



**General Certificate of Education (A-level)
June 2011**

Statistics

SS06

(Specification 6380)

Statistics 6

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 candidates' 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 candidates' 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 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 from: aqa.org.uk

Copyright © 2011 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 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)				
(i)	Each child <u>was allocated</u> to take the amoxicillin or the placebo in a <u>random</u> manner, as if by chance alone.	E1		Need idea of <u>allocation to drug/treatment</u> Disallow 'same chance' unless ref to being allocated Allow 'allocated with no pattern' or 'no order'
(ii)	A harmless substance that looks like real medication but <u>does not contain any drug</u> .	E1		Allow 'has no effect', 'fake' treatment, 'looks same but doesn't do same thing'
(iii)	Neither the medical staff involved nor the children and their parents know whether the child is receiving the amoxicillin or the placebo.	E1	3	Must make clear that neither group knows <u>which treatment</u>
(b)				
(i)	To avoid parents/children <u>anticipating</u> a slow recovery simply because no drug treatment is offered. All children will appear to be taking the same medication if the placebo is used.	E1		(b) answers <u>do not need</u> to be in context of this trial Child/parent 'Not knowing' which drug so won't know what to expect.
(ii)	To <u>avoid expectations</u> by children and parents or by medical staff, of success/failure of the treatment <u>outcome</u> due to a knowledge of whether an active drug or a placebo was given, influencing the outcome.	E1	2	Doctors and child/parent have no particular expectations due to knowledge of treatment Disallow 'Fair', 'reduces bias', 'reduces experimental error'
	Total		5	

Q	Solution	Marks	Total	Comments																								
2(a)	<p>H_0 pop mean diff $\mu_d = 0$ H_1 pop mean diff $\mu_d < 0$ 1 tail 5%</p> <p>$d = \text{Yorks} - \text{Surrey}$</p> <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>d</td> <td>-10.09</td> <td>-1.63</td> <td>1.35</td> <td>-1.13</td> <td>-2.67</td> </tr> <tr> <th></th> <th>F</th> <th>G</th> <th>H</th> <th>I</th> <th></th> </tr> <tr> <td>d</td> <td>3.75</td> <td>-8.23</td> <td>-11.43</td> <td>1.23</td> <td></td> </tr> </tbody> </table> <p>$\bar{d} = -3.21$ $s = 5.43$ $n = 9$ $t = \frac{-3.21 - 0}{\frac{5.43}{\sqrt{9}}} = -1.77$</p> <p>$df = 8$ $cv = 1.86$ $1.77 < 1.86$ oe</p> <p>Accept H_0 +/- and +/- comp only No significant evidence to suggest that mean prices are lower in Yorkshire.</p>		A	B	C	D	E	d	-10.09	-1.63	1.35	-1.13	-2.67		F	G	H	I		d	3.75	-8.23	-11.43	1.23		<p>B1</p> <p>M1</p> <p>m1</p> <p>M1</p> <p>m1</p> <p>A1</p> <p>B1</p> <p>B1</p> <p>A1</p>	<p>9</p>	<p>Condone μ or population mean/average Condone explained in words ref to average price/difference or pop average allow $\mu_d > 0$ if consistent H_1 must be consistent with their differences</p> <p>Differences – either way</p> <p>attempt to find \bar{d}, s</p> <p>Use of $\frac{s}{\sqrt{9}}$ ft</p> <p>Method for t ft</p> <p>(\pm) 1.77 (1.7 – 1.8) (\pm) 1.77 no method 5 marks</p> <p>df for <u>correct cv</u> or $p = 0.115$</p> <p>correct <u>conclusion in context</u> need ts/cv both correct</p>
	A	B	C	D	E																							
d	-10.09	-1.63	1.35	-1.13	-2.67																							
	F	G	H	I																								
d	3.75	-8.23	-11.43	1.23																								
(b)(i)	<p>(Differences between average prices for Yorkshire and Surrey are) <u>normally distributed</u>.</p>	E1	1	Allow <u>random and normally</u> dist																								
(ii)	<p><u>Sign test</u>. Not normal might well mean differences are not symmetrically distributed so sign test safer as <u>needs no assumption regarding distribution</u> of sample data.</p> <p>OR</p> <p>Wilcoxon Signed-Rank test No clear evidence differences are not symmetrically distributed so can <u>assume symmetry</u> and W S-R preferred to sign test as rank orders of differences taken into account.</p>	<p>B1</p> <p>E1</p> <p>OR</p> <p>(B1)</p> <p>(E1)</p>	<p>2</p>	<p>Sign test</p> <p>or ‘diffs/data do not need to be normally distributed’</p> <p>Wilcoxon Signed-Rank</p> <p>Allow’ W S-R does not require normally distributed data’</p> <p><u>Disallow W S-R more powerful</u> unless being compared to sign test</p>																								
Total			12																									

