

**Statistics (MEI)**

Advanced Subsidiary GCE AS H132

**Mark Scheme for the Units**

---

**June 2008**

**H132/MS/R/08J**

OCR (Oxford, Cambridge and RSA Examinations) is a unitary awarding body, established by the University of Cambridge Local Examinations Syndicate and the RSA Examinations Board in January 1998. OCR provides a full range of GCSE, A level, GNVQ, Key Skills and other qualifications for schools and colleges in the United Kingdom, including those previously provided by MEG and OCEAC. It is also responsible for developing new syllabuses to meet national requirements and the needs of students and teachers.

This mark scheme is published as an aid to teachers and students, to indicate the requirements of the examination. It shows the basis on which marks were awarded by Examiners. It does not indicate the details of the discussions which took place at an Examiners' meeting before marking commenced.

All Examiners are instructed that alternative correct answers and unexpected approaches in candidates' scripts must be given marks that fairly reflect the relevant knowledge and skills demonstrated.

Mark schemes should be read in conjunction with the published question papers and the Report on the Examination.

OCR will not enter into any discussion or correspondence in connection with this mark scheme.

© OCR 2008

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](mailto:publications@ocr.org.uk)

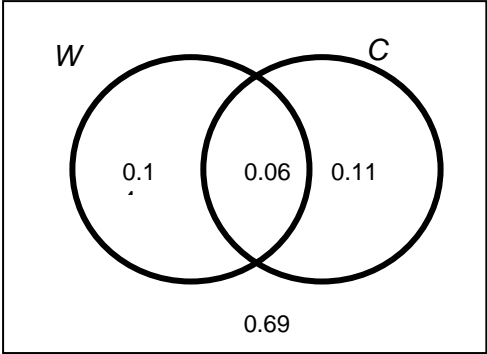
## CONTENTS

### Advanced Subsidiary GCE Statistics (H132)

#### MARK SCHEMES FOR THE UNITS

<b>Unit/Content</b>	<b>Page</b>
G241 Statistics 1	1
G242 Statistics 2	9
G243 Statistics 3	13
Grade Thresholds	17

## G241 Statistics 1

1	(i)	<p>Mean = 7.35 (or better)</p> <p>Standard deviation: 3.69 – 3.70 (awfw)</p> <p>Allow <math>s^2 = 13.62</math> to <math>13.68</math></p> <p>Allow <math>\text{rmsd} = 3.64 - 3.66</math> (awfw)</p> <p><b>After B0, B0 scored</b> then if at least 4 correct mid-points seen or used. {<b>1.5, 4, 6, 8.5, 15</b>}</p> <p>Attempt of their mean = <math>\frac{\sum fx}{44}</math>, with <math>301 \leq fx \leq 346</math> and <math>fx</math> strictly from mid-points not class widths or top/lower boundaries.</p>	<p>B2cao <math>\sum fx = 323.5</math></p> <p>B2cao <math>\sum fx^2 = 2964.25</math></p> <p>(B1) for variance s.o.i.o</p> <p>(B1) for rmsd</p> <p>(B1) mid-points</p> <p>(B1) <math>6.84 \leq \text{mean} \leq 7.86</math></p>	4
	(ii)	<p>Upper limit = <math>7.35 + 2 \times 3.69 = 14.73</math> or ‘their sensible mean’ + <math>2 \times</math> ‘their sensible s.d.’</p> <p>So there could be one or more outliers</p>	<p>M1 ( with s.d. &lt; mean)</p> <p>E1dep on B2, B2 earned and comment</p>	2
			<b>TOTAL</b>	<b>6</b>
2	(i)	<p><math>P(W) \times P(C) = 0.20 \times 0.17 = 0.034</math></p> <p><math>P(W \cap C) = 0.06</math> (given in the question)</p> <p>Not equal so not independent (Allow <math>0.20 \times 0.17 \neq 0.06</math> or <math>\neq p(W \cap C)</math> so not independent).</p>	<p>M1 for multiplying or 0.034 seen</p> <p>A1 (numerical justification needed)</p>	2
	(ii)	<div style="text-align: center;">  <p>The last two G marks are independent of the labels</p> </div>	<p>G1 for two <b>overlapping</b> circles labelled</p> <p>G1 for 0.06 and either 0.14 or 0.11 in the <b>correct places</b></p> <p>G1 for all 4 <b>correct</b> probs in the <b>correct</b> places (<b>including the 0.69</b>)</p> <p>NB No credit for Karnaugh maps here</p>	3
	(iii)	$P(W C) = \frac{P(W \cap C)}{P(C)} = \frac{0.06}{0.17} = \frac{6}{17} = 0.353 \text{ (awrt 0.35)}$	<p>M1 for 0.06 / 0.17</p> <p>A1 cao</p>	2

