



---

AS

# Statistics

SS03

Final Mark Scheme

---

6380

June 2017

---

Version/Stage: v1.0

---

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](http://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)	$H_0: \eta_d \text{ or } \mu_d = 0$ $H_1: \eta_d \text{ or } \mu_d < 0$ 1 tail test 5 % level	B1		$H_0$ : Population median/mean/average <b>difference = 0</b> $H_1$ : Population median/mean/average <b>difference &lt; 0</b> consistent with <b>differences</b>																																				
	<table border="1"> <thead> <tr> <th>Man</th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <td><i>d</i></td> <td>-9</td> <td>-4</td> <td>-8</td> <td>+1</td> <td>.</td> </tr> <tr> <td>rank</td> <td>7</td> <td>3</td> <td>6</td> <td>1</td> <td>.</td> </tr> <tr> <th>Man</th> <th>6</th> <th>7</th> <th>8</th> <th>9</th> <th>10</th> </tr> <tr> <td><i>d</i></td> <td>+5</td> <td>-18</td> <td>-10</td> <td>+3</td> <td>-7</td> </tr> <tr> <td>rank</td> <td>4</td> <td>9</td> <td>8</td> <td>2</td> <td>5</td> </tr> </tbody> </table>	Man	1	2	3	4	5	<i>d</i>	-9	-4	-8	+1	.	rank	7	3	6	1	.	Man	6	7	8	9	10	<i>d</i>	+5	-18	-10	+3	-7	rank	4	9	8	2	5	M1		Differences (ignore sign)
	Man	1	2	3	4	5																																		
	<i>d</i>	-9	-4	-8	+1	.																																		
	rank	7	3	6	1	.																																		
	Man	6	7	8	9	10																																		
	<i>d</i>	+5	-18	-10	+3	-7																																		
	rank	4	9	8	2	5																																		
			m1		Ranks (smallest abs diff = rank 1)																																			
	$T_+ = 1 + 4 + 2 = 7$ $T_- = 7 + 3 + 6 + 9 + 8 + 5 = 38$ test stat $T = 7$		m1		Effort at total for <b>ranks</b>																																			
$cv = 8 \quad n = 9$		A1		<b>dep +1 gains rank 1</b>																																				
$T < 8$		B1		One total correct																																				
Reject $H_0$		B1		cao for cv																																				
Significant evidence to suggest that the <b>average systolic blood pressure is lower after three hours of not smoking</b> (a cigarette than immediately after smoking a cigarette.)		dM1		Correct comparison lower/lower or upper/upper tail ts and dep cv correct																																				
		E1		Condone small slip (reasonable $T$ ) for M1																																				
			<b>8</b>	Conclusion correct in context Dep ts/cv correct																																				
(b)	<p>Males and females may react differently in terms of blood pressure when they smoke a cigarette and this would mean that a difference in blood pressure, if one exists, may not be identified.</p> <p>A difference in blood pressure after smoking a cigarette may be more likely to be identified, if one exists, because the people in the trial were all 'matched' in terms of sex.</p>	B1 E1		B1 valid reason- males/females different or 'matching' better																																				
		or		E1 Explained in context of the test that difference more likely to be detected																																				
		B1 E1		Allow 'smoking is the only variable involved'																																				
			<b>2</b>	Allow 'Reduces experimental error due to factor of sex/gender removed'																																				
	<b>Total</b>		<b>10</b>																																					

