
A-LEVEL

Statistics

Statistics 3 – SS03

Mark scheme

6380
June 2015

Version 1.0 Final

Mark schemes are prepared by the Lead Assessment Writer 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 associates 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 associate understands and applies it in the same correct way. As preparation for standardisation each associate 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, associates encounter unusual answers which have not been raised they are required to refer these to the Lead Assessment Writer.

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

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.

Q1	Solution	Marks	Total	Comments
1(a)	$r = 0.809$ from calculator or $\sum xy = 158928$ B1 and $r = \frac{\frac{158928}{12} - \left(\frac{1495}{12} \times \frac{1271}{12}\right)}{\sqrt{\left(\frac{189473}{12} - \left(\frac{1495}{12}\right)^2\right) \times \left(\frac{134781}{12} - \left(\frac{1271}{12}\right)^2\right)}}$ oe $= \frac{48.549}{16.383 \times 3.662} = 0.809$ M1 A1	B3		sc 0.81 no workings B2 sc 0.8 no working B1 (0.799 – 0.815)
1(b)	$H_0: \rho = 0$ $H_1: \rho > 0$ 1 tail 1% test stat $r = 0.809$ critical value = 0.658 $0.809 > 0.658$ so significant evidence exists to reject H_0 This suggests that there is a positive correlation between height and systolic blood pressure for healthy boys aged between 5 years and 10 years.	B1 B1 M1 dep cv E1	3 4	Hypotheses oe Correct value for cv Comparison 'ts'/cv [or Reject H_0] Conclusion correct in context Dep ts and cv correct
1(c)	Conclusion can only refer to healthy/boys no girls , not all children		1	
		Total	8	

Mark Scheme

Q2	Solution			Marks	Total	Comments									
2(a)	Answer to Question			M1 M1 A1	3	For 35 correctly place For 12 correctly placed All correct									
		Yes	No												
	Year 13	28	12				40								
	Year 12	7	18				25								
	35	30	65												
2(b)	<p>H₀ Answer to question asked independent of year of study H₁ Answer to question asked not independent of year of study</p> <p>1 tail 1%</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>Expected</th> <th>Yes</th> <th>No</th> </tr> </thead> <tbody> <tr> <td>Year 13</td> <td style="text-align: center;">21.54</td> <td style="text-align: center;">18.46</td> </tr> <tr> <td>Year 12</td> <td style="text-align: center;">13.46</td> <td style="text-align: center;">11.54</td> </tr> </tbody> </table> $ts = \sum \frac{(O - E - 0.5)^2}{E}$ $= \frac{5.96^2}{21.54} + \frac{5.96^2}{18.46} + \frac{5.96^2}{13.46} + \frac{5.96^2}{11.54}$ $= 1.65 + 1.92 + 2.64 + 3.08$ $= \underline{9.29}$ <p>cv df = 1 1% <u>cv = 6.635</u></p> <p>ts > 6.635 Reject H₀</p> <p>Significant evidence to suggest that the answer to the question, "Do you have part-time employment?" is not independent of (is associated with) the year of study.</p>			Expected	Yes	No	Year 13	21.54	18.46	Year 12	13.46	11.54	B1		oe H ₀ No association/Independent H ₁ Association/Not Independent B0 if nonsense
Expected	Yes	No													
Year 13	21.54	18.46													
Year 12	13.46	11.54													
				M1 A1		Method for expected freqs (can be implied) All correct – at least 1 dp									
				M1		ts 'correct'									
				A1		without Yates $\frac{6.46^2}{21.54} \dots = \frac{41.7}{21.54} \dots$ oe (1.94+2.26+3.10+3.62 = 10.9*) Yates used correctly awfw 9.10 -9.50									
				B1		cv cao p= <u>0.0023 < 0.01</u>									
				E1		No Yates used can gain M1A1 M1A0 B1E0 ts = 10.9 Note $\frac{(28 - 21.54)^2 - 0.5}{21.54} \dots = 10.79$ oe M1A1M0A0B1E0									
					7	Conclusion correct in context									

		Total	10	
	<ul style="list-style-type: none"> M1 A1 if scaled correctly x 0.851 			

Mark Scheme

Q3	Solution	Marks	Total	Comments
(a)	$\text{Min } T = 1 + 2 + 3 + 4 + 5 + 6 = 21$	M1 A1	2	M1 for addition effort 1 to 6 oe sc 1 $21-21 / 21 - \frac{6 \times 7}{2} = 0$