		<b>(iv)</b> 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	<b>1</b>
			<b>TOTAL</b>	<b>8</b>

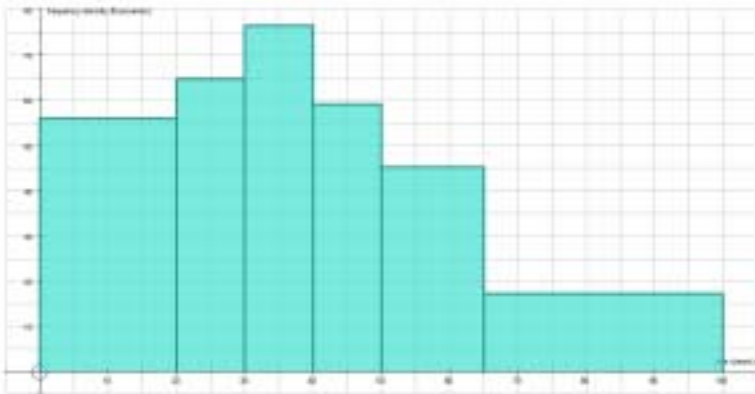
3	(i)	<p>(A) <math>0.5 + 0.35 + p + q = 1</math> so <math>p + q = 0.15</math></p> <p>(B) <math>0 \times 0.5 + 1 \times 0.35 + 2p + 3q = 0.67</math> so <math>2p + 3q = 0.32</math></p> <p>(C) from above <math>2p + 2q = 0.30</math> so <math>q = 0.02, p = 0.13</math></p>	<p>B1 <math>p + q</math> in a correct equation before they reach <math>p + q = 0.15</math></p> <p>B1 <math>2p + 3q</math> in a correct equation before they reach <math>2p + 3q = 0.32</math></p> <p>(B1) for any 1 correct answer B2 for both correct answers</p>	<p>1</p> <p>1</p> <p>2</p>
	(ii)	<p><math>E(X^2) = 0 \times 0.5 + 1 \times 0.35 + 4 \times 0.13 + 9 \times 0.02 = 1.05</math></p> <p><math>\text{Var}(X) = \text{'their } 1.05\text{' } - 0.67^2 = 0.6011</math> (awrt 0.6)</p> <p>(M1, M1 can be earned with their <math>p^+</math> and <math>q^+</math> but not A mark)</p>	<p>M1 <math>\sum x^2 p</math> (at least 2 non zero terms correct) M1dep for <math>(- 0.67^2)</math>, provided <math>\text{Var}(X) &gt; 0</math> A1 cao (No n or n-1 divisors)</p>	3
			<b>TOTAL</b>	<b>7</b>
4	(i)	<p><math>X \sim B(8, 0.05)</math></p> <p>(A) <math>P(X = 0) = 0.95^8 = 0.6634</math>      0.663 or better</p> <p>Or using tables <math>P(X = 0) = 0.6634</math></p> <p>(B) <math>P(X = 1) = \binom{8}{1} \times 0.05 \times 0.95^7 = 0.2793</math></p> <p><math>P(X &gt; 1) = 1 - (0.6634 + 0.2793) = 0.0573</math></p> <p>Or using tables <math>P(X &gt; 1) = 1 - 0.9428 = 0.0572</math></p>	<p>M1 <math>0.95^8</math> A1 CAO Or B2 (tables)</p> <p>M1 for <math>P(X = 1)</math> (allow 0.28 or better) M1 for <math>1 - P(X \leq 1)</math> must have both probabilities A1 cao (0.0572 – 0.0573)</p> <p>M1 for <math>P(X \leq 1)</math> 0.9428 M1 for <math>1 - P(X \leq 1)</math> A1 cao (must end in...2)</p>	<p>2</p> <p>3</p>
	(ii)	<p>Expected number of days = <math>250 \times 0.0572 = 14.3</math> awrt</p>	<p>M1 for <math>250 \times \text{prob}(B)</math> A1 FT but no rounding at end</p>	2
			<b>TOTAL</b>	<b>7</b>
5	(i)	<p>Let <math>p</math> = probability of remembering or naming all items (for population) (whilst listening to music.) <math>H_0: p = 0.35</math> <math>H_1: p &gt; 0.35</math></p> <p><math>H_1</math> has this form since the student believes that the</p>	<p>B1 for definition of <math>p</math> B1 for <math>H_0</math> B1 for <math>H_1</math></p> <p>E1dep on <math>p &gt; 0.35</math> in</p>	