Q	Solution	Marks	Total	Comments																
3(a)	$T_A = 2857 \quad T_B = 2490 \quad T_C = 3190$ $n_A = 6 \quad n_B = 5 \quad n_C = 7$ $T = 8537$ $\sum \sum x_{ij}^2 = 4\,067\,243 \quad N = 18$ $\sum \frac{T_i^2}{n_i} = \frac{2857^2}{6} + \frac{2490^2}{5} + \frac{3190^2}{7}$ $= 4054156.7$ $SS_{\text{Methods}} = 4054156.7 - \frac{8537^2}{18}$ $= 5247.3$ $SS_{\text{Total}} = 4\,067\,243 - \frac{8537^2}{18}$ $= 18333.6$																			
		M1		SS for methods																
		M1		SS for total																
	<table border="1"> <thead> <tr> <th></th> <th>SS</th> <th>df</th> <th>ms</th> </tr> </thead> <tbody> <tr> <td>Methods</td> <td>5247.3</td> <td>2</td> <td>2623.6</td> </tr> <tr> <td>Error</td> <td>13086.3</td> <td>15</td> <td>872.4</td> </tr> <tr> <td>Total</td> <td>18333.6</td> <td>17</td> <td></td> </tr> </tbody> </table>		SS	df	ms	Methods	5247.3	2	2623.6	Error	13086.3	15	872.4	Total	18333.6	17		m1		Error SS ft (not -ve)
		SS	df	ms																
	Methods	5247.3	2	2623.6																
	Error	13086.3	15	872.4																
	Total	18333.6	17																	
		m1		Method for MS – both correct ft incorrect df																
	$F = \frac{2623.6}{872.4} = 3.01$	m1 A1		Method for F ft 2.8 – 3.2 3.01/in range with <u>no method</u> seen allow 6 marks (or p = 0.080)																
$H_0 \mu_A = \mu_B = \mu_C$ H_1 at least 2 of the means differ																				
$F_{15}^2 = 3.682 > 3.01$ <u>Accept H_0. There is no significant evidence of a difference in(mean) reading achievement scores for the 3 methods. Allow no difference in teaching methods.</u>	B1 B1 A1	9	df correct cv correct correct ts/cv and conclusion <u>in context</u>																	
(b)	Assumptions:																			
	Reading scores are <u>normally distributed</u> for each method The normal populations of reading scores have a <u>common variance</u>	E1 E1	2	Normal mentioned Explanations <u>in some sort of context (scores appears) in one of the comments here</u> Disallow ‘random’, ‘no interaction’																
	Total		11																	