Q2	Solution	Marks	Total	Comments																																																		
2(a)	<table border="1"> <thead> <tr> <th></th> <th colspan="2">x</th> <th colspan="2">y</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>2</td> <td>8</td> <td>1</td> <td>9</td> </tr> <tr> <td>B</td> <td>4</td> <td>6</td> <td>4½</td> <td>5½</td> </tr> <tr> <td>C</td> <td>8½</td> <td>1½</td> <td>6</td> <td>4</td> </tr> <tr> <td>D</td> <td>5</td> <td>5</td> <td>7</td> <td>3</td> </tr> <tr> <td>E</td> <td>1</td> <td>9</td> <td>3</td> <td>7</td> </tr> <tr> <td>F</td> <td>8½</td> <td>1½</td> <td>8</td> <td>2</td> </tr> <tr> <td>G</td> <td>7</td> <td>3</td> <td>9</td> <td>1</td> </tr> <tr> <td>H</td> <td>3</td> <td>7</td> <td>2</td> <td>8</td> </tr> <tr> <td>I</td> <td>6</td> <td>4</td> <td>4½</td> <td>5½</td> </tr> </tbody> </table> <p> <math>r_s = \mathbf{0.807}</math> from calculator  or  <math>d = 1, -\frac{1}{2}, 2\frac{1}{2}, -2, -2, \frac{1}{2}, -2, 1, 1\frac{1}{2}</math>  <math>\sum d^2 = 23</math>  SRCC <math>r_s = 1 - \frac{6 \times 23}{9 \times 80} = \mathbf{0.808}</math> </p>		x		y		A	2	8	1	9	B	4	6	4½	5½	C	8½	1½	6	4	D	5	5	7	3	E	1	9	3	7	F	8½	1½	8	2	G	7	3	9	1	H	3	7	2	8	I	6	4	4½	5½	M1 m1 A1 B3 (m1) (m1) (A1)	6	Attempt to rank one column/category Correct use of ties at least once All correct and consistent sc Ranked as one group gains M1 m1 only (condone <b>0.806</b> ) AWRT Differences (ignore sign) Formula correct sc2 no method $r_s = 0.81$ sc1 no method $r_s = 0.8$ (-B1 if negative) sc 4 -0.806/7/8 nms Note PMCC 0.691 gains 0
	x		y																																																			
A	2	8	1	9																																																		
B	4	6	4½	5½																																																		
C	8½	1½	6	4																																																		
D	5	5	7	3																																																		
E	1	9	3	7																																																		
F	8½	1½	8	2																																																		
G	7	3	9	1																																																		
H	3	7	2	8																																																		
I	6	4	4½	5½																																																		
(b)(i)	<p> <math>H_0</math> Rank orders of price and cocoa content are not associated  <math>H_1</math> Rank orders of price and cocoa content have a positive association </p> <p> 1 tail 1% <math> cv  = 0.7667</math>  <math> r_s  = 0.808</math> or <math>0.807</math> or <math>0.806 &gt; 0.7667</math> </p>	B1 B1 M1	4	Correct hypotheses oe $H_0$ no correlation $H_1$ positive correlation cv correct cao awrt 0.767 consistent comparison with correct cv +/- or -/- (allow small slip in part (a))																																																		
(ii)	<p> Reject <math>H_0</math>  Significant evidence to suggest that rank orders of recommended <b>retail price</b> and percentage cocoa/ <b>cocoa content</b> have a positive association. </p> <p> <math>H_0</math> rejected in error. </p> <p> Conclusion that there is a positive association between recommended retail price and percentage cocoa content, when, in fact, there is no positive association. </p>	E1 E1 E1	2	Correct conclusion in context need price and <u>cocoa percentage</u> or <u>cocoa content</u> Dep ts/cv correct Type I error defined In context need price and <u>cocoa percentage</u> or <u>cocoa content</u>																																																		
	<b>Total</b>		<b>12</b>																																																			





