RECOGNISING ACHIEVEMENT

GCE

Statistics (MEI)

Advanced Subsidiary GCE

Unit G241: Statistics 1 (Z1)

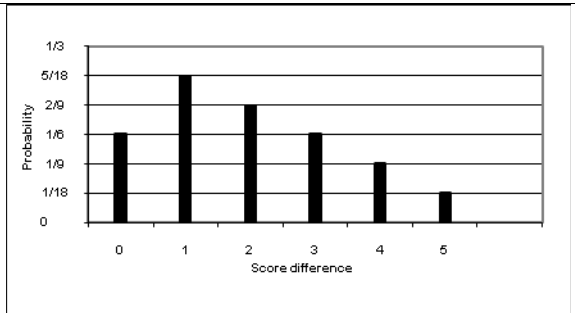
Mark Scheme for June 2011

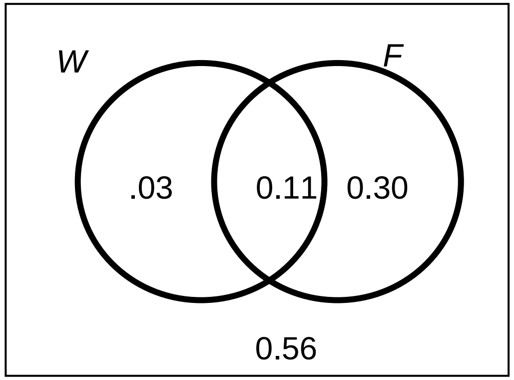
Any enquiries about publications should be addressed to:

OCR Publications
PO Box 5050
Annesley
NOTTINGHAM
NG15 0DL

Telephone: 0870 770 6622
Facsimile: 01223 552610
E-mail: publications@ocr.org.uk

	SECTION A			
Q1 (i)	$1000 \times 0.013 = 13$ Or $0.2 \times 65 = 13$ Or $0.2 \times 5 \times 13 = 13$	M1 A1 M1 for 0.2×65	**2**	Allow with or without working For MR $1000 \times 0.13 = 130$ Allow M1A0 Allow M1A0 if extra terms added eg 1000×0.004 SC1 for $1000 \times 0.014 = 14$ For whole calculation
(ii)	Positive	B1	**1**	Allow +ve but NOT skewed to the right Do not allow 'positive correlation'
(iii)	Minimum value = 1500 Maximum value = 2500	B1 Without wrong working B1 Without wrong working	**2**	Exact answers only unless good explanation such as eg no road has length zero so min is eg 1501 SC1 for lower answer between 1499 and 1501 and upper between 2499 and 2501 Allow answer given as inequality
		TOTAL	**5**	
Q2 (i)	Either $\text{P(alphabetic order)} = \frac{1}{5}\times\frac{1}{4}\times\frac{1}{3}\times\frac{1}{2}\times\frac{1}{1} = \frac{1}{120}$ or $\text{P(alphabetic order)} = \frac{1}{5!} = \frac{1}{120} = 0.00833$	M1 for 5! or 120 or ${}^5P_5$ seen or product of correct fractions A1 CAO	**2**	Allow 0.0083 or 1/120 but not 0.008
(ii)	Either $\text{P(picks Austen and Bronte)} = \frac{2}{5}\times\frac{1}{4} = \frac{1}{10}$ or $\text{P(picks Austen and Bronte)} = \frac{1}{5}\times\frac{1}{4}\times 2 = \frac{1}{10}$ or $\text{P(picks Austen and Bronte)} = \frac{1}{\binom{5}{2}} = \frac{1}{10}$	M1 for denominators M1 for 2× *dep on correct denominators* A1 CAO Or M1 for $\binom{5}{2}$ or 10 M1 for $1/\binom{5}{2}$ A1 CAO	**3**	$1/{}_5P_2$ scores M1 also 1/20 oe scores M1 even if followed by further incorrect working $\binom{5}{2}$ seen as part of a binomial expression gets M0M0A0
		TOTAL	**5**	

