AQA

ASSESSMENT and
QUALIFICATIONS
ALLIANCE

General Certificate of Education

Statistics 6380

SS03 Statistics 3

Mark Scheme

2005 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.

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	OE	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)

Application of Mark Scheme

No method shown:

Correct answer without working	mark as in scheme
Incorrect answer without working	zero marks unless specified otherwise

More than one method / choice of solution:

2 or more complete attempts, neither/none crossed out	mark both/all fully and award the mean mark rounded down
1 complete and 1 partial attempt, neither crossed out	award credit for the complete solution only

Crossed out work	do not mark unless it has not been replaced

Alternative solution using a correct or partially correct method	award method and accuracy marks as appropriate

SSO3

Q	Solution	Marks	Total	Comments
1(a) reversed 6 13 2 12 3 10.5 1 10.5 4 9 8 8 7 7 10 6 9 5 11 4 5 3 12 2 13 1	<table><tr><td>Team</td><td>Bo</td><td>Mil</td><td>Ka</td><td>Se</td><td>Ca</td></tr><tr><td>x rank</td><td>8</td><td>12</td><td>11</td><td>13</td><td>10</td></tr><tr><td>y rank</td><td>1</td><td>2</td><td>3½</td><td>3½</td><td>5</td></tr><tr><td>Team</td><td>Tex</td><td>Cle</td><td>NY</td><td>Mi</td><td>Oa</td></tr><tr><td>x rank</td><td>6</td><td>7</td><td>4</td><td>5</td><td>3</td></tr><tr><td>y rank</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr><tr><td>Team</td><td>Ch</td><td>Bal</td><td>To</td><td></td><td></td></tr><tr><td>x rank</td><td>9</td><td>2</td><td>1</td><td></td><td></td></tr><tr><td>y rank</td><td>11</td><td>12</td><td>13</td><td></td><td></td></tr></table> r_s = - 0.787 (3 sf from calc) sc -0.78/9 no method M1M1A1B1 subtract 1 mark if not negative	M1 M1 A1 B3	6	Attempt at ranks Tied ranks Reverse ranks acceptable Inconsistent ranks gain M1, 1, 1 Alternative: d = 7, 10,7½, 9½, 5, 0, 0, 4, 4, 7, 2, 10, 12 $\sum d^2$ =649½ B1 $r_s = 1 - \frac{6 \times 649.5}{13 \times 168} = -0.784$ M1, A1
(b)	H_o Rank orders number of home runs and batting averages are independent. H_1 Rank orders of home runs and batting averages are not independent. 2 tail 1% cv = ± 0.6978 test stat r_s = - 0.787 r_s < cv or $\lvert r_s \rvert$ > $\lvert$ cv $\rvert$ Reject H_o Significant evidence at 1% level to suggest an association between rank orders of number of home runs and batting average. [Results suggests that teams that score more home runs tend to have lower batting averages.]	B1 B1 M1 A1 E1	5	Generous For cv – ignore sign (range allowed B0M1) For comparison ts/cv r_s = - 0.787 or - 0.784 can be implied +/- cv/ts comparison M0 A1E1 poss Ft – not if contradictory in context – need not be correct but ft in context
	Total		**11**	
2	H_o pop median, η = 42 H_1 pop median, η > 42 1 tail 10% Signs - + + + + - + - + - + + n = 12 ts = 8+/ 4 – Binomial model B (12, 0.5) P ($\geq$ 8 +) = P($\leq$ 4 -) = 0.1938 > 0.10 for one tail test Accept H_o . There is insufficient evidence, at the 10% level, to suggest that the median is greater than 42.	B1 M1 A1 M1 M1 A1 E1	7	Clear η required or use of 'population' Signs or differences for 8+ and 4- Binomial model used and probability attempted and seen Comparison of Binomial probability 0.1938 (or awrt 0.19 with 0.10) Alternative – cr identified as {9,10,11,12} with prob 0.0730 used Interpretation in context – need not be correct but ft in context