(b)	<p>H_0 The two populations have identical distributions H_1 The two populations do not have identical distributions 2 tail 5%</p> $\begin{matrix} T_A = 46 & T_B = 74 \\ n_A = 7 & n_B = 8 \end{matrix}$ $U_A = 46 - \frac{7 \times 8}{2} = 18$ $U_B = 74 - \frac{8 \times 9}{2} = 38$ <p>Test stat $U = 18$</p> <p>$cv = 11$ for $n = 7, m = 8$ 2 tail 5%</p> <p>$U > 11$</p> <p>Accept H_0</p> <p>No significant evidence of a difference in accuracy for probes for the two manufacturers, A and B.</p>	<p>B1</p> <p>M1 A1</p> <p>B1</p> <p>m1dep</p> <p>E1</p>	<p>only</p> <p>Attempt to find U Either U correct</p> <p>cv correct (or 45)</p> <p>comparison consistent clear ('18' with 11 or '38' with 45) cv correct</p> <p>6</p>	<p>Correct conclusion in context (E0 if reference to mean or probes/manufacturers the same)</p>
		Total	8	

Mark Scheme

Q4	Solution	Marks	Total	Comments																																																																	
4(a)	<table border="1"> <thead> <tr> <th rowspan="2">Bank</th> <th colspan="2">Ranks</th> <th colspan="2"></th> <th rowspan="2">Assets</th> </tr> <tr> <th>Complaint</th> <th>Satisfaction</th> <th colspan="2"></th> </tr> </thead> <tbody> <tr> <td>A</td> <td>2 8</td> <td>8 2</td> <td colspan="2"></td> <td>1 9</td> </tr> <tr> <td>B</td> <td>3 7</td> <td>7 3</td> <td colspan="2"></td> <td>2 8</td> </tr> <tr> <td>C</td> <td>1 9</td> <td>9 1</td> <td colspan="2"></td> <td>3 7</td> </tr> <tr> <td>D</td> <td>5 5</td> <td>2½ 7½</td> <td colspan="2"></td> <td>4 6</td> </tr> <tr> <td>E</td> <td>6 4</td> <td>1 9</td> <td colspan="2"></td> <td>5 5</td> </tr> <tr> <td>F</td> <td>9 1</td> <td>4 6</td> <td colspan="2"></td> <td>6 4</td> </tr> <tr> <td>G</td> <td>4 6</td> <td>6 4</td> <td colspan="2"></td> <td>7 3</td> </tr> <tr> <td>H</td> <td>7 3</td> <td>5 5</td> <td colspan="2"></td> <td>8 2</td> </tr> <tr> <td>I</td> <td>8 2</td> <td>2½ 7½</td> <td colspan="2"></td> <td>9 1</td> </tr> </tbody> </table>	Bank	Ranks				Assets	Complaint	Satisfaction			A	2 8	8 2			1 9	B	3 7	7 3			2 8	C	1 9	9 1			3 7	D	5 5	2½ 7½			4 6	E	6 4	1 9			5 5	F	9 1	4 6			6 4	G	4 6	6 4			7 3	H	7 3	5 5			8 2	I	8 2	2½ 7½			9 1				
	Bank		Ranks					Assets																																																													
		Complaint	Satisfaction																																																																		
	A	2 8	8 2			1 9																																																															
	B	3 7	7 3			2 8																																																															
	C	1 9	9 1			3 7																																																															
	D	5 5	2½ 7½			4 6																																																															
	E	6 4	1 9			5 5																																																															
	F	9 1	4 6			6 4																																																															
	G	4 6	6 4			7 3																																																															
H	7 3	5 5			8 2																																																																
I	8 2	2½ 7½			9 1																																																																
(i)	$r_s = 0.767$ (or 0.766) from calculator or $d = 1, 1, -2, 1, 1, 3, -3, -1, -1$ $\sum d^2 = 28$ SRCC $r_s = 1 - \frac{6 \times 28}{9 \times 80} = 0.767$	B3 or M1 m1 A1		Attempt to rank – any order Consistent ranking – all three Ties correct Can be implied by correct r values sc2 no method $r_s = 0.77$ or 0.76 sc1 - 0.767 Differences and effort $\sum d^2$ Formula correct awrt																																																																	
(ii)	$r_s = -0.544$ (or -0.543) from calculator or $d = 7, 5, 6, -1.5, -4, -2, -1, -3, -6.5$ $\sum d^2 = 184.5$ SRCC $r_s = 1 - \frac{6 \times 184.5}{9 \times 80} = -0.537(5)$ or -0.538	B3 or M1 m1 A1		sc2 no method $r_s = -0.54$ sc1 + $0.544/3$ Differences $\sum d^2$ Formula awrt																																																																	
			9																																																																		

