



General Certificate of Education

Statistics 6380

SS03 Statistics 3

Mark Scheme

2009 examination – June series

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 meeting attended by all examiners and is the scheme which was used by them in this examination. The standardisation meeting ensures that the mark scheme covers the candidates' responses to questions and that every examiner understands and applies it in the same correct way. As preparation for the standardisation meeting each examiner analyses a number of candidates' scripts: alternative answers not already covered by the mark scheme are discussed at the meeting and legislated for. If, after this meeting, examiners encounter unusual answers which have not been discussed at the meeting 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 candidates' 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 to download from the AQA Website: www.aqa.org.uk

Copyright © 2009 AQA and its licensors. All rights reserved.

COPYRIGHT

AQA retains the copyright on all its publications. However, registered centres 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 centres 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 and abbreviations used in marking

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	MC	mis-copy
CAO	correct answer only	MR	mis-read
CSO	correct solution only	RA	required accuracy
AWFW	anything which falls within	FW	further work
AWRT	anything which rounds to	ISW	ignore subsequent work
ACF	any correct form	FIW	from incorrect work
AG	answer given	BOD	given benefit of doubt
SC	special case	WR	work replaced by candidate
OE	or equivalent	FB	formulae book
A2,1	2 or 1 (or 0) accuracy marks	NOS	not on scheme
-x EE	deduct x marks for each error	G	graph
NMS	no method shown	c	candidate
PI	possibly implied	sf	significant figure(s)
SCA	substantially correct approach	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. However, there are situations in some units where part marks would be appropriate, particularly when similar techniques are involved. Your Principal Examiner will alert you to these and details will be provided on the mark scheme.

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	Solutions	Marks	Total	Comments																						
1(a)	<table border="1"> <thead> <tr> <th>Rank x</th> <th>Rank y</th> </tr> </thead> <tbody> <tr><td>10</td><td>8</td></tr> <tr><td>9</td><td>6</td></tr> <tr><td>8</td><td>9</td></tr> <tr><td>7</td><td>10</td></tr> <tr><td>6</td><td>2</td></tr> <tr><td>5</td><td>5</td></tr> <tr><td>4</td><td>7</td></tr> <tr><td>3</td><td>4</td></tr> <tr><td>2</td><td>1</td></tr> <tr><td>1</td><td>3</td></tr> </tbody> </table>	Rank x	Rank y	10	8	9	6	8	9	7	10	6	2	5	5	4	7	3	4	2	1	1	3	M1		attempt at ranks inconsistent (can be reversed) SC M1M1 B2
	Rank x	Rank y																								
	10	8																								
	9	6																								
	8	9																								
	7	10																								
	6	2																								
	5	5																								
	4	7																								
	3	4																								
2	1																									
1	3																									
		M1																								
		A1		for 16 correct																						
	$r_s = 0.673$ (3 sf from calc)	B3	6	AWRT B2 0.67 B1 0.7 ft B2 from wrong ranks (small slip) No ranks seen, SC 0.67 B4 0.7 B3 alternative $d = 2, 3, 1, 3, 4, ., 3, 1, 1, 2$ $\sum d^2 = 54$ B1 $r_s = 1 - \frac{6 \times 54}{10 \times 99} = 0.673$ M1, A1																						
(b)	H_0 Rank orders of age and percentage body fat in females are independent.	B1		or equivalent																						
	H_1 Rank orders of age and percentage body fat in females are not independent – there is an association																									
	2 tail 10%																									
	$cv = \pm 0.5636$ $n = 10$	B1		for cv																						
test stat $r_s = 0.673$ $r_s > 0.5636$	M1		for comparison ts/cv SC Allow M1 0.593/0.5494 (pmcc)																							
Reject H_0 . Significant evidence at 10% level to suggest an association between rank orders of age and percentage body fat in females.(or positive association)	E1	4	correct and in context																							
	Total		10																							

SS03 (cont)