Q3 (i)	$P(X = 0) = 0.75^6 = 0.178$	M1 for 0.75^6 A1 CAO	**2**	Or from tables 0.1780 Or 729/4096 Allow 0.18 with working
(ii)	$E(X) = np = 50 \times 0.178 = 8.9$	M1 for product A1 FT	**2**	FT their answer to (i) providing it's a probability NB A0 if subsequently rounded
		TOTAL	**4**	
Q4 (i)		G1 labelled linear scales on both axes G1 heights	**2**	Accept *r* or *x* for horizontal label and p or better for vertical including probability distribution Visual check only Allow G1G0 for points rather than lines Bars must not be wider than gaps for second G1 Condone vertical scale 1, 2, 3, 4, 5 and Probability (×) 1/18 as label BOD for height of $r = 0$ on vertical axis
(ii)	(*A*) If $X = 1$, possible scores are (1,2), (2,3), (3,4), (4,5), (5,6) and (2,1), (3,2), (4,3), (5,4), (6,5) (All are equally likely) so probability = $\frac{10}{36} = \frac{5}{18}$	M1 A1	**2**	Also M1 for a clear correct sample space seen with the ten 1's identified by means of circles or ticks oe soi. Must be convincing. No additional values such as 0,1 and 1,0 Do not allow ' just 10 ways you can have a difference of 1 so 10/36' or equivalent SC1 for possible scores are (1,2), (2,3), (3,4), (4,5), (5,6) so probability = 2 × 5 × 1/36 with no explanation for 2 ×
	(*B*) If $X = 0$, possible scores are (1,1), (2,2), (3,3), (4,4), (5,5), (6,6) so probability = $\frac{6}{36} = \frac{1}{6}$	B1	**1**	Also B1 for a clear correct sample space seen with the six 0's identified by means of circles or ticks oe soi. Must be convincing. No additional values. Allow both dice must be the same so probability = 6/36 = 1/6. Allow 1 × 1/6= 1/6 BOD
(iii)	Mean value of X = $0\times\frac{1}{6}+1\times\frac{5}{18}+2\times\frac{2}{9}+3\times\frac{1}{6}+4\times\frac{1}{9}+5\times\frac{1}{18}=1\frac{17}{18}=1.94$	M1 for Σrp (at least 3 terms correct) A1 CAO	**2**	Or 35/18 Division by 6 or other spurious factor gets MAX M1A0
		TOTAL	**7**	

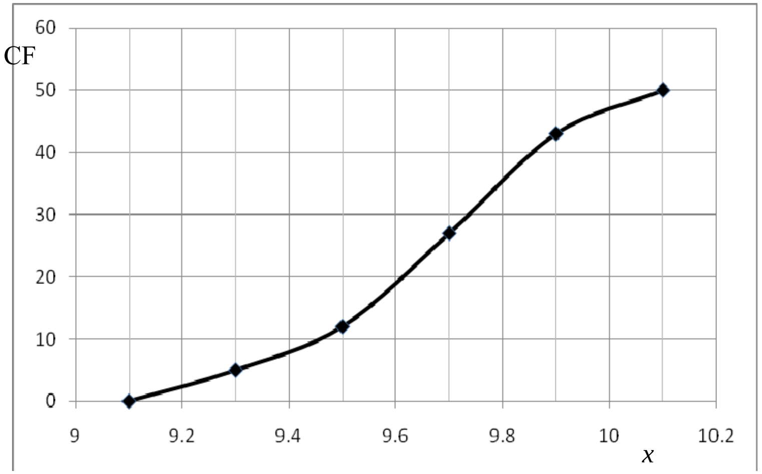
Q5 (i)		G1 for two labelled intersecting circles G1 for at least 2 correct probabilities. G1 for remaining correct probabilities	**3**	Allow labels such as P(W) and P(F) Allow other sensible shapes in place of circles
(ii)	$P(W) \times P(F) = 0.14 \times 0.41 = 0.0574 \neq P(W \cap F) = 0.11$ So not independent.	M1 for 0.41×0.14 A1 Condone dependent Must have full method www Must have either $P(W \cap F)$ or 0.11	**2**	Answer of 0.574 gets Max M1A0 Omission of 0.0574 gets M1A0 Max Or: $P(W\|F) = 0.11/0.41 = 0.268 \neq P(W)$ (= 0.14) M1 for full working $P(F\|W) = 0.11/0.14 = 0.786 \neq P(F)$ (= 0.41) M1 for full working No marks without correct working
(iii)	$P(W \mid F) = \frac{P(W \cap F)}{P(F)} = \frac{0.11}{0.41} = \frac{11}{41} = 0.268$ This is the probability that a randomly selected respondent works (part time), given that the respondent is female.	M1 for correct fraction A1 E1 For E1 must be in context – not just talking about events F and W	**3**	Allow 0.27 with working Allow 11/41 as final answer Condone 'if' or 'when' for 'given that' but not the words 'and' or 'because' or 'due to' for E1. E1 (independent of M1): the order/structure must be correct i.e. no reverse statement Allow 'The probability that a randomly selected female respondent works part time' oe
		TOTAL	**8**	