Q4	Solution	Marks	Total	Comments																														
4(a)	Ranks																																	
	<table border="1"> <thead> <tr> <th>V Low</th> <th>Low</th> <th>Moderate</th> <th>High</th> </tr> </thead> <tbody> <tr> <td>3</td> <td><u>15</u></td> <td><u>14</u></td> <td><u>20</u></td> </tr> <tr> <td>4</td> <td>9</td> <td><u>17</u></td> <td><u>24</u></td> </tr> <tr> <td>8</td> <td>5</td> <td><u>23</u></td> <td><u>21½</u></td> </tr> <tr> <td>1</td> <td>11</td> <td>12</td> <td><u>19</u></td> </tr> <tr> <td>7</td> <td>6</td> <td>13</td> <td><u>21½</u></td> </tr> <tr> <td>10</td> <td>2</td> <td><u>17</u></td> <td><u>17</u></td> </tr> </tbody> </table>	V Low	Low	Moderate	High	3	<u>15</u>	<u>14</u>	<u>20</u>	4	9	<u>17</u>	<u>24</u>	8	5	<u>23</u>	<u>21½</u>	1	11	12	<u>19</u>	7	6	13	<u>21½</u>	10	2	<u>17</u>	<u>17</u>					
	V Low	Low	Moderate	High																														
	3	<u>15</u>	<u>14</u>	<u>20</u>																														
	4	9	<u>17</u>	<u>24</u>																														
	8	5	<u>23</u>	<u>21½</u>																														
	1	11	12	<u>19</u>																														
	7	6	13	<u>21½</u>																														
	10	2	<u>17</u>	<u>17</u>																														
			M1		Any ranks effort																													
		M1		5 correct ( Allow 21, 22 rather than tied ranks)																														
		A1	<b>3</b>	All correct																														
(b)	H <sub>0</sub> : Samples from identical populations H <sub>1</sub> : Samples not from identical populations 1% sig level			oe																														
	Totals T <sub>0</sub> = 33    T <sub>1</sub> = 48    T <sub>3</sub> = 96    T <sub>9</sub> = 123  n <sub>0</sub> = 6        n <sub>1</sub> = 6        n <sub>3</sub> = 6        n <sub>9</sub> = 6		m1		Attempt at totals dep ranks																													
			A1		One total correct																													
	$\sum \frac{T_i^2}{n_i} = \frac{33^2}{6} + \frac{48^2}{6} + \frac{96^2}{6} + \frac{123^2}{6} = 4623$		m1		effort PI																													
	$H = \frac{12}{24 \times 25} \times 4623 - 3 \times 25 = 17.46$		m1		Formula for H dep previous m1																													
			A1		awfw 17.3 – 17.6																													
	Critical value from v = 3 cv = 11.345 H > 11.345		B1		v = 3 PI 6.251,7.815,9.348,12.838																													
			B1		cv cao																													
	Significant evidence to reject H <sub>0</sub> . There is significant evidence of a difference between <b>average</b> (not mean) <b>yield</b> for <b>at least 2 categories of weed density</b> ( very Low different to High)		A1		Reject H <sub>0</sub> Dep ts/cv correct																													
	<b>Very low weed density results in highest average corn yield</b>		E1		In context and must include 'at least two differ' Dep ts/cv correct																													
as rank average 5.5/total 33 oe		B1		In context – must identify Very low as highest/most yield																														
<table border="1"> <thead> <tr> <th></th> <th>VL</th> <th>L</th> <th>M</th> <th>H</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><b>Mean</b></td> <td>5.5</td> <td>8.0</td> <td>16</td> <td>20.5</td> </tr> <tr> <td><b>Med</b></td> <td>5.5</td> <td>7.5</td> <td>15.5</td> <td>20.75</td> </tr> <tr> <td><b>Total</b></td> <td>33</td> <td>48</td> <td>96</td> <td>123</td> </tr> <tr> <td><b>Raw total</b></td> <td>88.6</td> <td>48</td> <td>96</td> <td>123</td> </tr> </tbody> </table>		VL	L	M	H		rank score	rank score	rank score	rank score	<b>Mean</b>	5.5	8.0	16	20.5	<b>Med</b>	5.5	7.5	15.5	20.75	<b>Total</b>	33	48	96	123	<b>Raw total</b>	88.6	48	96	123		E1		Considering mean/median (allow from raw data) or totals
	VL	L	M	H																														
	rank score	rank score	rank score	rank score																														
<b>Mean</b>	5.5	8.0	16	20.5																														
<b>Med</b>	5.5	7.5	15.5	20.75																														
<b>Total</b>	33	48	96	123																														
<b>Raw total</b>	88.6	48	96	123																														
			<b>12</b>																															
	<b>Total</b>		<b>15</b>																															



