



**General Certificate of Education (A-level)
June 2013**

Statistics

SS03

(Specification 6380)

Statistics 3

Final

Mark Scheme

Mark schemes are prepared by the Principal Examiner and considered, together with the relevant questions, by a panel of subject teachers. This mark scheme includes any amendments made at the standardisation events which all examiners participate in and is the scheme which was used by them in this examination. The standardisation process ensures that the mark scheme covers the students' responses to questions and that every examiner understands and applies it in the same correct way. As preparation for standardisation each examiner analyses a number of students' scripts: alternative answers not already covered by the mark scheme are discussed and legislated for. If, after the standardisation process, examiners encounter unusual answers which have not been raised they are required to refer these to the Principal Examiner.

It must be stressed that a mark scheme is a working document, in many cases further developed and expanded on the basis of students' reactions to a particular paper. Assumptions about future mark schemes on the basis of one year's document should be avoided; whilst the guiding principles of assessment remain constant, details will change, depending on the content of a particular examination paper.

Further copies of this Mark Scheme are available from: aqa.org.uk

Copyright © 2013 AQA and its licensors. All rights reserved.

Copyright

AQA retains the copyright on all its publications. However, registered schools/colleges for AQA are permitted to copy material from this booklet for their own internal use, with the following important exception: AQA cannot give permission to schools/colleges to photocopy any material that is acknowledged to a third party even for internal use within the centre.

Set and published by the Assessment and Qualifications Alliance.

Key to mark scheme abbreviations

M	mark is for method
m or dM	mark is dependent on one or more M marks and is for method
A	mark is dependent on M or m marks and is for accuracy
B	mark is independent of M or m marks and is for method and accuracy
E	mark is for explanation
✓ or ft or F	follow through from previous incorrect result
CAO	correct answer only
CSO	correct solution only
AWFW	anything which falls within
AWRT	anything which rounds to
ACF	any correct form
AG	answer given
SC	special case
OE	or equivalent
A2,1	2 or 1 (or 0) accuracy marks
-x EE	deduct x marks for each error
NMS	no method shown
PI	possibly implied
SCA	substantially correct approach
c	candidate
sf	significant figure(s)
dp	decimal place(s)

No Method Shown

Where the question specifically requires a particular method to be used, we must usually see evidence of use of this method for any marks to be awarded.

Where the answer can be reasonably obtained without showing working and it is very unlikely that the correct answer can be obtained by using an incorrect method, we must award **full marks**. However, the obvious penalty to candidates showing no working is that incorrect answers, however close, earn **no marks**.

Where a question asks the candidate to state or write down a result, no method need be shown for full marks.

Where the permitted calculator has functions which reasonably allow the solution of the question directly, the correct answer without working earns **full marks**, unless it is given to less than the degree of accuracy accepted in the mark scheme, when it gains **no marks**.

Otherwise we require evidence of a correct method for any marks to be awarded.

Q	Solution	Marks	Total	Comments
1(a)	<p>H_0: Population median time = 32 H_1: Population median time < 32 1 tail test 10% level signs + - - - - + + + - - - -</p> <p>test stat = 8 -/ 4+</p> <p>Bin (12, 0.5) model</p> <p>$P(\leq 4+) = 0.194 > 0.10$</p> <p>Accept H_0. No significant evidence to suggest median time to complete crossword has decreased.</p>	<p>B1</p> <p>M1</p> <p>A1</p> <p>M1</p> <p>M1</p> <p>A1</p>	<p>6</p>	<p>must mention population</p> <p>for signs or signed differences</p> <p>for test stat 8 or 4</p> <p>for use of Bin model</p> <p>for comparison ts, 0.193-0.194, and 10%</p> <p>ts/cv correct</p> <p>Alt method Use of cr {0, 1, 2, 3} or {9, 10, 11, 12} with prob 0.073 used.</p>
	Total		6	