Q6 (i)	Mean $= \dfrac{1\times10+2\times40+3\times15+4\times5}{70}=\dfrac{155}{70}=2.214$ S_{xx} = $1^2\times10+2^2\times40+3^2\times15+4^2\times5-\dfrac{155^2}{70}=385-343.21=41.79$ $s=\sqrt{\dfrac{41.79}{69}}=0.778$	M1 A1 CAO M1 for Σfx^2 s.o.i. M1 for attempt at S_{xx} Dep on first M1 A1 CAO If 0.778 or better seen ignore previous incorrect working (calculator answer) Allow final answer to 2 sig fig (www)	**5**	For M1 allow sight of at least 3 double pairs seen from $1\times10+2\times40+3\times15+4\times5$ with divisor 70. Allow answer of 155/70 or 2.2 or 2.21 or 31/14 oe For 155/70 = eg 2.3 , allow A1 isw M1 for $1^2\times10+2^2\times40+3^2\times15+4^2\times5$ with at least three correct terms Using exact mean leads to S_{xx} = 41.79 , s=0.778, Using mean 2.214 leads to S_{xx} = 41.87 , s=0.779, Using mean 2.21 leads to S_{xx} = 43.11 and s = 0.790 Using mean 2.2 leads to S_{xx} = 46.2 and s = 0.818 Using mean 2 leads to S_{xx} = 105 and s = 1.233 All the above get M1M1A1 except the last one which gets M1M1A0 RMSD(divisor n rather than n – 1) =√(41.79/70) = 0.772 gets M1M1A0 Alternative method, award M1for at least 3 terms of and second M1 for all 4 terms of $(1-2.214)^2\times10+(2-2.214)^2\times40+(3-2.214)^2\times15$ $+(4-2.214)^2\times5(=41.79)$ NB Allow full credit for correct answers without working (calculator used)
(ii)	Mean would decrease Standard deviation would increase	B1 B1	**2**	Do not accept increase/decrease seen on their own – must be linked to mean and SD. Allow eg 'It would skew the mean towards zero' And eg ' It would stretch the SD' SC1 for justified argument that standard deviation might either increase or decrease according to number with no eggs (n≤496 increase, n≥497 decrease)
		TOTAL	**7**	