	Total		**7**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
3(a)(i)	H_0 $\eta_{\text{difference}} = 0$ H_1 $\eta_{\text{difference}} < 0$ 2 tail 5%	B1		H_0 $\mu_{\text{difference}} = 0$ or $\eta_{1=}\eta_2$ H_1 $\mu_{\text{difference}} < 0$ or $\eta_1 < \eta_2$ etc or in words
	<table><tr><td>Carpet Type</td><td>Difference 1 – 2</td><td colspan="2">Rank</td></tr><tr><td></td><td></td><td>-</td><td>+</td></tr><tr><td>A</td><td>- 4</td><td>2</td><td></td></tr><tr><td>B</td><td>- 14</td><td>7</td><td></td></tr><tr><td>C</td><td>- 15</td><td>8</td><td></td></tr><tr><td>D</td><td>- 11</td><td>6</td><td></td></tr><tr><td>E</td><td>- 45</td><td>9</td><td></td></tr><tr><td>F</td><td>- 7</td><td>5</td><td></td></tr><tr><td>G</td><td>+ 6</td><td></td><td>4</td></tr><tr><td>H</td><td>- 5</td><td>3</td><td></td></tr><tr><td>I</td><td>+ 3</td><td></td><td>1</td></tr></table>	M1 m1		For differences – ignore signs (1 – 2 or 2 – 1) For ranks (1 = smallest |diff|)
	Rank totals $T_- = 40$ $T_+ = 5$ Test stat $T = 5$ critical value = 6 $T < cv$ Reject H_0 There is significant evidence to suggest that there is a difference in the average fixing times for the two adhesives.	m1 A1 B1 M1 A1 E1	 9	For total of + / - ranks (even if ranked incorrectly) For cv (range allowed B0M1) For comparison ts/cv must be in context and refer to average
(ii)	Cost of adhesive, size of carpet Long term reliability Method of application etc	E1	 1	Disallow 'carpets the same' allow 'surface carpet laid on' any relevant comment
(b)	By using each adhesive on each type of carpet, experimental error is reduced and any difference in fixing time detected is due to adhesive used.	B1 E1	 2	For idea of reducing experimental error For clear explanation in context disallow 'more accurate' unless in explanation
	Total		**12**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
4(a)(i)	H_o Returns status is independent of when order was placed H_1 Returns status is not independent of when order was placed 1 tail 5%	B1		Only H_o sufficient – must be clear on independence/no association Condone use of return/order in H_o even if not worded completely correctly
	<table><tr><td></td><td>**Prom**</td><td>**Sale**</td><td>**Stand**</td></tr><tr><td>**No items returned**</td><td>49.2</td><td>32.8</td><td>82</td></tr><tr><td>**Some items returned**</td><td>10.8</td><td>7.2</td><td>18</td></tr></table>	M1 m1		E method for 1 correct For all E correct ft arithmetic error
	ts = $\sum \dfrac{(O-E)^2}{E}$	m1		ts sum with correct denominators disallow Yates' correction
	$= \dfrac{5.8^2}{49.2} + \dfrac{4.8^2}{32.8} + \dfrac{1^2}{82} + \dfrac{5.8^2}{10.8} + \dfrac{4.8^2}{7.2} + \dfrac{1^2}{18}$	A1		For ts in range 7.70 - 7.80 awfw
	$= 7.77$			
	cv df = 2 5% cv = 5.991	B1		For cv (4.605, 9.210, 7.378 B0m1)
	ts > 5.991	m1		For comparison ts/cv
	Reject H_o Sig evidence to suggest returns status is not independent of when order was placed	A1	8	**No context** required Can imply B1 for H_o
(ii)	Orders placed during a Promotion are much **more likely** to result in **no items** being returned whereas orders placed during the Sale are far **less likely** to result in **no items** being returned.	E1,1	2	E1 if inconsistent/inadequate Or greater likelihood of some returns for orders placed during the Sale etc. Must mention promotion and sale for E2. E1 generous if in context
	Total		**10**	

SS03 (cont)