Q	Solution	Marks	Total	Comments																		
2(a)	H_0 : Population average difference = 0 H_1 : Population average difference \neq 0 2 tail test 5% level	B1		May refer to mean/median																		
	<table border="1"> <thead> <tr> <th></th> <th>A</th> <th>B</th> <th>C</th> <th>D</th> <th>E</th> </tr> </thead> <tbody> <tr> <td>diff</td> <td>+3.3</td> <td>.</td> <td>-0.8</td> <td>-0.9</td> <td>+0.3</td> </tr> <tr> <td>rank</td> <td>9</td> <td></td> <td>5</td> <td>6</td> <td>2</td> </tr> </tbody> </table>		A	B	C	D	E	diff	+3.3	.	-0.8	-0.9	+0.3	rank	9		5	6	2	M1 m1		For differences Ranks: smallest rank 1
		A	B	C	D	E																
	diff	+3.3	.	-0.8	-0.9	+0.3																
	rank	9		5	6	2																
	<table border="1"> <thead> <tr> <th></th> <th>F</th> <th>G</th> <th>H</th> <th>I</th> <th>J</th> </tr> </thead> <tbody> <tr> <td>diff</td> <td>+0.7</td> <td>+0.4</td> <td>+1.7</td> <td>-0.1</td> <td>+1.1</td> </tr> <tr> <td>rank</td> <td>4</td> <td>3</td> <td>8</td> <td>1</td> <td>7</td> </tr> </tbody> </table>		F	G	H	I	J	diff	+0.7	+0.4	+1.7	-0.1	+1.1	rank	4	3	8	1	7	m1 A1		Total of ranks (any) One correct
		F	G	H	I	J																
	diff	+0.7	+0.4	+1.7	-0.1	+1.1																
	rank	4	3	8	1	7																
	$T_+ = 9 + 2 + 4 + 3 + 8 + 7 = 33$ $T_- = 5 + 6 + 1 = 12$ test stat $T = 12$ critical value = 6	B1		For cv																		
test stat > 6	m1		ft (must be positive ts) 'correct' T with cv comparison (smaller T / smaller cv larger T / larger cv)																			
Accept H_0 There is no significant evidence of a difference in average percentage of total expenditure spent on 'Highways' between 2002 and 2012.	A1		ts/cv correct																			
	E1	9	In context – only if conclusion correct																			
(b)(i) <u>Differences</u> are symmetrically distributed.	B1																					
(ii) A <u>paired</u> sign test	B1	2																				
(c)(i) 0	B1																					
(ii) $1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 = 55$	M1A1	3																				
	Total		14																			

Q	Solution	Marks	Total	Comments																												
3(a)(i)	H_0 : Samples are taken from identical populations	B1		Or equivalent referring to <u>population medians</u>																												
	H_1 : Samples are not taken from identical populations (average emissions lower when device fitted) 1 tail 5%																															
	Ranks	M1		Ranks as one group either way																												
	<table border="1"> <thead> <tr> <th colspan="2">New device fitted</th> <th colspan="2">Device not fitted</th> </tr> </thead> <tbody> <tr> <td>139.1</td> <td>5 8</td> <td>145.4</td> <td>12 1</td> </tr> <tr> <td>134.6</td> <td>3 10</td> <td>144.0</td> <td>11 2</td> </tr> <tr> <td>128.9</td> <td>1 12</td> <td>138.7</td> <td>4 9</td> </tr> <tr> <td>139.8</td> <td>8 5</td> <td>139.7</td> <td>7 6</td> </tr> <tr> <td>129.5</td> <td>2 11</td> <td>139.6</td> <td>6 7</td> </tr> <tr> <td>140.9</td> <td>10 3</td> <td>140.5</td> <td>9 4</td> </tr> </tbody> </table>	New device fitted		Device not fitted		139.1	5 8	145.4	12 1	134.6	3 10	144.0	11 2	128.9	1 12	138.7	4 9	139.8	8 5	139.7	7 6	129.5	2 11	139.6	6 7	140.9	10 3	140.5	9 4	A1		Ranks correct
	New device fitted		Device not fitted																													
	139.1	5 8	145.4	12 1																												
	134.6	3 10	144.0	11 2																												
	128.9	1 12	138.7	4 9																												
	139.8	8 5	139.7	7 6																												
	129.5	2 11	139.6	6 7																												
140.9	10 3	140.5	9 4																													
$T_{\text{fitted}} = 29$ 49 $T_{\text{not}} = 49$ 29 $n_{\text{fitted}} = 6$ $n_{\text{not}} = 6$	m1		Ranks totalled or reversed																													
$U_{\text{fitted}} = 29 - \frac{6 \times 7}{2} = 8$	m1		Attempt to find U																													
$U_{\text{not}} = 49 - \frac{6 \times 7}{2} = 28$	A1		Either U correct																													
$U = 8$																																
$cv = 7$ for $n = 6, m = 6$ 1 tail 5%	B1		cv correct cao																													
$U > 7$ (or $cv = 29$ and $28 < 29$)	M1		correct comparison ft – upper ts /29 lower ts /7																													
Accept H_0	A1		ts/cv correct																													
No significant evidence of a reduction in average CO ₂ emissions for cars fitted with new device.	E1	10	In context – only if conclusion correct																													
(ii) The 12 new cars can be regarded as a random sample. OR They were randomly selected.	B1	1	Disallow ‘random’ no context																													
(b) A Type II error would be to conclude that H_0 is true, that is there is no reduction in average CO ₂ emissions for cars fitted with new device, when in fact H_0 is not true and there is a reduction in average CO ₂ emissions for cars fitted with new device	B1		Type II correctly explained																													
	E1	2	In context																													
	Total		13																													