Q	Solutions	Marks	Total	Comments																														
<p>2.(a)</p>	<p>H_0 pop median/mean $\eta, \mu = 9$ H_1 pop median/mean $\eta, \mu > 9$ 1 tail 5% (d is result -9)</p> <table border="1" data-bbox="225 434 695 544"> <tr> <td>diff</td> <td>0.5</td> <td></td> <td>0.2</td> <td>0.6</td> <td></td> </tr> <tr> <td></td> <td></td> <td>-0.4</td> <td></td> <td></td> <td>-0.1</td> </tr> <tr> <td>rank</td> <td>4</td> <td>-3</td> <td>2</td> <td>5½</td> <td>-1</td> </tr> </table> <table border="1" data-bbox="225 577 603 687"> <tr> <td>diff</td> <td>0.7</td> <td></td> <td>0.8</td> </tr> <tr> <td></td> <td></td> <td>-0.6</td> <td></td> </tr> <tr> <td>rank</td> <td>7</td> <td>-5½</td> <td>8</td> </tr> </table> <p>$T_+ = 4 + 2 + \dots + 8 = 26\frac{1}{2}$ $T_- = 3 + 1 + 5\frac{1}{2} = 9\frac{1}{2}$</p> <p>Test stat $T = 9\frac{1}{2}$ $n = 8$ 1 tail 5% $cv = 6$ $T > 6$</p> <p>No significant evidence at 5% level to reject H_0. Conclude that there is no significant evidence to suggest that the average time to complete the task is greater than 9 minutes.</p>	diff	0.5		0.2	0.6				-0.4			-0.1	rank	4	-3	2	5½	-1	diff	0.7		0.8			-0.6		rank	7	-5½	8	<p>B1</p> <p>M1</p> <p>m1</p> <p>M1</p> <p>m1</p> <p>A1</p> <p>B1</p> <p>M1</p> <p>E1</p> <p>B1</p>	<p></p> <p></p> <p></p> <p></p> <p></p> <p></p> <p></p> <p></p> <p>9</p> <p>1</p> <p>10</p>	<p></p> <p>For differences (result -9) - ignore signs</p> <p>For ranks</p> <p>For ties</p> <p>For total attempted from ranks For one correct total</p> <p>For cv Comparison cv/ts (consistent)</p> <p>In context</p> <p>Disallow 'normally distributed'</p>
diff	0.5		0.2	0.6																														
		-0.4			-0.1																													
rank	4	-3	2	5½	-1																													
diff	0.7		0.8																															
		-0.6																																
rank	7	-5½	8																															
	Total		10																															

SS03 (cont)

Q	Solutions	Marks	Total	Comments
4(a)	$H_0 \eta_d = 0$	B1	6	Signs (allow signed differences) test stat correct and identified Binomial model used and probability attempted Comparison of Binomial probability with 0.05 (or 0.1) Identified correct critical region with probability given also M1m1
	$H_1 \eta_d \neq 0$ 2 tail 10%			
	Signs + . + + - - + +	M1		
	$5^+ / 2^-$ signs – test values	A1		
	Binomial (7, 0.5) model	M1		
	$P(\geq 5 +) = P(\leq 2 -) = 0.227 > 0.05$ or $P(\geq 5 +) = P(\leq 2 -) = 0.453 > 0.10$ one tail test	m1		
	Accept H_0 There is not sufficient evidence, at the 10% level, to suggest that the average cornea thickness differs between the normal eye and the eye with glaucoma.	E1		Interpretation in context
4(b)	$H_0 \eta_d = 0$		5	One tail – either way if consistent test stat identified ft incorrect ts from (a) Binomial model used and probability attempted Comparison of Binomial probability with 0.05 Identified correct critical region with probability given also M1m1 SC $n = 8$ in part(a) Allow part(b) $n = 13$ M1, M1 for $0.0112 < 0.05$
	$H_1 \eta_d > 0$ 1 tail 5%	B1		
	$10^+ / 2^-$ signs – test values	B1		
	Binomial (12, 0.5) model			
	$P(\geq 10 +) = P(\leq 2 -) = 0.0193 < 0.05$ one tail test	M1 m1		
	Reject H_0 . There is sufficient evidence, at the 5% level, to suggest that the average cornea thickness is greater for the normal eye than for the eye with glaucoma.	A1		
	Total		11	

SS03 (cont)

Q	Solutions	Marks	Total	Comments																												
5(a)	Ho Samples are taken from identical populations H ₁ Samples are not taken from identical populations – population average bottle cap productions differ for the 3 machines. 1 tail 1%	B1		or H ₀ $\eta_A = \eta_B = \eta_C = \eta_D$																												
	Ranks	B1		H ₁ at least two of $\eta_A, \eta_B, \eta_C, \eta_D$ do differ																												
	<table border="1"> <thead> <tr> <th>Machine A</th> <th>Machine B</th> <th>Machine C</th> <th>Machine D</th> </tr> </thead> <tbody> <tr> <td>2½</td> <td>15</td> <td>1</td> <td>8</td> </tr> <tr> <td>5</td> <td>16</td> <td>2½</td> <td>9</td> </tr> <tr> <td>10</td> <td>18</td> <td>4</td> <td>11</td> </tr> <tr> <td>14</td> <td>20</td> <td>6</td> <td>12</td> </tr> <tr> <td>17</td> <td>21</td> <td>7</td> <td>13</td> </tr> <tr> <td>19</td> <td></td> <td></td> <td></td> </tr> </tbody> </table>	Machine A	Machine B	Machine C	Machine D	2½	15	1	8	5	16	2½	9	10	18	4	11	14	20	6	12	17	21	7	13	19						
	Machine A	Machine B	Machine C	Machine D																												
	2½	15	1	8																												
	5	16	2½	9																												
	10	18	4	11																												
	14	20	6	12																												
	17	21	7	13																												
	19																															
T _A = 67½ T _B = 90 T _C = 20½ T _D = 53 n _A = 6 n _B = 5 n _C = 5 n _D = 5	M1		Totals																													
$\sum_{i=1}^m \frac{T_i^2}{n_i} = \frac{67.5^2}{6} + \frac{90^2}{5} + \frac{20.5^2}{5} + \frac{53^2}{5}$	m1		Method for $\sum_{i=1}^m \frac{T_i^2}{n_i}$																													
= 3025.225	m1		n _i correct																													
H = $\frac{12}{21 \times 22} \times 3025.225 - (3 \times 22)$ = 12.58	m1		test stat H																													
Critical value from $\chi^2_3 = 11.345$ H > 11.345	A1 B1 M1		12.3 – 12.9 3 df cv comparison cv/ts																													
Sig evidence to reject H ₀ and conclude that samples are not from identical populations	A1																															
Significant evidence at the 1% level to suggest that the population average bottle cap productions differ for the 4 machines. At least two machines have different averages	E1	12	Difference in context Mention of ‘at least two’																													
(b) Machine C had the lowest average rank score so it would seem likely that machine C is the lowest producer of bottle caps and therefore this machine is the most obvious one to suggest for replacement	B1		Machine C																													
	E1	2	Reason – must refer to the lowest average score or lowest production																													
	Total		14																													