	SECTION B			
Q7 (i)	$X \sim \mathrm{B}(20, 0.15)$ (***A***) *Either* $\mathrm{P}(\boldsymbol{X} = 1) = \binom{20}{1} \times 0.15^1 \times 0.85^{19} = 0.1368$ *or* $\mathrm{P}(X = 1) = \mathrm{P}(X \le 1) - \mathrm{P}(X \le 0)$ $= 0.1756 - 0.0388 = 0.1368$	M1 $0.15^1 \times 0.85^{19}$ M1 $\binom{20}{1} \times p^1 q^{19}$ A1 CAO OR: M2 for 0.1756 – 0.0388 A1 CAO	**3**	With $p + q = 1$ Allow answer 0.137 with or without working or 0.14 if correct working shown See tables at the website http://www.mei.org.uk/files/pdf/formula_book_mf2.pdf For misread of tables 0.3917 – 0.1216 = 0.2701 allow M1M1A0 also for 0.1304 – 0.0261 = 0.1043
	(***B***) $\mathrm{P}(X \ge 2) = 1 - \mathrm{P}(X \le 1)$ $= 1 - 0.1756 = 0.8244$	M1 for 1 – their 0.1756 A1 CAO	**2**	Provided 0.1756 comes from $\mathrm{P}(X{=}0) + \mathrm{P}(X{=}1)$ Allow answer 0.824 with or without working or 0.82 if correct working shown Point probability method: P(1) = 0.1368 , P(0) = 0.0388 So 1 - $\mathrm{P}(X \le 1)$ = 1 - 0.1756 gets M1 then mark as per scheme M0A0 for $1 - \mathrm{P}(X \le 1) = 1 - 0.4049 = 0.5951$ For misread of tables 1 – 0.3917 = 0.6083 allow M1A1 also for 1 – 0. 1304 = 0.8696 provided consistent with part (*A*) OR M1A0 if formula used in part (*A*)

(ii)	Let $X \sim B(n, p)$ Let p = probability of a 'no-show' (for population) H_0: $p = 0.15$ H_1: $p < 0.15$ H_1 has this form because the hospital management hopes to reduce the proportion of no-shows.	B1 for definition of p B1 for H_0 B1 for H_1 E1 Allow correct answer even if H_1 wrong	**4**	Allow p = P(no-show) for B1 Definition of p must include word probability (or chance or proportion or percentage or likelihood but NOT possibility). Preferably as a separate comment. However can be at end of H_0 as long as it is a clear definition ' p = the probability of no-show, NOT just a sentence 'probability is 0.15' H_0 : p(no-show) = 0.15, H_1 : p(no-show) < 0.15 gets B0B1B1 Allow p=15%, allow θ or π and ρ but not x. However allow any single symbol <u>if defined</u> Allow $H_0 = p$=0.15, Do not allow $H_0 : P(X=x) = 0.15$, $H_1 : P(X=x) < 0.15$ Do not allow H_0: =0.15, =15%, P(0.15), p(0.15), p(x)=0.15, x=0.15 (unless x correctly defined as a probability) Do not allow H_1:$p \le 0.15$, Do not allow H_0 and H_1 reversed for B marks but can still get E1 Allow NH and AH in place of H_0 and H_1 For hypotheses given in words allow Maximum B0B1B1E1 Hypotheses in words must include probability (or chance or proportion or percentage) and the figure 0.15 oe.
(iii)	$P(X \le 1) = 0.1756 > 5\%$ So not enough evidence to reject H_0. Not significant. Conclude that there is not enough evidence to indicate that the proportion of no-shows has decreased. Note: use of critical region method scores M1 for region {0} M1 for 1 does not lie in critical region, then A1 E1 as per scheme	M1 for probability seen, but not in calculation for point probability M1 dep for comparison A1 E1 dep for conclusion in context.	**4**	Zero for use of point prob - $P(X = 1) = 0.1368$ Do <u>NOT</u> FT wrong H_1 Allow accept H_0, or reject H_1 Full marks only available if 'not enough evidence to…' oe mentioned somewhere Do not allow 'enough evidence to reject H_1' for final mark but can still get 3/4 Upper end comparison: $1 - 0.1756 = 0.8244 < 95\%$ gets M2 then A1E1 as per scheme

