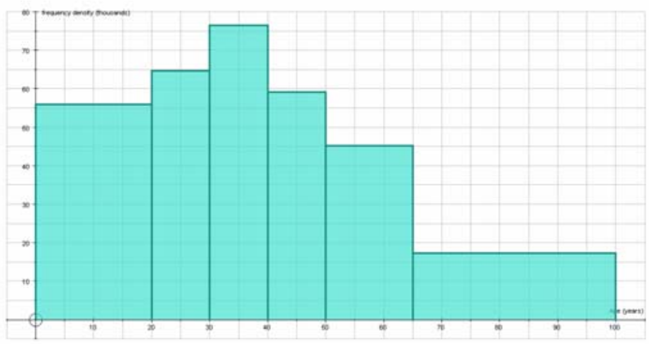
Q1 (i)	Mean = 7.35 (or better) Standard deviation: 3.69 – 3.70 (awfw) Allow s^2 = 13.62 to 13.68 Allow rmsd = 3.64 – 3.66 (awfw) **After B0, B0 scored** then if at least 4 correct mid-points seen or used.{**1.5, 4, 6, 8.5, 15**} Attempt of their mean = $\frac{\sum fx}{\mathbf{44}}$, with $301 \leq fx \leq 346$ and fx strictly from mid-points not class widths or top/lower boundaries.	B2cao $\sum fx$ = 323.5 B2cao $\sum fx^2$ = 2964.25 (B1) for variance s.o.i.o (B1) for rmsd (B1) mid-points (B1) $6.84 \leq \text{mean} \leq 7.86$	4
(ii)	Upper limit = 7.35 + 2 × 3.69 = 14.73 or 'their sensible mean' + 2 × 'their sensible s.d.' So there could be one or more outliers	M1 (with s.d. < mean) E1**dep** on B2, B2 earned and comment	2
		TOTAL	**6**
Q2 (i)	P(*W*) × P(*C*) = 0.20 × 0.17 = 0.034 P(*W*∩*C*) = 0.06 (given in the question) Not equal so not independent (Allow $0.20 \times 0.17 \neq 0.06$ or $\neq$ p ($W \cap C$) so not independent).	M1 for multiplying or 0.034 seen A1 (numerical justification needed)	2
(ii)	 The last two G marks are independent of the labels	G1 for two **overlapping** circles labelled G1 for 0.06 and either 0.14 or 0.11 in the **correct places** G1 for all 4 **correct** probs in the **correct** places **(including the 0.69)** NB No credit for Karnaugh maps here	3
(iii)	P(*W*|*C*) = $\frac{P(W \cap C)}{P(C)} = \frac{0.06}{0.17} = \frac{6}{17} = 0.353$ (awrt 0.35)	M1 for 0.06 / 0.17 A1 cao	2

(iv)	Children are more likely than adults to be able to speak Welsh or 'proportionally more children speak Welsh than adults' Do not accept: 'more Welsh children speak Welsh than adults'	E1FT Once the correct idea is seen, apply ISW	**1**
		TOTAL	**8**
Q3 (i)	$(\boldsymbol{A})$ $0.5 + 0.35 + \boldsymbol{p} + \boldsymbol{q} = 1$ so $\boldsymbol{p} + \boldsymbol{q} = 0.15$ $(\boldsymbol{B})$ $0 \times 0.5 + 1 \times 0.35 + 2\boldsymbol{p} + 3\boldsymbol{q} = 0.67$ so $2\boldsymbol{p} + 3\boldsymbol{q} = 0.32$ $(\boldsymbol{C})$ from above $2\boldsymbol{p} + 2\boldsymbol{q} = 0.30$ so $\boldsymbol{q} = 0.02, \boldsymbol{p} = 0.13$	B1 p + q in a correct equation before they reach p + q =0.15 B1 2p + 3q in a correct equation before they reach 2p + 3q = 0.32 (B1) for any 1 correct answer B2 for both correct answers	**1** **1** **2**
(ii)	E(X^2) = $0 \times 0.5 + 1 \times 0.35 + 4 \times 0.13 + 9 \times 0.02 = 1.05$ Var(X) = 'their 1.05' – 0.67^2 = 0.6011 (awrt 0.6) (M1, M1 can be earned with their p^+ and q^+ but not A mark)	M1 $\Sigma x^2 p$ (at least 2 non zero terms correct) M1dep for (– 0.67^2), provided Var(X) > 0 A1 cao (No n or n-1 divisors)	**3**
		TOTAL	**7**
Q4 (i)	$X \sim$ B(8, 0.05) $(\boldsymbol{A})$ $\text{P}(\boldsymbol{X} = 0) = 0.95^8 = 0.6634$ 0.663 or better *Or* using tables $\text{P}(\boldsymbol{X} = 0) = 0.6634$ $(\boldsymbol{B})$ $\text{P}(\boldsymbol{X} = 1) = \binom{8}{1} \times 0.05 \times 0.95^7 = 0.2793$ $\text{P}(\boldsymbol{X} > 1) = 1 - (0.6634 + 0.2793) = 0.0573$ *Or* using tables P(X > 1) = 1 – 0.9428 = 0.0572	M1 0.95^8 A1 CAO Or B2 (tables) M1 for P(X = 1)(allow 0.28 or better) M1 for 1 – P(X ≤ 1) must have both probabilities A1cao (0.0572 – 0.0573) M1 for P(X ≤ 1) 0.9428 M1 for 1 – P(X ≤ 1) A1 cao (must end in…2)	**2** **3**
(ii)	Expected number of days = 250 × 0.0572 = 14.3 awrt	M1 for 250 x prob(B) A1 FT but no rounding at end	**2**
		TOTAL	**7**

