AQA

AS
STATISTICS

SS03 Statistics
Mark scheme

6380
June 2018

Version/Stage: 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

Copyright © 2018 AQA and its licensors. All rights reserved.
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.

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	Mark	Total	Comment
	H_0: Reaction is independent of location of injection. H_1: Reaction is not independent of location of injection.	B1		For H_0 oe ref to no association
	1 tail 10%	M1PI		Attempt to find expected frequencies $\frac{4957\times3244}{5000}$ = 3216.1 etc
	Expected frequencies **No**: Arm 3216.1, Thigh 1740.9 **Yes**: Arm 27.9, Thigh 15.1	A1		All expected freq correct AWRT 1 decimal place (A0 if integers)
	$\chi^2 = \sum \frac{(\lvert(O-E)\rvert-0.5)^2}{E}$	m1PI		Attempt at test statistic
	$= \frac{3.4^2}{3216.1}+\frac{3.4^2}{1740.9}+\frac{3.4^2}{27.9}+\frac{3.4^2}{15.1}$	m1PI		Correct use of Yates' correction Allow 'their' $\lvert(O-E)\rvert-0.5$
	= 1.19	A1		awrt 1.19 (note that no Yates used gives test stat = 1.56)
	ts χ^2 = 1.19			
	df = 1 10% cv = 2.706 p =0.275 ts < 2.706 p > 0.10	B1		For correct cv (awrt 2.7) or correct p value (awfw 0.27 – 0.28)
	Accept H_0			
	No significant evidence to suggest that reaction and location of injection are associated (are not independent)	E1dep	**8**	Correct conclusion in context dep all previous working correct Condone B0 for hypotheses No Yates gains B1,M1,A1,m1,A0,B1,E0
	Total		**8**	

Q2	Solution	Mark	Total	Comment
	H_0: η_d , $\mu_d = 0$ H_1: η_d , $\mu_d > 0$ where *d* is Nov – Aug difference	B1		
	Where N represents November and A represents August 1 tail test 1 % level Differences *N – A*			
	Area / **Difference N -A** / **Ranks of Diffs** **1** -5.6 8 **2** -2.2 5 **3** -1.2 3 **4** 0.1 1 **5** 1.0 2 **6** 2.0 4 **7** 3.1 6 **8** 5.3 7 **9** 6.3 9 **10** 7.2 10 **11** 12.0 11 **12** 12.3 12 **13** 23.4 13	M1 m1 m1		For differences – ignore signs Can be implied by correct ranks For any ranks effort For correct ranks (smallest abs diff = rank1).
	$T_+ = 75$ $T_- = 16$ $T_+ = 1 + 2 + \ldots + 13 = 75$ $T_- = 8 + 5 + 3 = 16$ test stat T = 16	m1 A1		Effort at total of ranks + and/or - Either total correct
	critical value = 13 (or 78 if ts = 75 used) test stat 16 >13 (or 78 > 75 for upper tail) Accept H_o	B1 M1		cv = 13 correct correct lower tail ts used/identified (or upper tail/upper cv correct)
	There is no significant evidence to suggest that the average aluminium content in poplar trees was higher in November 2014 than in August 2014.	E1dep	**9**	Correct conclusion dep ts and cv correct
	Total		**9**	

Q3	Solution	Mark	Total	Comment
(a)(i)	Ranks: **A** 11, 3; **B** 10, 1; **C** 9, 7; **D** 8, 2; **E** 7, 4; **F** 6, 5; **G** 5, 6; **H** 4, 9; **I** 3, 8; **J** 2, 11; **K** 1, 10	M1 M1	5	Effort to rank one category Rank of other category – consistent (condone inconsistent if clear understanding of nature of correlation explained in context in part (a)(ii))
	r_S = - 0.855 from calculator OR d = 8, 9, 2, 6, 3, 1, -1, -5, -5, -9, -9 $\sum d^2$ = 408 SRCC $r_s = 1-\frac{6\times 408}{11\times 120}$ = - 0.855	B3 OR m1 B1 A1		SRCC from calculator directly using ranks -0.854 or -0.855 (note PMCC = - 0.867) Obtains differences $\sum d^2$ correct SRCC correct -0.854 or -0.855
(ii)	A strong negative rank correlation between <u>volume</u> of inflated throat sac and drumming <u>frequency</u>. Males with larger inflated throat sacs tend to produce lower frequency drumming sounds.	B1 E1	2	Strong negative mentioned (allow strong positive if inconsistent ranking and value correctly interpreted in context) Explained in context
(b)	H_o Rank orders of sac volume and drumming sound frequency are independent. (no association) H_1 Rank orders of sac volume and drumming sound frequency are not independent. (negative association) 1 tail 5% |cv| = 0.5273 $\|r_S\|$ = 0.855 > 0.5273	B1 B1 m1dep		Hypotheses cv correct Consistent comparison SRCC/cv dep cv correct