Q	Solution				Marks	Total	Comments
4(b)(i)		**First time**	**Repeat**		B1		For raw data numbers correct For contingency table headings, data all correct
	No items returned	20 (17.64)	43 (45.36)				
	Some items returned	8 (10.36)	29 (26.64)			2	(E values bracketed)
(ii)	H_o Returns status is independent of type of customer H_1 Returns status is not independent of type of customer 1 tail 5%				B1		Only H_o sufficient – must be clear on independence/no association Condone use of order/customer in H_o even if not worded completely correctly
					M1		For E values method
	$\text{ts} = \sum \dfrac{(|O-E|-0.5)^2}{E} =$ $\dfrac{1.86^2}{17.64}+\dfrac{1.86^2}{45.36}+\dfrac{1.86^2}{10.36}+\dfrac{1.86^2}{26.64}$ = 0.736				M1 m1 A1		For ts for Yates' corr attempt include: O-E-0.5 or (O-E)² - 0.5 For ts 0.730 -0.750 awft Condone if not 3 sf
	cv df = 1 5% cv =3.841 ts < 3.84				B1 m1		For cv CAO For comparison ts/cv ft
	Accept H_o No sig evidence to doubt that returns status is independent of type of customer. New customers are no more/less likely to return items.				A1	8	Allow A1 if H_o fully correct and stated earlier In context
				Total		**10**	

SS03 (cont)

Q	Solution	Marks	Total	Comments

5(a) B1 Only if μ	H_0 Samples from identical populations
	H_1 Samples not from identical
	     populations 5% sig level
		B1
		
		B1
				Or H_0 $\eta_A = \eta_B = \eta_C$
				  H_1 at least two of η_A, η_B, η_C
				   do differ
				B1,0 if no 'population' used
				H_0 No difference ………
				H_1 Difference…………
				In context used B1,0(B1 generous)

	Ranks 	Fertiliser A	Fertiliser B	Fertiliser C
---	---	---
1	9	3
2	11	4
5	13	7
6	15	8
10	16	12
	17	14		M1 A2		A1 $\geq$ 10 A2 all correct
	$T_A = 24$ $T_B = 81$ $T_C = 48$ $n_A = 5$ $n_B = 6$ $n_C = 6$	m1 A1		Totals Any one correct
	$\sum_{i=1}^{m} \frac{T_i^2}{n_i} = \frac{24^2}{5}+\frac{81^2}{6}+\frac{48^2}{5} = 1592.7$	m1		
	$H = \frac{12}{17 \times 18} \times 1592.7 - (3 \times 18) = 8.46$	m1 A1		ft for test stat H with previous result substituted AWFW 8.40 – 8.50
	Critical value from $\chi_2^2 = 5.99$ H > 5.99	B1 M1		For cv (range allowed B0M1)
	Sig evidence to reject H_0 and conclude that samples are not from identical populations	A1	13	Can imply B1,B0 at start of solution
(b)(i)	There is significant evidence that at least two of the median yields (from plants fed with Fertilisers A, B or C) do differ. Medians 26 , 42, 28½	E1 E1	2	Ft for E1 Difference in context Mention of 'at least two'
(ii)	It would appear that those plants that were fed Fertiliser B produced a significantly higher yield on average.	B1	1	Identification of B
	Total		**16**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
6(a)	H_0 Samples are taken from identical populations H_1 Samples are not taken from identical populations 2 tail 5%	B1		Or H_0 $\eta_H = \eta_C$ H_1 $\eta_H \neq \eta_C$ need both or population average used – need both
	$U_H = 66 - \dfrac{7 \times 8}{2} = 38$	M1		For attempt at U
	$U_C = 70 - \dfrac{9 \times 10}{2} = 25$	A1		For either correct
	$U = 25$ cv = 13 for n = 7, m = 9 5% $U > 13$ Accept H_0 No significant evidence of any difference between the two techniques.	B1 B1 M1 A1	 7	For correct U and consistent cv For cv (range allowed B0M1) For comparison ts/cv
(b)	A Type II error would be to conclude that H_0 is true, that is there is no difference between the two techniques, when in fact H_0 is not true and the techniques do differ	B1 E1	 2	Concept of Type II correct In context
	Total		**9**	
	Total		**75**	