		probability will be increased/ improved/ got better /gone up.	$H_0$ In words not just because $p > 0.35$	4
	(ii)	<p>Let <math>X \sim B(15, 0.35)</math>  <b>Either:</b> <math>P(X \geq 8) = 1 - 0.8868 = 0.1132 &gt; 5\%</math>                      Or <math>0.8868 &lt; 95\%</math></p> <p>So not enough evidence to reject <math>H_0</math> (Accept <math>H_0</math>)</p> <p>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.)</p> <p>-----</p> <p><b>Or:</b></p> <p>Critical region for the test is <math>\{9,10,11,12,13,14,15\}</math>                      8 does not lie in the critical region.</p> <p>So not enough evidence to reject <math>H_0</math></p> <p>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.)</p> <p>-----</p> <p><b>Or:</b></p> <p>The smallest critical region that 8 could fall into is <math>\{8, 9, 10, 11, 12, 13, 14, \text{ and } 15\}</math>. The size of this region is 0.1132</p> <p><math>0.1132 &gt; 5\%</math></p> <p>So not enough evidence to reject <math>H_0</math></p> <p>Conclude that there is not enough evidence to indicate that the probability of remembering all of the items is improved (when listening to music)</p>	<p><b>Either:</b>                      M1 for probability (0.1132)                      M1dep for comparison</p> <p>A1dep</p> <p>E1dep on all previous marks for conclusion in context</p> <p>-----</p> <p><b>Or:</b></p> <p>M1 for correct CR(no omissions or additions)                      M1dep for 8 does not lie in CR                      A1dep</p> <p>E1dep on all previous marks for conclusion in context</p> <p>-----</p> <p><b>Or:</b></p> <p>M1 for CR <math>\{8,9,\dots,15\}</math> and size = 0.1132                      M1 dep for comparison</p> <p>A1dep</p> <p>E1dep on all previous marks for conclusion in context</p>	4
			<b>TOTAL</b>	<b>8</b>

<b>Section B</b>				
<b>6</b>	<b>(i)</b>	<p>(A) <math>P(\text{both rest of UK}) = 0.20 \times 0.20</math>  <math>= 0.04</math></p> <p>(B) <i>Either: All 5 case</i>  <math>P(\text{at least one England}) =</math>  <math>(0.79 \times 0.20) + (0.79 \times 0.01) + (0.20 \times 0.79) + (0.01 \times 0.79) + (0.79 \times 0.79)</math>  <math>= 0.158 + 0.0079 + 0.158 + 0.0079 + 0.6241 = 0.9559</math></p> <p><i>Or</i></p> <p><math>P(\text{at least one England}) = 1 - P(\text{neither England})</math>  <math>= 1 - (0.21 \times 0.21) = 1 - 0.0441 = 0.9559</math>                      or listing all  <math>= 1 - \{ (0.2 \times 0.2) + (0.2 \times 0.01) + (0.01 \times 0.20) + (0.01 \times 0.01) \}</math>  <math>= 1 - (**)</math>  <math>= 1 - \{ 0.04 + 0.002 + 0.002 + 0.0001 \}</math>  <math>= 1 - 0.0441</math>  <math>= 0.9559</math></p> <p><i>Or: All 3 case</i>  <math>P(\text{at least one England}) =</math>  <math>= 0.79 \times 0.21 + 0.21 \times 0.79 + 0.79^2</math>  <math>= 0.1659 + 0.1659 + 0.6241</math>  <math>= 0.9559</math></p> <hr style="border-top: 1px dashed black;"/> <p>(C) <i>Either</i>  <math>0.79 \times 0.79 + 0.79 \times 0.2 + 0.2 \times 0.79 + 0.2 \times 0.2 = 0.9801</math></p> <p><i>Or</i>  <math>0.99 \times 0.99 = 0.9801</math></p> <p><i>Or</i>  <math>1 - \{ 0.79 \times 0.01 + 0.2 \times 0.01 + 0.01 \times 0.79 + 0.01 \times 0.02 + 0.01^2 \} = 1 - 0.0199</math>  <math>= 0.9801</math></p>	<p>M1 for multiplying A1cao</p> <p>M1 for any correct term (3case or 5case) M1 for correct sum of all 3 (or of all 5) with <b>no extras</b> A1cao (condone 0.96 www)</p> <p><i>Or</i> M1 for <math>0.21 \times 0.21</math> or for (***) fully enumerated or 0.0441 seen M1dep for <math>1 - (1^{\text{st}} \text{ part})</math> A1cao</p> <p>See above for 3 case</p> <hr style="border-top: 1px dashed black;"/> <p>M1 for sight of all 4 correct terms summed A1cao (condone 0.98 www) <i>or</i> M1 for <math>0.99 \times 0.99</math> A1cao <b>Or</b> M1 for everything <math>1 - \{ \dots \}</math> A1cao</p>	<p>2</p> <p>3</p> <p>2</p>