	Reject H_o It appears that Diego may be correct since there is **significant evidence** that male birds with larger inflated sac volume tend to produce drumming sounds with lower frequency so females may be able to detect lower drumming frequency and identify males with larger inflated throat sacs. However, there is no evidence to suggest that the female birds do use the association to identify males.	A1 E1	5	Conclusion correct– dep cv and SRCC correct Valid comment regarding negative association significant in context with reference to Diego's claim OR comment regarding correlation not implying causation OR comment regarding insufficient information available regarding behaviour of frigate birds Note – comment may have been made in part(a)(ii). Allow E1 if this is the case.
	Total		**12**	

Q4	Solution	Mark	Total	Comment
(a)	H_0: Samples from identical populations H_1: Samples not from identical populations 5% sig level	B1		oe referring to population medians (η symbol or words)
	Ranks			
	No exercise | **20 minutes** | **60 minutes** 1 | 16 | 2 | 15 | 11½ | 5½ 6½ | 10½ | 4½ | 12½ | 13 | 4 8 | 9 | 11½ | 5½ | 15 | 2 4½ | 12½ | 10 | 7 | 16 | 1 3 | 14 | 6½ | 10½ | 9 | 8 | | 3 | 14 | |	M1 A1		Attempt at ranking with some ties (can be reversed) All correct
	23 | **62** | **48½** | **53½** | **64½** | **20½**	m1 A1		Attempt at totals One total correct
	$T_{no\ ex}$ T_{20} T_{60} $n_{no\ ex}$= 6 n_{20} = 5 n_{60} = 5			
	$\sum \frac{T_i^2}{n_i} = \frac{62^2}{5} + \frac{53.5^2}{6} + \frac{20.5^2}{5}$ =1329.9	m1		effort at $\sum \frac{T_i^2}{n_i}$ ft
	$H = \frac{12}{16\times17} \times 1329.9 - (3\times17)$ = 7.67	m1 A1		Formula for H awfw 7.5 – 7.9
	Critical value from v = 2 cv = 5.991 H > 5.991	B1		cao for cv or p = 0.02158
	Significant evidence to reject H_o There is significant evidence of a difference between average scores for <u>at least two</u> exercise regimes	A1dep B1		Reject H_o Dep all previous method correct At least two differ mentioned PI explanation in context

		No ex	**20**	**60**		
		rank	rank	rank		
	Mean	12.4	8.9	4.1		
	Med	12.5	8.75	4		
		score	**score**	**score**		
	Mean	10.2	13.8	18.6		
	Med	11	14	18		
	60 minutes of exercise appears to result in a higher score for current state of mind (meaning people undertaking 60 minutes exercise each day appear, on average, to be happier) than those undertaking no exercise at all.				E1 12	In context – must identify '60 minutes exercise' as leading to higher score and supply some numerical justification
(b)	Removing her score of 15 would strengthen conclusion …….hence no effect on conclusion OR May need to find out why she stopped (perhaps 60 minutes of jogging daily is not successful for all as it may be too much for some people and they would not stick with the regime) ……..hence exercise may not help all people (beneficial effect of 60 minutes jogging daily may not be as important in general if such patients give up completely so conclusion that 60 minutes jogging makes depressed people feel happier may not be true for all such patients.)				E1 E1 (E1) (E1) 2	E1 for valid reason E1 for effect on conclusion (ft small slip in part (a))
(c)	Randomly assigning patients means that no deliberate, or unconscious, bias can occur in allocation to exercise regime. For example, fitter/healthier/happier looking patients being assigned to the '60 minutes of jogging' regime				E1 E1 2	Reason of 'no bias' in allocation Example in context – any valid reason in context accepted
	Total				**16**	