Q5 (i)	Let p = probability of remembering or naming all items (for population) (whilst listening to music.) H_0: $p = 0.35$ H_1: $p > 0.35$ H_1 has this form since the student believes that the probability will be increased/ improved/ got better /gone up.	B1 for definition of p B1 for H_0 B1 for H_1 E1dep on p>0.35 in H_0 In words not just because p > 0.35	**4**
(ii)	Let X ~ B(15, 0.35) ***Either***: $P(X \geq 8) = 1 - 0.8868 = 0.1132 > 5\%$ Or 0.8868 < 95% So not enough evidence to reject H_0 (Accept H_o) Conclude that there is not enough evidence to indicate that the probability of remembering all of the items is improved / improved/ got better /gone up. (when listening to music.) --- ***Or***: Critical region for the test is {9,10,11,12,13,14,15} 8 does not lie in the critical region. So not enough evidence to reject H_0 Conclude that there is not enough evidence to indicate that the probability of remembering all of the items is improved / improved/ got better /gone up. (when listening to music.) -- ***Or*:** The smallest critical region that 8 could fall into is {8, 9, 10, 11, 12, 13, 14, and 15}. The size of this region is 0.1132 0.1132 > 5% So not enough evidence to reject H_0 Conclude that there is not enough evidence to indicate that the probability of remembering all of the items is improved (when listening to music)	***Either:*** M1 for probability (0.1132) M1**dep** for comparison A1**dep** E1**dep on all previous marks** for conclusion in **context** ----------------------------- ***Or***: M1 for correct CR(no omissions or additions) M1**dep** for 8 does not lie in CR A1**dep** E1**dep on all previous marks** for conclusion in **context** **-----------------------------** **Or**: M1 for CR{8,9,…15}**and** size = 0.1132 M1 **dep** for comparison A1**dep** E1**dep on all previous marks** for conclusion in **context**	**4**
		TOTAL	**8**

	Section B		
Q6 (i)	(*A*) P(both rest of UK) = 0.20×0.20 = 0.04	M1 for multiplying A1cao	**2**
	(*B*) *Either:* **All 5 case** P(at least one England) = (0.79 x 0.20) + (0.79 x 0.01) + (0.20 x 0.79) + (0.01 x 0.79) + (0.79 x 0.79) = 0.158 + 0.0079 + 0.158 + 0.0079 + 0.6241 = 0.9559 *Or* P(at least one England) = 1 – P(neither England) = 1 – (0.21×0.21) = 1 – 0.0441 = 0.9559 or listing all = 1 –{ (0.2 x 0.2) + (0.2 x 0.01) + (0.01 x 0.20) + (0.01x 0.01)} = 1 – (**) = 1 – { 0.04 + 0.002 + 0.002 + 0.0001) = 1 – 0.0441 = 0.9559 Or: **All 3 case** P(at least one England) = = $0.79 \times 0.21 + 0.21 \times 0.79 + 0.79^2$ = 0.1659 + 0.1659 + 0.6241 = 0.9559	M1 for any correct term (3case or 5case) M1 for correct sum of all 3 (or of all 5) with **no extras** A1cao (condone 0.96 www) *Or* M1 for 0.21×0.21 or for (**) fully enumerated or 0.0441 seen M1**dep** for 1 – (1st part) A1cao See above for 3 case	**3**
	(*C*)*Either* 0.79 x 0.79 + 0.79 x 0.2 + 0.2 x 0.79 + 0.2 x 0.2 = 0.9801 *Or* 0.99×0.99 = 0.9801 Or 1 – { 0.79 x 0.01 + 0.2 x 0.01 + 0.01 x 0.79 + 0.01 x 0.02 + 0.01^2 } = 1 – 0.0199 = 0.9801	M1 for sight of all 4 correct terms summed A1 cao (condone 0.98 www) *or* M1 for 0.99 x 0.99 A1cao ***Or*** M1 for everything 1 – {…..} A1cao	**2**
(ii)	P(both the rest of the UK | neither overseas) $= \frac{\text{P(the rest of the UK } \textit{and} \text{ neither overseas)}}{\text{P(neither overseas)}}$ $= \frac{0.04}{0.9801} = 0.0408$ {Watch for: $\frac{answer(A)}{answer(C)}$ as evidence of method (p <1)}	M1 for numerator of 0.04 or 'their answer to (i)(A)' M1 for denominator of 0.9801 or 'their answer to (i) (C)' A1 FT (0 < p < 1) 0.041 at least	**3**