7	(i)	Positive	B1	1
	(ii)	<p>Number of people = <math>20 \times 33 (000) + 5 \times 58 (000)</math>  <math>= 660 (000) + 290 (000) = 950 000</math></p>	<p>M1 first term  M1(indep)  second term  A1 cao  NB answer of  950 scores  M2A0</p>	3
	(iii)	<p>(A) <math>a = 1810 + 340 = 2150</math>  (B) Median = age of 1 385 (000<sup>th</sup>) person or 1385.5 (000)  Age 30, cf = 1 240 (000); age 40, cf = 1 810 (000)  Estimate median = <math>(30) + \frac{145}{570} \times 10</math>    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  <b>no previous working</b> allow SC1</p>	<p>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  <math>\frac{145k}{570k} \times 10</math>  (2.54 or better  suggests this)  A1 cao min 1dp</p>	2  3
	(iv)	<p>Frequency densities: 56, 65, 77, 59, 45, 17  <i>(accept 45.33 and 17.43 for 45 and 17)</i></p> 	<p>B1 for any one  correct  B1 for all correct  (soi by listing or  from histogram)    <b>Note: all G  marks below  dep on attempt  at frequency  density, NOT  frequency</b>    G1 Linear scales  on both axes (no  inequalities)  G1 Heights FT  their listed fds or  all must be  correct. <b>Also  widths. All  blocks joined</b>    G1 Appropriate  label for vertical</p>	5

			scale eg 'Frequency density (thousands)', 'frequency (thousands) per 10 years', 'thousands of people per 10 years'. (allow key). <b>OR f.d.</b>	
	(v)	<p>Any two suitable comments such as:</p> <p>Outer London has a greater proportion (or %) of people under 20 (or almost equal proportion)</p> <p>The modal group in Inner London is 20-30 but in Outer London it is 30-40</p> <p>Outer London has a greater proportion (14%) of aged 65+</p> <p><b>All</b> populations in <b>each</b> age group are higher in Outer London</p> <p>Outer London has a more evenly spread distribution or balanced distribution (ages) o.e.</p>	<p>E1</p> <p>E1</p>	2
	(vi)	<p>Mean increase ↑ median unchanged (-) midrange increase ↑</p> <p>standard deviation increase ↑ interquartile range unchanged. (-)</p>	<p>Any one correct B1 Any two correct B2 Any three correct B3 All <b>five</b> correct B4</p>	4
			<b>TOTAL</b>	<b>20</b>