Q	Solution	Marks	Total	Comments																
4(a)	<table border="1"> <thead> <tr> <th></th> <th>Minor</th> <th>Serious</th> <th>total</th> </tr> </thead> <tbody> <tr> <td>New</td> <td>9</td> <td>3</td> <td>12</td> </tr> <tr> <td>Current</td> <td>4</td> <td>9</td> <td>13</td> </tr> <tr> <td>Total</td> <td>13</td> <td>12</td> <td>25</td> </tr> </tbody> </table>		Minor	Serious	total	New	9	3	12	Current	4	9	13	Total	13	12	25	B1 M1	3	One total correct 13/12 One frequency correct
		Minor	Serious	total																
New	9	3	12																	
Current	4	9	13																	
Total	13	12	25																	
		A1	All Correct																	
(b)	<p>H_0: Mouthwash type not associated with category of mouth infection H_1: Mouthwash type is associated with category of mouth infection 1 tail 5%</p> <table border="1"> <thead> <tr> <th>Expected</th> <th>Minor</th> <th>Serious</th> </tr> </thead> <tbody> <tr> <td>New</td> <td>6.24</td> <td>5.76</td> </tr> <tr> <td>Current</td> <td>6.76</td> <td>6.24</td> </tr> </tbody> </table> $ts = \sum \frac{(O - E - 0.5)^2}{E}$ $= \frac{2.26^2}{6.24} + \frac{2.26^2}{5.76} + \frac{2.26^2}{6.76} + \frac{2.26^2}{6.24}$ $= 3.28$ <p>cv df = 1 5% cv = 3.84</p> <p>ts < 3.84 Accept H_0</p> <p>No significant evidence to suggest that mouthwash is associated with category of mouth infection</p>	Expected	Minor	Serious	New	6.24	5.76	Current	6.76	6.24	B1		H_0 Indep/No Assoc H_1 Not Indep/ Assoc							
	Expected	Minor	Serious																	
New	6.24	5.76																		
Current	6.76	6.24																		
		M1 A1		Method for expected freqs																
		M1		ts 'correct' with/without Yates																
		M1		Yates used correctly																
		A1		ts correct (3.1 – 3.4)																
		B1		cv cao																
				No Yates used can gain M1 A1 M1 B1 ts = 4.89																
		E1	8	Conclusion correct in contest and ts/cv correct																
	Total		11																	