(b)(i)	H ₀ Rank orders of upheld complaints and assets are independent. H ₁ Rank orders of upheld complaints and assets are not independent.	B1	Either pair of hypotheses correct or generic
	2 tail 5% $cv = 0.6833$ $r_s = 0.767$ Reject H ₀ . Significant evidence at 5% level to suggest an association/correlation between rank orders of upheld complaints and assets . Banks with higher assets tend to have a higher level of upheld complaints.	B1	cv correct (condone +/- consistent)
		E1	Conclusion in context ts/cv consistent Condone 'slight error' in part (a)
	(ii) H ₀ Rank orders of customer satisfaction ratings and assets are independent. H ₁ Rank orders of customer satisfaction ratings and assets are not independent.		
	2 tail 5% $cv = -0.6833$ $r_s = -0.544$ or $-0.537/8$ Accept H ₀ . No significant evidence at 5% level to suggest an association between rank orders of customer satisfaction ratings and assets .	B1	cv correct (condone +/- consistent)
		E1	Conclusion in context ts/cv correct Condone 'slight error' in part (a)
			5
(c)	H ₀ $\eta = 15$ H ₁ $\eta < 15$ 1 tail test 10% level Signs - - - + + + - - + test stat = 5 -/ 4+	B1	Allow pop median
		M1	for signs
		A1	for test stat
	Bin (9, 0.5) model $P(\leq 4+) = 0.500 > 0.10$ Or cr {0,1,2} or {7,8,9} inc probs seen		
		M1	for use of correct Bin model (allow sc B1 for 0.746/0.254 seen) and comparison ts and 10%
Accept H ₀ . No significant evidence to suggest that average customer satisfaction/ rating is less than 15 .		E1	oe dep all correct do not allow statement 'customer satisfaction rating is 15'
			5
			5
		Total	19

Mark Scheme

Q5	Solution	Marks	Total	Comments																											
(a)	Ranks																														
	<table border="1"> <thead> <tr> <th>Terrific Teen</th> <th>Beetleman</th> <th>Hunk</th> <th>Warrior Crab</th> </tr> </thead> <tbody> <tr> <td>4</td> <td>1</td> <td>6</td> <td>12</td> </tr> <tr> <td>7½</td> <td>2</td> <td>11</td> <td>16</td> </tr> <tr> <td>9</td> <td>3</td> <td>14</td> <td>18</td> </tr> <tr> <td>10</td> <td>5</td> <td>15</td> <td>19</td> </tr> <tr> <td>13</td> <td>7½</td> <td>20</td> <td>21</td> </tr> <tr> <td>17</td> <td></td> <td>22</td> <td></td> </tr> </tbody> </table>	Terrific Teen	Beetleman	Hunk	Warrior Crab	4	1	6	12	7½	2	11	16	9	3	14	18	10	5	15	19	13	7½	20	21	17		22		M1 m1 A1	
Terrific Teen	Beetleman	Hunk	Warrior Crab																												
4	1	6	12																												
7½	2	11	16																												
9	3	14	18																												
10	5	15	19																												
13	7½	20	21																												
17		22																													
5(b)	<p>H_0: Samples from identical populations H_1: Samples not from identical populations 1% sig level</p> <p>Totals $T_{Teen} = 60\frac{1}{2}$ $T_{Beetle} = 18\frac{1}{2}$ $T_{Hunk} = 88$ $T_{Crab} = 86$ $n_{Teen} = 6$ $n_{Beetle} = 5$ $n_{Hunk} = 6$ $n_{Crab} = 5$</p> $\sum_{i=1}^m \frac{T_i^2}{n_i} = \frac{60.5^2}{6} + \frac{18.5^2}{5} + \frac{88^2}{6} + \frac{86^2}{5} = 3448.36$ $H = \frac{12}{22 \times 23} \times 3448.36 - (3 \times 23) = 12.78$ Critical value from $\chi_3^2 = \underline{11.3(45)}$ $H > 11.345$ Significant evidence to <u>reject H_0</u> <p>There is significant evidence of a difference between average scores for at least 2 superhero costumes. (can be implied by comment that children wearing Beetleman costumes are clearly more exuberant than those wearing Warrior Crab costumes.)</p> <table border="1"> <thead> <tr> <th></th> <th>T Teen</th> <th>Beetle</th> <th>Hunk</th> <th>Crab</th> </tr> </thead> <tbody> <tr> <td></td> <td>rank score</td> <td>rank score</td> <td>rank score</td> <td>rank score</td> </tr> <tr> <td>Mean</td> <td>10 45.7</td> <td>3.7 66.2</td> <td>14.7 34.5</td> <td>17.2 29.6</td> </tr> <tr> <td>Med</td> <td>9.5 46</td> <td>3 65</td> <td>14.5 37</td> <td>18 30</td> </tr> </tbody> </table>		T Teen	Beetle	Hunk	Crab		rank score	rank score	rank score	rank score	Mean	10 45.7	3.7 66.2	14.7 34.5	17.2 29.6	Med	9.5 46	3 65	14.5 37	18 30	B1 M1 m1 M1 A1 B1 A1dep E1dep Rej H_0 B1	3	oe Allow η or pop median but need 'at least two differ' [not 'at least one differs'] Totals of <u>rank</u> s 56½ 18½ 83 81 ft sc $\sum_{i=1}^m \frac{T_i^2}{n_i}$ m1 (sc1 3060.8 seen) H formula attempt correct A1 awfw 12.4 – 13.8 Cao 11.3 or better Reject H_0 There is significant evidence of a difference between average scores for at least 2 superhero costumes. (or ref to difference between Beetleman and Warrior crab.) Difference in exuberance and means/medians considered – ranks or raw scores <u>seen</u> considered							
	T Teen	Beetle	Hunk	Crab																											
	rank score	rank score	rank score	rank score																											
Mean	10 45.7	3.7 66.2	14.7 34.5	17.2 29.6																											
Med	9.5 46	3 65	14.5 37	18 30																											