(iii)	(*A*) Probability = $1 - 0.79^5$ $= 1 - 0.3077$ $= 0.6923$ (accept awrt 0.69) see additional notes for alternative solution	M1 for 0.79^5 or 0.3077... M1 for $1 - 0.79^5$ dep A1 CAO	
	(*B*) $1 - 0.79^n > 0.9$ **EITHER:** $1 - 0.79^n > 0.9$ or $0.79^n < 0.1$ (condone = and $\geq$ throughout) but not reverse inequality $n > \frac{\log \mathbf{0.1}}{\log \mathbf{0.79}}$, so $n > 9.768\ldots$ Minimum $n = 10$ Accept $n \geq 10$	M1 for equation/inequality in n (accept either statement opposite) M1(indep) for process of using logs i.e. $\frac{\log a}{\log b}$ A1 CAO	3
	--- **OR** (using trial and improvement): Trial with 0.79^9 or 0.79^{10} $1 - 0.79^9 = 0.8801$ (< 0.9) or $0.79^9 = 0.1198$ (> 0.1) $1 - 0.79^{10} = 0.9053$ (> 0.9) or $0.79^{10} = 0.09468$ (< 0.1) Minimum $n = 10$ Accept $n \geq 10$ --- NOTE: $n = 10$ unsupported scores SC1 only	----------------------------- M1(indep) for sight of 0.8801 **or** 0.1198 M1(indep) for sight of 0.9053 **or** 0.09468 A1 dep **on both M's** cao -----------------------------	3
		TOTAL	**16**

Q7 (i)	Positive	B1	**1**
(ii)	Number of people = 20 × 33 (000) + 5 × 58 (000) = 660 (000) + 290 (000) = 950 000	M1 first term M1(indep) second term A1 cao NB answer of 950 scores M2A0	**3**
(iii)	(*A*) $a = 1810 + 340 = 2150$ (*B*) Median = age of 1 385 (000th) person or 1385.5 (000) Age 30, cf = 1 240 (000); age 40, cf = 1 810 (000) Estimate median = (30) + $\frac{\mathbf{145}}{\mathbf{570}} \times \mathbf{10}$ Median = 32.5 years (32.54...) If no working shown then 32.54 or better is needed to gain the M1A1. If 32.5 seen with **no previous working** allow SC1	M1 for sum A1 cao 2150 or 2150 thousand but not 215000 B1 for 1 385 (000) or 1385.5 M1 for attempt to interpolate $\frac{145k}{570k} \times 10$ (2.54 or better suggests this) A1 cao min 1dp	**2** **3**
(iv)	Frequency densities: 56, 65, 77, 59, 45, 17 *(accept 45.33 and 17.43 for 45 and 17)* 	B1 for any one correct B1 for all correct (soi by listing or from histogram) **Note: all G marks below *dep* on attempt at frequency density, NOT frequency** G1 Linear scales on both axes (no inequalities) G1 Heights FT their listed fds or all must be correct. **Also widths. All blocks joined** G1 Appropriate label for vertical scale eg 'Frequency density (thousands)', 'frequency (thousands) per 10 years', 'thousands of people per 10 years'. (allow key). **OR f.d.**	**5**

PMT

(v)	Any two suitable comments such as: Outer London has a greater proportion (or %) of people under 20 (or almost equal proportion) The modal group in Inner London is 20-30 but in Outer London it is 30-40 Outer London has a greater proportion (14%) of aged 65+ **All** populations in **each** age group are higher in Outer London Outer London has a more evenly spread distribution or balanced distribution (ages) o.e.	E1 E1	**2**
(vi)	Mean increase ↑ median unchanged (-) midrange increase ↑ standard deviation increase ↑ interquartile range unchanged. (-)	Any one correct B1 Any two correct B2 Any three correct B3 All **five** correct B4	**4**
		TOTAL	**20**