SS03 (cont)

Q	Solutions	Marks	Total	Comments																		
<p>6(a)(i)</p>	<p>H_0 No association between place of occurrence of fall and sex of person who falls. H_1 Association exists between place of occurrence of fall and sex of person who falls.</p> <p>1 tail 1%</p> <table border="1" data-bbox="228 539 676 927"> <thead> <tr> <th>Expected freqs</th> <th>Male</th> <th>Female</th> </tr> </thead> <tbody> <tr> <td>Home</td> <td>1306.75</td> <td>355.25</td> </tr> <tr> <td>School, other</td> <td>31.45</td> <td>8.55</td> </tr> <tr> <td>Trade and serv</td> <td>72.34</td> <td>19.66</td> </tr> <tr> <td>Industrial etc</td> <td>127.37</td> <td>34.63</td> </tr> <tr> <td>Farm</td> <td>29.09</td> <td>7.91</td> </tr> </tbody> </table> $ts = \sum \frac{(O - E)^2}{E}$ $= \frac{(1269 - 1306.75)^2}{1306.75} + \frac{(393 - 355.25)^2}{355.25}$ $\dots + \frac{(6 - 7.91)^2}{7.91}$ <p>= 38.5</p> <p>df = 4 1% cv = 13.277</p> <p>ts > 13.277</p> <p>Significant evidence to reject H_0 and conclude that there is an association between place of occurrence of fall and sex of person who falls.</p>	Expected freqs	Male	Female	Home	1306.75	355.25	School, other	31.45	8.55	Trade and serv	72.34	19.66	Industrial etc	127.37	34.63	Farm	29.09	7.91	<p>B1</p> <p>M1 m1</p> <p>A1</p> <p>m1</p> <p>A1</p> <p>B1 B1 m1</p> <p>A1</p>	<p>10</p>	<p>For method for E for 3 correct</p> <p>All correct (1dp required minimum except 1307/355)</p> <p>ts sum with correct denominators</p> <p>for ts in range 34 - 43</p> <p>for df = 4 for cv for comparison ts/cv</p>
Expected freqs	Male	Female																				
Home	1306.75	355.25																				
School, other	31.45	8.55																				
Trade and serv	72.34	19.66																				
Industrial etc	127.37	34.63																				
Farm	29.09	7.91																				
<p>(ii)</p>	<p>Far fewer than expected females have ladder-related falls in the Industrial or Construction area and far more males than expected do have falls in this area. More females were observed to have falls at home than would be expected.</p>	<p>E1</p> <p>E1</p>	<p>2</p>	<p>Any two points made</p>																		

SS03 (cont)

Q	Solutions	Marks	Total	Comments												
6(a)(iii)	Many more males involved in ladder-related falls than females. Most falls occurred at home.	E1	1	Any one point made												
(b)(i)	<table border="1"> <thead> <tr> <th></th> <th>Male</th> <th>Female</th> </tr> </thead> <tbody> <tr> <td>Direct</td> <td>141</td> <td>29</td> </tr> <tr> <td>Transfer</td> <td>28</td> <td>6</td> </tr> <tr> <td>Other</td> <td>58</td> <td>22</td> </tr> </tbody> </table>		Male	Female	Direct	141	29	Transfer	28	6	Other	58	22	M1		% of 227 or % of 57
		Male	Female													
	Direct	141	29													
	Transfer	28	6													
Other	58	22														
		m1		4 correct (not necessarily integers)												
		A1	3	all correct and integers												
(ii)	<p>H_0 No association between admission category and sex of person who falls. H_1 Association exists between admission category and sex of person who falls.</p> <p>$t_s = 3.84$ $df = 2$ 5% $cv = 5.991$</p> <p>$3.84 < 5.991$ No significant evidence to reject H_0. Conclude that there is no evidence of association between admission category and sex of person who falls.</p>	B1														
		B1														
		M1														
		A1	4													
	Total		20													
	TOTAL		75													