Q5	Solution	Mark	Total	Comment
(a)	H_0: Tyre quality is independent of shift H_1: Tyre quality is not independent of shift 1 tail 5% Expected freqs	B1		oe
		M1		One exp freq correct or method seen
		m1		At least 6 exp freq correct (allow integers)
		A1		All correct to 1 dp
	$\chi^2 = \sum \frac{(O-E)^2}{E} =$	m1		Numerator correct in ts
	$\frac{(77-77.95)^2}{77.95} + \frac{(89-77.95)^2}{77.95} +$	m1		Denominator correct in ts
	$\ldots\ldots \frac{(7-9.42)^2}{9.42}$	m1		Summation $\sum \frac{(O-E)^2}{E}$
	= 0.0116 + 1.566 + 1.497 + 0.0006 + 2.062 + 2.105 + 0.004 + 0.454 + 0.623			
	ts χ^2 = 8.33	A1		awfw 8.10 – 8.50
	df = 4 10% cv = 7.779 p = 0.080	B1		df correct PI (or p value correct)
	ts > 7.779 or p < 0.10	M1		cv correct and comparison ts and cv or p value compared with 0.10
	Reject H_0	A1dep		Conclusion correct
	Sig evidence to suggest that tyre quality is not independent of shift	E1		Correct and in context
			12	
(b)	Biggest sources of association are indicated by 2.062 and 2.105 (followed by 1.566 and 1.500)	B1		Ref to contributions to ts / largest $\frac{(O-E)^2}{E}$ PI
		B1		Correct contribution to ts stated or correct obs and expected frequencies comparison stated anywhere

	1	2	3
Excellent	77.95	77.95	68.10
Satisfactory	85.26	85.26	74.48
Defect	10.79	10.79	9.42

	It appears that: **Shift 2** produced **fewer** tyres than **expected** (obs 72, exp 85.26) in the **'Satisfactory'** regime [and **more** than **expected** in the (obs 89, exp 77.9)5 **'Excellent'** regime] **Shift 3** produced **more** tyres than **expected** (obs 87, exp 74.48) in the **'Satisfactory'** regime [and **fewer** than **expected** in the (obs 58, exp 68.10) **'Excellent'** regime]	E1 E1	4	E1, E1 Must ref to obs and/or exp frequencies (not necessary to quote values) Any two different valid points made
	Total		**16**	

<table>
<tr><th>Q6</th><th>Solution</th><th>Mark</th><th>Total</th><th>Comment</th></tr>
<tr><td>(a)</td><td>H_o The two populations have identical distributions
H_1 The two populations do not have identical distributions
(attendants with short recovery periods have, on average, slower reaction times than those with long recovery periods.)

Ranks
<table><tr><th colspan="2">Short recovery</th><th colspan="2">Long recovery</th></tr><tr><td>15</td><td>6</td><td>20</td><td>1</td></tr><tr><td>12</td><td>9</td><td>19</td><td>2</td></tr><tr><td>10</td><td>11</td><td>18</td><td>3</td></tr><tr><td>9</td><td>12</td><td>17</td><td>4</td></tr><tr><td>8</td><td>13</td><td>16</td><td>5</td></tr><tr><td>6</td><td>15</td><td>14</td><td>7</td></tr><tr><td>5</td><td>16</td><td>13</td><td>8</td></tr><tr><td>4</td><td>17</td><td>11</td><td>10</td></tr><tr><td>2</td><td>19</td><td>7</td><td>14</td></tr><tr><td>1</td><td>20</td><td>3</td><td>18</td></tr></table>
$U_s = 72 - \frac{10\times11}{2} = 17$

$U_L = 138 - \frac{10\times11}{2} = 83$

cv = 24 (or upper tail 76)
17 < 24 (or 83 > 76)

Reject H_o

Significant evidence to suggest that attendants with <u>short recovery</u> periods have, on average, <u>slower reaction times</u> than those with <u>long recovery</u> periods</td><td>B1

M1

m1

m1

A1

B1
m1

A1

E1</td><td>9</td><td>oe referring to population medians (η symbol or words)

Ranks effort as one group– can be reversed

Totals of ranks effort

Method correct for U

Either U correct

cv correct cao
comparison consistent

Conclusion correct

Explanation in context correct</td></tr>
</table>

(b)	H_0: $\eta_d = 0$ H_1: $\eta_d > 0$ 1 tail 10% + + - + - + + + - + +	B1		difference refers to Flight Attendant – Server
	ts = 8+ / 3 –	M1		Signs consistent with H_1
	B (11, 0.5)	M1		0.113 obtained
	P(≤ 3 -) = 0.113 > 0.10 Accept H_0	m1		Correct comparison 0.113 and 0.10
	No significant evidence to support suggestion that flight attendants suffer, on average, from greater hearing loss than servers in busy restaurants.	E1	5	Conclusion in context dep all previous method correct
	Total		**14**	

	TOTAL MARK FOR PAPER		**75**	