				Line diagram method M1 for squiggly line between 0 and 1 with arrow pointing to left, M1 0.0388 seen on diagram from squiggly line or from 0, A1E1 for correct conclusion Bar chart method M1 for line clearly on boundary between 0 and 1 and arrow pointing to left, M1 0.0388 seen on diagram from boundary line or from 0, A1E1 for correct conclusion
(iv)	$6 < 8$ So there is sufficient evidence to reject H_0 Conclude that there is enough evidence to indicate that the proportion of no-shows appears to have decreased.	M1 for comparison seen A1 E1 for conclusion in context	**3**	Allow '6 lies in the CR' Do NOT insist on 'not enough evidence' here Do not FT wrong H_1:p>0.15 but may get M1 In part (iv) ignore any interchanged H_0 and H_1 seen in part (ii)
(v)	For $n \leq 18$, $P(X \leq 0) > 0.05$ so the critical region is empty.	E1 for $P(X \leq 0) > 0.05$ E1 indep for critical region is empty	**2**	E1 also for sight of 0.0536 Condone $P(X = 0) > 0.05$ or all probabilities or values, (but not outcomes) in table (for $n \leq 18$) > 0.05 Or 'There is no critical region' For second E1 accept 'H_0 would always be accepted' Do NOT FT wrong H_1 Use professional judgement - allow other convincing answers
		TOTAL	**18**	

<table>
<tr>
<td>Q8
(i)</td>
<td>
<table>
<tr><td>Upper Bound</td><td>9.1</td><td>9.3</td><td>9.5</td><td>9.7</td><td>9.9</td><td>10.1</td></tr>
<tr><td>Cumulative frequency</td><td>0</td><td>5</td><td>12</td><td>27</td><td>43</td><td>50</td></tr>
</table>


</td>
<td>B1 for cumulative frequencies

G1 for scales

G1 for labels

G1 for points (Provided plotted at correct UCB positions)

G1 for joining points

All G’s dep on attempt at cumulative frequency but not cumulative fx’s or other spurious values.</td>
<td>5</td>
<td>May be implied from graph. Condone omission of 0 at this stage.

Linear horizontal scale.
Linear vertical scale: 0 to 50
(no inequality scales - Not even <9.1, <9.3, <9.5 …)

Heating quality or x and Cumulative frequency or just CF or similar but not just frequency or fd nor cumulative fd

Plotted as (UCB, their cf). Ignore (9.1,0) at this stage. No midpoint or LCB plots.
Plotted within ½ small square

For joining all of ‘their points’ (line or smooth curve) AND now including (9.1,0) dep on previous G1

Mid point or LCB plots may score first three marks

Can get up to 3/5 for cum freq bars
Allow full credit if axes reversed correctly

Lines of best fit could attract max 4 out of 5.</td>
</tr>
</table>

(ii)	Median = 9.67 Q1 = 9.51 Q3 = 9.83 Inter-quartile range = 9.83 – 9.51 = 0.32	B1 FT Allow answers between 9.66 and 9.68 without checking curve. Otherwise check curve. B1 FT for Q3 or Q1 B1 FT for IQR providing both Q1 and Q3 are correct Allow answers between 9.50 and 9.52 and between 9.82 and 9.84 without checking curve. Otherwise check curve.	**3**	Based on 25th to 26th value on a cumulative frequency graph ft their mid-point plot (not LCB's) approx 9.57 for m.p. plot Allow 9.56 to 9.58 without checking B0 for interpolation Based on 12th to 13th and 37th to 38th values on a cumulative frequency graph ft their mid -point plot (not LCB's) approx Q1 = 9.42; Q3 = 9.73 Allow 9.41 to 9.43 and 9.72 to 9.74 without checking B0 for interpolation Allow correct IQR from graph if quartiles not stated Lines of best fit: B0 B0 B0 here.
(iii)	Lower limit 9.51 – 1.5 × 0.32 = 9.03 Upper limit 9.83 + 1.5 × 0.32 = 10.31 Thus there are no outliers in the sample.	B1 FT their Q_1, IQR B1 FT their Q_3, IQR E1 NB E mark dep on both B marks	**3**	Any use of <u>median</u> ± 1.5 IQR scores B0 B0 E0 If FT leads to limits above 9.1 or below 10.1 then E0 No marks for ± 2 or 3 IQR In this part FT their values from (ii) if sensibly obtained (eg from LCB plot) or lines of best fit, but not from location ie 12.5, 37.5 or cumulative fx's or similar. For use of mean ± 2s, Mean = 9.652, s = 0.235, Limits 9.182, 10.122 gets M1 for correct lower limit, M1 for correct upper limit, zero otherwise, but E0 since there could be outliers using this definition