## G242 Statistics 2

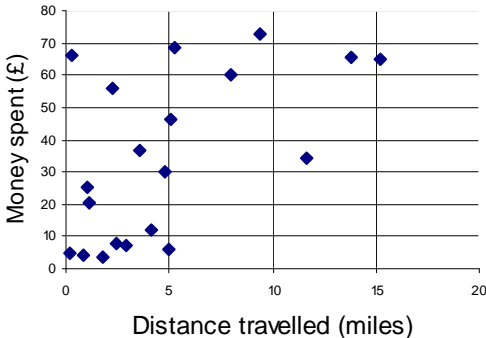
1		(i)	Scots pine seedlings occur randomly and independently with uniform mean rate.	B1 B1	2
1	A	(ii)	$e^{-8} \times 8^7 \div 7!$ (0.4530-03134 from tables)	M1	2
	B		0.1396 $1 - P(X \leq 7)$ $1 - 0.4530$ 0.547	A1 M1 A1	
1		(iii)	$(1 - (ii)B)^5$ [ $= 0.453^5$ ] $1 - 0.453^5$ 0.9809	M1 M1 A1	3
1		(iv)	$P(\text{height} > 70) = P(Z > \frac{70 - 56}{20}) = P(Z > 0.7)$ $= 1 - \Phi(0.7) = 1 - 0.7580 (= 0.242 \text{ (answer given)})$	M1 A1	2
1		(v)	$z = 1.645$ $-1.645 \times 20 + 56$ 23.1	B1 ( $\pm$ ) M1 (-ive z) A1	3
2		(i)	Mean = 185 Variance = 210.727	B1 B1	2
2		(ii)	$H_0 : \mu = 175$ & $H_1 : \mu > 175$ Where $\mu$ represents the mean decrease for the underlying population. $t = \frac{185 - 175}{SD / \sqrt{12}} = 2.39$ (3s.f.) 11 degrees of freedom At 5% level, critical value of $t$ is 1.796 $2.39 > 1.796$ so the result is significant. Evidence suggests the mean decrease in cholesterol level is more than 175.	B1 B1 M1 A1 FT B1 B1 M1A1 A1	9
2		(iii)	The decrease in cholesterol level in the underlying population follows a Normal distribution. The sample is assumed to be random.	E1(Normal) E1(Random)	2

3	(i)	<p><math>H_0</math>: there is no association between gum disease and coronary artery disease  <math>H_1</math>: there is an association between gum disease and coronary artery disease</p> <p>Expected frequencies</p> <table border="1" data-bbox="427 450 1002 566"> <thead> <tr> <th></th> <th>With g.d.</th> <th>Without g.d.</th> </tr> </thead> <tbody> <tr> <td>With c.a.d</td> <td>48.62</td> <td>36.38</td> </tr> <tr> <td>Without c.a.d.</td> <td>94.38</td> <td>70.62</td> </tr> </tbody> </table> <p>Contributions to <math>X^2</math> (without Yates' correction)</p> <table border="1" data-bbox="427 674 1002 790"> <thead> <tr> <th></th> <th>With g.d.</th> <th>Without g.d.</th> </tr> </thead> <tbody> <tr> <td>With c.a.d</td> <td>1.52827</td> <td>2.04245</td> </tr> <tr> <td>Without c.a.d.</td> <td>0.78729</td> <td>1.05217</td> </tr> </tbody> </table> <p>Contributions to <math>X^2</math> (with Yates' correction)</p> <table border="1" data-bbox="427 898 1002 1014"> <thead> <tr> <th></th> <th>With g.d.</th> <th>Without g.d.</th> </tr> </thead> <tbody> <tr> <td>With c.a.d</td> <td>1.35612</td> <td>1.81238</td> </tr> <tr> <td>Without c.a.d.</td> <td>0.69861</td> <td>0.93365</td> </tr> </tbody> </table> <p><math>X^2 = 5.4102</math> (or 4.8008 with Yates' correction)  1 degree of freedom  Critical value at 5% level is 3.841  As 5.4102 (or 4.8008) &gt; 3.841 the result is significant  There is evidence of an association between gum disease and coronary artery disease</p>		With g.d.	Without g.d.	With c.a.d	48.62	36.38	Without c.a.d.	94.38	70.62		With g.d.	Without g.d.	With c.a.d	1.52827	2.04245	Without c.a.d.	0.78729	1.05217		With g.d.	Without g.d.	With c.a.d	1.35612	1.81238	Without c.a.d.	0.69861	0.93365	<p>B1</p> <p>M1 A1</p> <p>M1 M1 (summation)</p> <p>A1 CAO B1 B1 M1 A1(in context)</p>	<p><b>10</b></p>
	With g.d.	Without g.d.																													
With c.a.d	48.62	36.38																													
Without c.a.d.	94.38	70.62																													
	With g.d.	Without g.d.																													
With c.a.d	1.52827	2.04245																													
Without c.a.d.	0.78729	1.05217																													
	With g.d.	Without g.d.																													
With c.a.d	1.35612	1.81238																													
Without c.a.d.	0.69861	0.93365																													
3	(ii)	<p>2 degrees of freedom  Critical value at 5% level is 5.991  8.2808 &gt; 5.991 the result is significant  There is evidence of an association between age and coronary artery disease</p>	<p>B1 B1 M1 A1(in context)</p>	<p><b>4</b></p>																											
3	(iii)	<p>Suitable comments (in context)</p>	<p>E1 E1</p>	<p><b>2</b></p>																											