	Children wearing Beetleman costumes clearly displayed more exuberance (than those wearing Warrior Crab)	E1		Mention Beetleman most exuberant (allow without backup)
		Total	10	
			13	

Mark Scheme

Q6	Solution	Marks	Total	Comments
-----------	-----------------	--------------	--------------	-----------------

6(a)	$H_0: \eta \text{ or } \mu = 6.5$ $H_1: \eta \text{ or } \mu \neq 6.5$ 2 tail test 1 % level								B1	oe ref population medians/means Differences $x - 6.5$ (disregard sign) Ranks (smallest abs diff = rank 1) Disallow assigning rank 1 to 0 Effort at total of any ranks allow m1 here if zero included either total correct cao for cv Correct comparison 9,13 or 96,92 Correct conclusion in context	
	diff $x-6.5$	-	+	-	-	+	-	-	M1		
	rank	2.3	0.9	1.5	3.7	0.8	3	2.6			
		7	4	5	13	3	12	9	m1dep		
	diff $x-6.5$	+	-	-	-	-	-	-			
	rank	0.7	2.9	4.3	2.8	0	1.7	0.6			
		2	11	14	10	.	6	1			
	diff $x-6.5$	-									m1dep
	rank	2.4									
		8									A1
$T_+ = 4 + 3 + 2 = 9$ $T_- = 7 + 5 + 13 + \dots + 8 = 96$ test stat $T = 9$ critical value = <u>13</u> test stat $9 < cv$ 13 Reject H_0 There is significant evidence that the median (average) time for healthy adults taking the new drug to achieve persistent sleep when their regular bedtime is shifted earlier by five hours differs (is lower) from 6.5 minutes.								B1 M1			
								E1dep			
								8			

Q6	Solution	Marks	Total	Comments
6(b) (i)	So that any influence of the order of taking the different levels of the drug does not affect the outcome of the investigation.	1		Condone decrease of 'demand characteristics' by volunteers
6(b) (ii)	<p>H_0: Population mean μ/median η(difference)=0 H_1: Population mean μ/median η(difference)>0 1 tail test 5 % level</p> <p>Differences 20mg – 50mg A B C D E F G H I J -0.2 2.1 0.9 -0.6 1.7 -1 -0.3 1.9 2.2 0.4</p> <p>Ranks 1 9 5 4 7 6 2 8 10 3</p> <p>$T_+ = 9 + 5 + 7 + 8 + 10 + 3 = 42$ $T_- = 1 + 4 + 6 + 2 = 13$ test stat $T = 13$</p> <p>critical value = <u>11</u> test stat 13 > 11</p> <p>Accept H_0 No significant evidence to suggest that the average number of minutes/time taken by healthy adults to achieve persistent sleep is lower when taking 50mg of the new drug half an hour before bedtime than when taking 20mg of the new drug half an hour before bedtime.</p>	<p>B1</p> <p>M1</p> <p>m1dep</p> <p>m1dep</p> <p>A1</p> <p>B1</p> <p>M1</p> <p>E1</p>	<p>9</p> <p>17</p>	<p>difference 20mg – 50mg or reverse H_1 and 50mg – 20mg consistent with signs of differences</p> <p>For differences</p> <p>For ranks of any differences (smallest abs diff = rank1). Effort at total of any ranks (dep ranks any effort)</p> <p>Either total correct</p> <p>cv correct consistent 13 ,11 or 42,44 comparison</p> <p>Must be <u>consistent with H_1</u> Disallow 'times taken are same' stated</p>
		Total	17	