Q5	Solution			Marks	Total	Comments																				
(a)(i)		Reduced	Not reduced	Total	M1	M1 17 or 19 placed correctly																				
	CT	<u>14</u>	3	<u>17</u>	M1	M1 23 placed correctly																				
	RT	12	7	<u>19</u>	M1	M1 14 or 9 placed correctly																				
	SL	<u>9</u>	13	<u>22</u>	A1	A1 all correct																				
	Total	35	<u>23</u>	58																						
						<b>4</b>																				
(ii)	<p><math>H_0</math>: Reduction or not of frequency of <b>panic attacks is independent of treatment</b> type</p> <p><math>H_1</math>: Reduction or not of frequency of <b>panic attacks is not independent of treatment</b> type</p> <p>1 tail 5%</p> <p>Expected freqs</p> <table border="1"> <thead> <tr> <th></th> <th>Reduced</th> <th>Not reduced</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>CT</td> <td>10.26</td> <td>6.74</td> <td>17</td> </tr> <tr> <td>RT</td> <td>11.47</td> <td>7.53</td> <td>19</td> </tr> <tr> <td>SL</td> <td>13.28</td> <td>8.72</td> <td>22</td> </tr> <tr> <td>Total</td> <td>35</td> <td>23</td> <td>58</td> </tr> </tbody> </table>					Reduced	Not reduced	Total	CT	10.26	6.74	17	RT	11.47	7.53	19	SL	13.28	8.72	22	Total	35	23	58	B1	oe
		Reduced	Not reduced	Total																						
	CT	10.26	6.74	17																						
	RT	11.47	7.53	19																						
	SL	13.28	8.72	22																						
	Total	35	23	58																						
	$\chi^2 = \sum \frac{(O-E)^2}{E}$ $= \frac{(14-10.26)^2}{10.26} + \frac{(3-6.74)^2}{6.74} + \dots + \frac{(13-8.72)^2}{8.72}$ $= 1.36 + \mathbf{2.07} + 0.02 + 0.04 + 1.38 + \mathbf{2.10}$				M1	One exp freq correct or method seen PI																				
	<p>ts <math>\chi^2 = 6.98</math> df = 2</p> <p>cv = 5.991      <math>p = 0.031 &lt; 0.05</math></p>				A1	All correct to at least 1 dp PI																				
	<p>6.98 &gt; 5.991    Reject <math>H_0</math>.</p> <p>Sig evidence to suggest that <b>reduction</b> or not of (frequency of ) <b>panic attacks/panic/attacks</b> is not independent of/associated with <b>treatment</b> type.</p>				m1	Numerator and denominator correct PI																				
					A1	awfw 6.8 – 7.1																				
				B1	df = 2 correct PI (4.605, 7.378, 9.21, 10.597)																					
				B1	cv correct or B1 B1 p value comp 0.05																					
				A1dep	<b>9</b> Conclusion correct dep ts/cv correct																					
				E1dep	Correct and in context dep ts/cv correct																					
(iii)	<p><b>For patients having Cognitive Therapy</b></p> <p>More (14) than expected (10.26) had reduced panic attacks <b>or</b></p> <p>Fewer (3) than expected (6.74) did not have reduced panic attacks</p>			B1	B1 ref to contributions to ts or largest																					
	<p><b>For patients having Sympathetic Listening</b></p> <p>Fewer (9) than expected (13.28) had reduced panic attacks <b>or</b></p> <p>More (13) than expected (8.72) did not have reduced panic attacks</p>			E1	$\frac{(O-E)^2}{E}$ or mention of																					
					<b>4</b> Exp/Obs freqs PI eg 'more than expected'																					
				B1	E1 correct identification of main sources of assoc in context (with or without back up)																					

For (iii)

- E1, E1 max for FT wrong initial table ONLY IF **clear** identification of main contribution to ts can be seen
- Subtract B1 if **ALL 3** CT, RT and SL included
- Subtract B1 if 'reduction' 'no reduction' only in explanation and panic attacks/panic/attacks **does not appear AT LEAST ONCE** ( just once is OK)

Q5	Solution	Marks	Total	Comments
(b)(i)	$H_0 \eta_d = 0$ $H_1 \eta_d > 0$ 1 tail test 5 % level  Differences sign ( After –Before)  A B C D E F G H I J K L - + + + + + . + . + + -  $t_s = 8+ / 2 -$  $B ( 10, 0.5 )$  $P( \leq 2 - ) = P( \geq 8 + ) = 0.0547 > 0.05$ Accept $H_0$  <b>No significant evidence to suggest that social competence/ score increases, on average, following the training.</b>	B1		oe Must be consistent with differences
		M1		Differences
		m1 A1		G and I discarded ts correct 8+ 2- PI
		M1		Correct comparison
		E1	6	Correct conclusion in context dep ts/cv correct disallow 'mean'
(ii)	The patients may be regarded as a random sample of such patients undertaking SCT	E1 E1	2	Random patients/sample idea In context of SCT
	<b>Total</b>		<b>25</b>	