4	(i)	<p><math>H_0</math>: population median = 26  <math>H_1</math>: population median &lt; 26</p> <p>Actual differences  -3 17 -10 -11 -4 7 1 -15 -8 -2 -16 -6</p> <p>Associated ranks  3 12 8 9 4 6 1 10 7 2 11 5</p> <p><math>T = 3 + 8 + 9 + 4 + 10 + 7 + 2 + 11 + 5 = 59</math></p> <p><math>T^+ = 12 + 6 + 1 = 19</math></p> <p><math>\therefore T = 19</math></p> <p>From tables – at the 5% level of significance in a one-tailed Wilcoxon signed rank test, the critical value of <math>T</math> is 17</p> <p><math>19 &gt; 17 \therefore</math> the result is not significant</p> <p>The evidence does not suggest a decrease in the numbers of ants this year.</p>	<p>B1  B1</p> <p>B1</p> <p>M1 A1</p> <p>B1</p> <p>B1  B1</p> <p>B1</p> <p>M1 A1</p> <p>E1</p>	<p><b>12</b></p>
4	(ii)	<p>Variable - symmetry  Sample - random</p>	<p>E1  E1</p>	<p><b>2</b></p>

5	(i)	If hamsters choose their bedding randomly then the three options will be equally likely to occur. Hence the probability that a hamster chooses the new material will be $\frac{1}{3}$ .	E1	<b>1</b>
5	(ii)	0.0390, 0.1561, 0.2731, 0.2732, 0.1707, 0.0682, 0.0197 5.85, 23.415, 40.965, 40.98, 25.605, 10.23, 2.955	M1 A2 M1 A1	<b>5</b>
5	(iii)	<p><math>H_0</math>: B(8, <math>\frac{1}{3}</math>) is a good model  <math>H_1</math>: B(8, <math>\frac{1}{3}</math>) is not a good model</p> <p>Merge final two cells.</p> $X^2 = \frac{(10 - 5.85)^2}{5.85} + \frac{(31 - 23.415)^2}{23.415} + \frac{(42 - 40.965)^2}{40.965} + \frac{(34 - 40.98)^2}{40.98} + \frac{(19 - 25.605)^2}{25.605} + \frac{(14 - 13.185)^2}{13.185}$ <p>= 8.370  5 degrees of freedom (6-1)  Critical value at 5% level is 11.07 (FT their dof)  8.370 &lt; 11.07 the result is not significant.  No evidence to suggest the binomial distribution is not a good model.</p>	<p>B1</p> <p>M1  M1 [(o-e)<sup>2</sup>/e]  M1 [sum]</p> <p>A1  B1  B1  M1  A1</p>	<b>9</b>