Q	Solution	Marks	Total	Comments
4(a)	Total with misregistration = $9 + 12 + \dots = 81$	M1		<u>Total</u> misregistered attempt
	Total wafers inspected = 480	m1		<u>ft total</u> misregistered \div 480
	Estimate for $p = \frac{81}{480}$ (0.16875)	A1		0.168/0.169 awrt
	Warning 95% control limits			Mark if seen in either limits
sc2 n=8 n=480 B1M1	$\frac{81}{480} \pm 1.96 \times \sqrt{\frac{81/480 \times 399/480}{60}}$	B1 M1		Use of <u>1.96 and 3.09</u>
		m1		Use of $\frac{p(1-p)}{n}$ for
	= (0.074 , 0.264)	A1		n = 60 or 8 or 480 Correct expression
	Action 99.8% control limits			0.073/0.075 0.263/0.265 awrt to 2 sf
	$\frac{81}{480} \pm 3.09 \times \sqrt{\frac{81/480 \times 399/480}{60}}$			
	= (0.019 , 0.318)	A1	8	0.018/0.020 0.317/0.319 awrt to 2 sf
(b)	(i) $p = \frac{20}{60}$ (= 0.333)	M1		<u>finding p correctly</u> in any part
	Stop process immediately as p above upper action limit	E1		<u>E1 no ft</u> unless small arithmetic slip and all methods OK
	(ii) $p = \frac{16}{60}$ (= 0.267)			
	Take another sample immediately as p between upper warning and action limits	E1		
	(iii) $p = \frac{3}{60}$ (= 0.05)			
	Below lower warning limit. <u>Investigate to find out why misregistration proportions have improved</u> so much	E1		For 'below warning' or 'take another sample'
		E1	5	For investigation of improvement
			13	

Q	Solution	Marks	Total	Comments																
5(a)	<table border="1"> <thead> <tr> <th>p</th> <th>0.01</th> <th>0.05</th> <th>0.10</th> </tr> </thead> <tbody> <tr> <td>P(acc) R</td> <td>0.7397</td> <td>0.2146</td> <td>0.0424</td> </tr> <tr> <td>P(acc) S</td> <td>0.9639</td> <td>0.5535</td> <td>0.1837</td> </tr> <tr> <td>P(acc) T</td> <td>0.9691</td> <td>0.5119</td> <td>0.1259</td> </tr> </tbody> </table>	p	0.01	0.05	0.10	P(acc) R	0.7397	0.2146	0.0424	P(acc) S	0.9639	0.5535	0.1837	P(acc) T	0.9691	0.5119	0.1259			
	p	0.01	0.05	0.10																
	P(acc) R	0.7397	0.2146	0.0424																
	P(acc) S	0.9639	0.5535	0.1837																
P(acc) T	0.9691	0.5119	0.1259																	
<p>Plan R P(0) $n = 30$</p>	M1		method seen/any one prob correct																	
<p>Plan S P(≤ 1) $n = 30$</p>	M1 A1		1 correct 2 correct M1 can be gained in Plan R or Plan S solutions																	
<p>Plan T $p = 0.05$ $= 0.2774 + 0.3650 \times 0.6424$</p>	B1		correct prob seen (.6424 or .2774)																	
	M1		double sample attempt																	
	m1		$n = 25$ throughout																	
	A1	7																		
(b)	<p>Plot: any attempt to plot one correct all correct</p>	M1 A1 A1	3	Any one thro (0,1) One All correct by eye T 'in middle'																
(c)	(i) 0.30	B1	1	(0.28 – 0.32)																
	(ii) 0.77	B1	1	(0.74 - 0.79)																
(d)	<p>Plan S P(rej) = $\frac{1 - 0.9639}{0.05} = 0.0361$</p>	M1		Prob rejection																
	P(0,1) for $n = 7$ and $p = 0.0361$	m1		Binomial used correctly $n = 7$																
	<p>Prob(more than 1) = $1 - P(0,1)$ $= 1 - 0.9758$ $= 0.0242$</p>	m1 A1		(0.0240 – 0.0245)																
			4																	
Total			16																	