(iv)	(A) P(All 3 more than 9.5) = $\frac{38}{50}\times\frac{37}{49}\times\frac{36}{48}=0.4304$ (=50616/117600 = 2109/4900)	M1 for 38/50 × (triple product) M1 for product of remaining fractions A1 CAO	**3**	$(38/50)^3$ which gives answer 0.4389 scores M1M0A0 so watch for this. M0M0A0 for binomial probability including 0.76^{100} but ${}^3C_0\times0.24^0\times0.76^3$ still scores M1 $(k/50)^3$ for values of k other than 38 scores M0M0A0 $\frac{k}{50}\times\frac{(k-1)}{49}\times\frac{(k-2)}{48}$ for values of k other than 38 scores M1M0A0 Correct working but then multiplied or divided by some factor scores M1M0A0 Accept 0.43 with working and 0.430 without working Or $\binom{38}{3}/\binom{50}{3}$ = 2109/4900 = 0.4304 Allow unsimplified fraction as final answer 50616/117600
	(B) P(At least 2 more than 9.5) $=3\times\frac{38}{50}\times\frac{37}{49}\times\frac{12}{48}+0.4304$ = 3 × 0.1435 + 0.4304 = 0.4304 + 0.4304 = 0.861 (=101232/117600 = 4218/4900 = 2109/2450)	M1 for product of 3 correct fractions seen M1 for 3 × a sensible triple or sum of 3 sensible triples M1 indep for + 0.4304 FT (providing it is a probability) A1 CAO	**4**	Or $\binom{38}{2}\binom{12}{1}/\binom{50}{3}$=0.4304 gets first two M1M1's SC1 for $3\times\frac{38}{50}\times\frac{38}{50}\times\frac{12}{50}$ or other sensible triple and SC2 if this + their 0.4304 (= 0.8549) Allow 0.86 or 2109/2450 or 4218/4900, but only M3A0 for other unsimplified fractions
	OR P(At least 2 more than 9.5) = 1 – (P(0) + P(1)) $=1-\left[\left(\frac{12}{50}\times\frac{11}{49}\times\frac{10}{48}\right)+\left(3\times\frac{12}{50}\times\frac{11}{49}\times\frac{38}{48}\right)\right]$ = 1 – [0.01122 + 0.12796] = 1 – 0.13918 = 0.861	M1 for 12/50×11/49×38/48 M1 for 3 × a sensible triple or sum of 3 sensible triples M1 dep on both previous M1's for 1 –[0.01122 + 0.12796] A1 CAO		Use of 1 – method 'with replacement' SC1 for $3\times\frac{12}{50}\times\frac{12}{50}\times\frac{38}{50}$ SC2 for whole of 1 - $3\times\frac{12}{50}\times\frac{12}{50}\times\frac{38}{50}+\frac{12}{50}\times\frac{12}{50}\times\frac{12}{50}$ (= 1 – (0.1313 + 0.0138) = 1 – 0.1451 = 0.8549)
		TOTAL	**18**	

NOTE RE OVER-SPECIFICATION OF ANSWERS

If answers are grossly over-specified (see instruction 8), deduct the final answer mark in every case. Probabilities should also be rounded to a sensible degree of accuracy. In general final non probability answers should not be given to more than 4 significant figures. Allow probabilities given to 5 sig fig. In general accept answers which are correct to 3 significant figures when given to 4 or 5 significant figures.

OCR (Oxford Cambridge and RSA Examinations)
1 Hills Road
Cambridge
CB1 2EU

OCR Customer Contact Centre

14 – 19 Qualifications (General)
Telephone: 01223 553998
Facsimile: 01223 552627
Email: general.qualifications@ocr.org.uk

www.ocr.org.uk