# G243 Statistics 3

1		(i)	<p>Quota sampling.  <b>Advantage</b> – probably the only realistic way to get a reasonably ‘representative’ sample in these circumstances.  <b>Disadvantage</b> – non-random, so statistical analysis is complicated.</p>	B1 E1 E1	<p>Or other sensible comments.                  Eg cost or time effective as an advantage</p>	3
	(a)	(ii)		G1 G1 G1	<p>Axes, including labels.                  Correct zero.                  All points correct (allow 2 errors).</p>	3
	(b)		<p>Ranks:                  1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20                  3 18 1 9 8 2 14 6 5 12 7 10 4 13 19 15 20 11 17 16  <math> d </math> 2 16 2 5 3 4 7 2 4 2 4 2 9 1 4 1 3 7 2 4</p> $r_s = 1 - \frac{6 \times 584}{20 \times 399} = 1 - 0.4391 = 0.5609$ <p>Critical point for <math>n=20</math> at two-sided 5% level is 0.4466                  Significant.                  Seems there is an association between distance travelled and money spent.</p>	B2 M1 A1 B1 E1 E1	<p>Allow B1 if one or two errors.                  CAO.                  No FT if wrong.                  No access to these marks if value of <math>r_s</math> is nonsense.</p>	7
	(c)		<p>Some sensible explanation of “no”.                  Scatter diagram does not suggest bivariate Normality.</p>	M1 A1	<p>SC1. Allow 1 out of 2 if “bivariate” missing.                  SC2. Allow 1 out of 2 for sensible comment re “outliers”.                  No marks for “data not linear”.</p>	2
2						
		(i)	<p><math>H_0 : \mu_A = \mu_B</math>  <math>H_1 : \mu_A \neq \mu_B</math></p> <p>Where <math>\mu_A, \mu_B</math> are the population mean lengths for the machines.</p>	B1 B1 B1	<p>Do not allow <math>\bar{A}, \bar{B}</math> or similar unless they are clearly and explicitly stated to be <u>population</u> means.                  Hypotheses in words must include “population”.</p>	



		$\left. \begin{aligned} \bar{x}_1 &= \frac{184.5}{90} = 2.05 \\ \bar{x}_2 &= \frac{156.0}{75} = 2.08 \end{aligned} \right\}$ $\left. \begin{aligned} s_1^2 &= \frac{1}{89} \left( 396.94 - \frac{184.5^2}{90} \right) = \frac{18.715}{89} = 0.2103 \\ s_2^2 &= \frac{1}{74} \left( 334.19 - \frac{156.0^2}{75} \right) = \frac{9.71}{74} = 0.1312 \end{aligned} \right\}$ <p>Because the samples are large, the values of <math>s_1^2</math> and <math>s_2^2</math> are taken as <math>\sigma_1^2</math> and <math>\sigma_2^2</math>.</p> <p>Two-sample test based on N(0,1).</p> <p>Test statistic is:</p> $\frac{2.05 - 2.08}{\sqrt{\frac{0.2103}{90} + \frac{0.1312}{75}}} = -\frac{0.03}{\sqrt{0.004086}} = -\frac{0.03}{0.0639} = -0.46 \text{ (93)}$ <p>Double-tailed 10% point of N(0,1) is 1.645. Not significant. No reason to suppose mean lengths differ.</p>	<p>B1</p> <p>M1</p> <p>A1</p> <p>M1</p> <p>M1</p> <p>M1</p> <p>A1</p> <p>A1</p> <p>E1</p> <p>E1</p>	<p>For adequate verbal definition. Allow absence of “population” here if correct notation <math>\mu</math> has been used.</p> <p>M0 A0 for divisor <math>n</math>, but FT.</p> <p>Accept as implicit if <math>s_1^2</math> and <math>s_2^2</math> are <u>correctly</u> used in sequel.</p> <p>Accept usual alternatives.</p> <p>No FT if wrong.</p>	<p>13</p>
	(ii)	<p>Samples are large, so by the Central Limit Theorem the underlying distribution of the sample means will be approximately Normal.</p>	E2	(2, 1, 0)	2
3					
	(i)	<p>Differences are:</p> <p style="text-align: center;">14 16 6 -8 -24 12 2 -15 10 -1</p> <p>Ranks of <math> d </math> 7 9 3 4 10 6 2 8 5 1</p> <p>Test statistic is 4+10+8+1=23 (or 7+9+3+6+2+5=32)</p> <p>Refer to paired Wilcoxon table with <math>n=10</math>.</p> <p>Need lower 2½% point which is 8 (or, if 32 used, upper 2½% point which is 47).</p> <p>Not significant.</p> <p>Seems underlying median total journey times may be assumed equal.</p>	<p>B1</p> <p>M1</p> <p>A1</p> <p>M1</p> <p>A1</p> <p>M1</p> <p>A1</p> <p>E1</p> <p>E1</p>	<p>FT if M1 earned or if d (not  d ) ranked.</p> <p>FT if M1 earned.</p> <p>No FT if wrong.</p> <p>No FT if wrong.</p>	9
	(ii)	<p>The “pairing” will eliminate any differences between the weeks - so can compare the two airlines.</p>	E1 <p>E1</p>		2