Q	Solution	Marks	Total	Comments																					
5(a)	H_0 : Samples from identical populations	B1																							
	H_1 : Samples not from identical populations 5% sig level																								
	Ranks																								
	<table border="1"> <thead> <tr> <th>I</th> <th>II</th> <th>III</th> </tr> </thead> <tbody> <tr> <td>16 1</td> <td>13 4</td> <td>11 6</td> </tr> <tr> <td>15 2</td> <td>9 8</td> <td>7 10</td> </tr> <tr> <td>14 3</td> <td>6 11</td> <td>5 12</td> </tr> <tr> <td>12 5</td> <td>4 13</td> <td>3 14</td> </tr> <tr> <td>10 7</td> <td>1 16</td> <td>2 15</td> </tr> <tr> <td>8 9</td> <td></td> <td></td> </tr> </tbody> </table>	I	II	III	16 1	13 4	11 6	15 2	9 8	7 10	14 3	6 11	5 12	12 5	4 13	3 14	10 7	1 16	2 15	8 9			M1 A1		For ranks as one group 10 or more correct
	I	II	III																						
	16 1	13 4	11 6																						
	15 2	9 8	7 10																						
	14 3	6 11	5 12																						
	12 5	4 13	3 14																						
	10 7	1 16	2 15																						
8 9																									
Totals of ranks		m1		Totals – can be 27 52 57																					
$T_I = 75$ 27 $T_{II} = 33$ 52 $T_{III} = 28$ 57 $n_I = 6$ $n_{II} = 5$ $n_{III} = 5$																									
$\sum_{i=1}^m \frac{T_i^2}{n_i} = \frac{75^2}{6} + \frac{33^2}{5} + \frac{28^2}{5} = 1312.1$		m1 m1		Numerators correct Denominators correct																					
$H = \frac{12}{16 \times 17} \times 1312.1 - (3 \times 17)$		m1		H formula <u>correctly</u> used																					
= 6.89		A1		(6.7 – 7.1) ts/cv correct																					
Critical value from $\chi_2^2 = 5.991$ $H > 5.991$		B1		For cv cao																					
Reject H_0 . Significant evidence to suggest that samples are not from identical populations. Significant difference in average score for at least 2 of the three methods involved.		E1	10	Conclusion correct in context																					
(b) Use Approach I since <u>average of ranks</u> is highest (lowest ft) (so <u>lowest average scores</u>)		B1 E1	2	Approach I Reason (vice versa for reversed ranks). Allow reference to average scores.																					
	Total		12																						

Q	Solution	Marks	Total	Comments															
6(a)(i)	H ₀ : Colour preference is independent of personality H ₁ : Colour preference is not independent of personality 1 tail 5%	B1		H ₀ Indep / No Assoc H ₁ Not Indep / Assoc															
	<table border="1"> <thead> <tr> <th>Exp</th> <th>R</th> <th>Y</th> <th>G</th> <th>B</th> </tr> </thead> <tbody> <tr> <td>Intro</td> <td>47</td> <td>9.4</td> <td>18.8</td> <td>18.8</td> </tr> <tr> <td>Extro</td> <td>153</td> <td>30.6</td> <td>61.2</td> <td>61.2</td> </tr> </tbody> </table>	Exp	R	Y	G	B	Intro	47	9.4	18.8	18.8	Extro	153	30.6	61.2	61.2	M1 M1 A1		Any one E _i correct At least 5 correct All correct SC1 integers
	Exp	R	Y	G	B														
	Intro	47	9.4	18.8	18.8														
	Extro	153	30.6	61.2	61.2														
	$ts = \sum \frac{(O - E)^2}{E}$																		
	$= \frac{11^2}{47} + \frac{1.4^2}{9.4} + \frac{5.2^2}{18.8} + \frac{7.2^2}{18.8} + \frac{11^2}{153} +$ $\frac{1.4^2}{30.6} + \frac{5.2^2}{61.2} + \frac{7.2^2}{61.2}$	M1 M1		Numerators OK ft Denominators OK ft and added															
	= 9.12	A1		(9.0-9.3)															
	df = 3 5% cv = 7.815 ts > 7.815	B1		For cv or = 0.0277															
	Reject H ₀ Sig evidence to suggest colour preference is not independent of personality	E1	9	In context															
(ii)	Introverts <u>far more likely than expected</u> to prefer blue or green (introverts far less likely than expected to choose red)	B1		Alt Extroverts are <u>more likely than expected</u> to prefer red															
		E1	2																

Q	Solution	Marks	Total	Comments
6(b)(i)	d 0 0 1.5 0 0 1 1 1 2.5 $\sum d^2 = 11.5$	M1	3	Differences
	SRCC $r_s = 1 - \frac{6 \times \sum d^2}{9 \times 80} = 0.904$ or SRCC $r_s = 0.904$ (from calc)	M1 A1 (B3)		Formula correct SC1 0.9 SC2 0.90 if no method shown
(ii)	H_0 : Rank orders of personality score and happiness score are independent. H_1 : Rank orders of personality score and happiness score are not independent. 2 tail 1% $cv = 0.8167$ test stat $r_s = 0.904$ $r_s > cv$	B1	5	Hypothesis
		B1		cv cao
		M1		comparison ft seen or implied
		A1		ts/cv correct
		E1		in context – vice versa OK
	Total		19	
	TOTAL		75	