Q	Solution	Marks	Total	Comments																				
6 (a)(i)	$T_{Tv}=2322.1$ $T_{News}=2461.2$ $T_{Rad}=2338.5$ $n_{Tv} = 4$ $n_{News} = 4$ $n_{Rad} = 4$ $T_{Con}=1692.6$ $T_{Qual}=1970.4$ $T_{Cost}=1829.6$ $n_{Conv} = 3$ $n_{Qual} = 3$ $n_{Cost} = 3$ $T_{Health}=1629.2$ $n_{Health} = 3$ $T=7121.8$ $\sum \sum x_{ij}^2 = 4253083.0$ $N=12$																							
	Total SS																							
	$4253083.0 - \frac{7121.8^2}{12} = 26413.4$	M1		Total SS method – must have correct given values																				
	Advertising approach SS C																							
	$\frac{1692.6^2}{3} + \frac{1970.4^2}{3} + \frac{1829.6^2}{3} + \frac{1629.2^2}{3}$ $- \frac{7121.8^2}{12} = 23030.3$	M1		Adv approach SS method ft																				
	Media SS R																							
	$\frac{2322.1^2}{4} + \frac{2461.2^2}{4} + \frac{2338.5^2}{4} - \frac{7121.8^2}{12}$ $= 2889.4$	M1		Media SS method ft																				
	<table border="1"> <thead> <tr> <th></th> <th>SS</th> <th>df</th> <th>ms</th> </tr> </thead> <tbody> <tr> <td>Advert Appr</td> <td>23030.3</td> <td>3</td> <td>7676.8</td> </tr> <tr> <td>Media</td> <td>2889.4</td> <td>2</td> <td>1444.7</td> </tr> <tr> <td>Error</td> <td>493.7</td> <td>6</td> <td>82.3</td> </tr> <tr> <td>Total</td> <td>26413.4</td> <td>11</td> <td></td> </tr> </tbody> </table>		SS	df	ms	Advert Appr	23030.3	3	7676.8	Media	2889.4	2	1444.7	Error	493.7	6	82.3	Total	26413.4	11		M1 B1		Error SS ft (not –ve) Error df= 6 correct
		SS	df	ms																				
	Advert Appr	23030.3	3	7676.8																				
	Media	2889.4	2	1444.7																				
	Error	493.7	6	82.3																				
	Total	26413.4	11																					
	$H_0 \mu_{con} = \mu_{qual} = \mu_{cost} = \mu_{health}$ H_1 at least 2 of the means differ	M1 m1			Method for Ad/Med MS ft Method (dep prev M's) for F (either)																			
	$F = \frac{7676.8}{82.3} = 93.3(91-95)$	A1			Adv approach F correct																			
$F \frac{3}{6} = 9.78$ $93.3 > 9.78$ Reject H_0	B1			cv = 9.78																				
$H_0 \mu_{Tv} = \mu_{Newspaper} = \mu_{Radio}$ H_1 at least 2 of the means differ				sc df muddled Allow B1, B1 for 9.78, 10.925 in any part																				
$F = \frac{1444.7}{82.3} = 17.6(16.5-18.5)$	A1			Media F correct																				
$F \frac{2}{6} = 10.925$ $17.6 > 10.925$ Reject H_0	B1			cv = 10.925																				
There is significant evidence of a difference in(mean) sales for the 4 adv approaches and also for the 3 media types used.	A1 E1			Correct conclusions (both reject) <u>ts/cv correct</u> In <u>context ref to sales and advert & media</u>																				
			13																					

6(a) cont.	(ii) part (i) indicates a significant difference between at least two advertising approaches. <u>Means</u> are $\bar{X}_{Con}=564.2$ $\bar{X}_{Qual}=656.8$ $\bar{X}_{Cost}=609.9$ and $\bar{X}_{Health}=543.1$ <u>Quality clearly has the highest</u> mean sales. Chose Quality approach	M1		Use of means or totals for comparisons
		A1	2	Quality chosen
(b)(i)	There is <u>no interaction</u> between the two factors.	E1	1	For no interaction
(ii)	There is no particular reason why one advertising approach should <u>be more or less suited to one media type</u> rather than any other.	E1		General idea of assumption that each advertising approach has the same effect no matter what media type is used
		E1	2	Explained precisely
	Total		18	
	TOTAL		75	