		(iii)	Two sensible comments such as: - check-in and waiting times not in airlines' control - time for collecting luggage not in airlines' control - other journey criteria might be of importance (e.g. departure time, on-board service, fares).	E2 E2	Reward any two sensible comments for (E2 each) (2,1,0). For 2 marks there must be some comment to the effect of <u>comparison</u> , not merely that a factor might affect both airlines.	4
4						
		(i)	Randomisation: to guard against possible unsuspected sources of bias - caused by fertility patterns among the plots.	E1 E1	Or equivalent comments.	2
		(ii)	Replication: so that natural variation can be measured, so that any observed inter-variety variation can be compared with it.	E1 E1	Or equivalent comments.	2
		(iii)	Normality of <u>both populations</u> , equal <u>population</u> variances.  $H_0 : \mu_A = \mu_B$ $H_1 : \mu_A \neq \mu_B$  where $\mu_A, \mu_B$ are the population means for varieties A and B.  A : $\bar{x} = 23.50$ , $s_{n-1} = 0.9529$ B : $\bar{y} = 21.94$ , $s_{n-1} = 0.8649$  Pooled $s^2 = \frac{(5 \times 0.908) + (4 \times 0.748)}{9} = \frac{7.532}{9} = 0.8368$  Test statistic is $\frac{23.50 - 21.94(-0)}{\sqrt{0.8368} \sqrt{\frac{1}{6} + \frac{1}{5}}}$  $= \frac{1.56}{0.5539(5)} = 2.816$  Refer to $t_9$ . Double-tailed 5% point is 2.262.	B1 B1 B1 B1  B1  M1 A1  M1 M1 M1 A1	Do not allow $\bar{A}, \bar{B}$ or similar unless they are clearly and explicitly stated to be <u>population</u> means. Hypotheses in words must include "population".  Do not allow $s_n$ : 0.8699, 0.7736  For any reasonable attempt at pooling. If correct.  For numerator. For $\sqrt{0.8368}$ (or cand's value). For $\sqrt{\frac{1}{6} + \frac{1}{5}}$ .  FT from here if all M marks earned.	

		Significant. Appears that population mean yields are different.	M1 A1 E1 E1	No FT if wrong. [accept usual No FT if wrong. [alternatives.]	15
	(iv)	The pairing will eliminate differences around the field. - can compare the plots within the pairs.	E1 E1		2
	(v)	Refer to $t_4$ . Single-tailed 5% point is 2.132. Not significant.  No evidence to reject $H_0$ that population mean yields of A and V are the same. Normality of underlying population of differences.	M1 A1 E1  E1 B1 B1	No FT if wrong. No FT if wrong.	6

# Grade Thresholds

Advanced Subsidiary GCE Statistics MEI (H132)  
June 2008 Examination Series

## Unit Threshold Marks

Unit		Maximum Mark	a	b	c	d	e	u
G241	Raw	72	53	45	38	31	24	0
	UMS	100	80	70	60	50	40	0
G242	Raw	72	56	49	42	35	28	0
	UMS	100	80	70	60	50	40	0
G243	Raw	72	56	48	40	33	26	0
	UMS	100	80	70	60	50	40	0

## Specification Aggregation Results

Overall threshold marks in UMS (ie after conversion of raw marks to uniform marks)

	Maximum Mark	A	B	C	D	E	U
H132	300	240	210	180	150	120	0

The cumulative percentage of candidates awarded each grade was as follows:

	A	B	C	D	E	U	Total Number of Candidates
H132	9.7	12.9	35.5	51.6	64.5	100	31

For a description of how UMS marks are calculated see:  
[http://www.ocr.org.uk/learners/ums\\_results.html](http://www.ocr.org.uk/learners/ums_results.html)

Statistics are correct at the time of publication.

**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](mailto:general.qualifications@ocr.org.uk)

**[www.ocr.org.uk](http://www.ocr.org.uk)**

For staff training purposes and as part of our quality assurance programme your call may be recorded or monitored

**Oxford Cambridge and RSA Examinations**  
is a Company Limited by Guarantee  
Registered in England  
Registered Office; 1 Hills Road, Cambridge, CB1 2EU  
Registered Company Number: 3484466  
OCR is an exempt Charity

**OCR (Oxford Cambridge and RSA Examinations)**  
Head office  
Telephone: 01223 552552  
Facsimile: 01223 552553

© OCR 2008

