Version 1.0:01.08

ASSESSMENT and QUALIFICATIONS ALLIANCE

General Certificate of Education

Statistics 6380

SS03 Statistics 3

Mark Scheme

2008 examination - January 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.

Set and published by the Assessment and Qualifications Alliance.

The Assessment and Qualifications Alliance (AQA) is a company limited by guarantee registered in England and Wales (company number 3644723) and a registered charity (registered charity number 1073334).
Registered address: AQA, Devas Street, Manchester M15 6EX *Dr Michael Cresswell Director General*

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.

SS03

Q	Solution	Marks	Total	Comments
1	H_o pop median = £5.60 H_1 pop median ≠ £5.60 2 tail 10%	B1		
	Signs - + + + + + + + + + - +	M1		
	test statistic 2^- / 10^+ $n = 12$	A1		ts correct
	$P(\leq 2^-) = 0.0193$ or $P(\geq 10^+) = 0.0193$	M1		Bin model seen to be used
	$0.0193 < 0.05$ for 2 tailed test at 10%	m1		Comparison of correct B(12, 0.5) prob with 0.05 or use of identified cv
	Significant evidence at 10% level to reject H_o	A1		
	There is significant evidence to suggest that the median weekly amount of pocket money given to 14 year-old children living in Brighton has changed (increased) since 2003	E1	7	
	Total		**7**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
2	H_o Samples are taken from identical populations	B1		Hypotheses referring to population averages also acceptable
	H_1 Samples are not taken from identical populations (rugby players have higher average scores)	B1		
	1 tail 5%			
	Golfers ranks / **Rugby players ranks**: 1 / 4; 2 / 6; 3 / $7\frac{1}{2}$; 5 / 10; $7\frac{1}{2}$ / 13; 9 / 14; 11 / 15; 12 /	M1		Attempt at Mann-Whitney; ranks as one group
		m1		For ties
	$T_G = 1 + 2 + ... + 12 = 50.5$ $T_R = 4 + 6 + ... + 15 = 69.5$	M1		For total attempt
	$U_G = 50.5 - \frac{8\times 9}{2} = 14.5$ $U_R = 69.5 - \frac{7\times 8}{2} = 41.5$	M1		For U formula correct
	Test statistic U = 14.5	B1		Correct/relevant cv used
	$n = 8$, $m = 7$, cv = 13	M1		
	U = 14.5 > 13	A1		
	Accept H_o No significant evidence at the 5% level to suggest that the average test score is higher for rugby players	E1	10	In context
	Total		**10**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
3(a)	H_0 No association between type of victim and type of offence			
	H_1 Association exists between type of victim and type of offence 1 tail 5%	B1		Independent / not independent: allow B1 once only
	Expected frequencies:	M1		E method (1dp allowed)
		m1		for 3 correct
		m1		for all E correct (SC2 if integers)
	$ts = \sum \frac{(O-E)^2}{E}$			
	$= \frac{(112-126.31)^2}{126.31} + \frac{(108-93.69)^2}{93.69} + \ldots$	m1		ts sum with correct denominators
	$= 8.013$	A1		ts in range 7.80 – 8.20
	df = 2 5% cv = 5.991	B1		For cv
	ts > 5.991	M1		For comparison ts/cv
	Reject H_o	A1		
	Significant evidence to suggest an association exists between type of victim and type of offence. Individuals much more likely to suffer arson / business much more likely to suffer robbery etc	E1	10	Any sensible interpretation in context

	Individual	Business
Robbery	126.31	93.69
Burglary	138.94	103.06
Arson	36.75	27.25

SS03 (cont)

Q	Solution	Marks	Total	Comments
3(b)(i)	Expected frequencies: |	**Under 25 years**	**25 years and over** | | **Aggravated** | 3.375 | 5.625 | | **Simple** | 11.625 | 19.375 | | **Intimidation** | 18 | 30 |	M1 m1 m1	 3	E method for 3 correct for all E correct (SC2 if integers)
(ii)	Pooling necessary because the expected frequency (3.375) for 'Under 25 years' 'Aggravated' assault is below 5	E1	1	
(iii)	2 assault categories should be pooled – both the same 'type' of offence: assault	E1	1	
(iv)	|	**Under 25 years**	**25 years and over** | | **Assaults – simple/aggravated** | 15 | 25 | | **Intimidation** | 18 | 30 |			
	H_o No association between age of offender and type of offence H_1 Association exists between age of offender and type of offence 1 tail 5%	B1		
	ts = $\sum \frac{(\lvert O-E\rvert - 0.5)^2}{E}$ = $\frac{1.5^2}{15}+\frac{1.5^2}{25}+\frac{1.5^2}{18}+\frac{1.5^2}{30}$	M1 M1		For ts correct denominators For Yates' correction
	= 0.44	A1		For ts 0.2 – 0.50 (SC2 ts = 0.782)
	df = 1 5% cv = 3.841	B1		For cv
	ts < 3.841	M1		For comparison ts/cv
	Accept H_o No significant evidence to suggest an association between age of offender and type of offence	A1	7	In context
	Total		**22**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
4(a)	H_o pop median/mean diff $\eta_d = 0$	B1		
	H_1 pop median/mean diff $\eta_d < 0$ 1 tail 5% (d is 2003 – 1999)	B1		Consistent with differences
	diff -5.4 -3.2 -3.8 -4.2 -2.4 **rank** 10 6 8 9 3	M1		For differences
	-2.1 -3.1 +0.3 -2.8 +3.4 2 5 1 4 7	M1		For ranks
	$T_+ = 1 + 7 = 8$	m1		For total
	$T_- = 10 + 6 + \ldots + 4 = 47$	A1		For one correct total
	ts T = 8 $n = 10$ cv =11	B1		For cv
	T < 11	M1		Comparison cv/ts
	Significant evidence at 5% level to reject H_o and conclude that average teenage conception rate has decreased between 1999 and 2003	E1	9	In context
(b)	A matched pairs design eliminates individual differences by comparing conception rates in the same regions for the two years. This means that any	B1		General idea of matched pairs reducing experimental error
	particular regional differences will not affect the comparisons and so a difference is more likely to be detected if one exists	E1	2	In context
(c)	A Type I error is when a correct H_0 is rejected. In this case it would mean that	B1		
	we conclude that the average conception rate has decreased when, in fact, it has not	E1	2	
	Total		**13**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
5(a)				
	Team A B C D E ***x* rank** 1 2 3 4 5 ***y* rank** 2 5 4 6 7 **Team** **F** **G** **H** **I** **J** **K** ***x* rank** 6 7 8 $9\frac{1}{2}$ $9\frac{1}{2}$ 11 ***y* rank** 10 3 11 9 1 8	M1 M1 A1		Attempt at ranks (can be reversed) For 8 correct
	$r_s =$ 0.355 (3sf from calc)	B3	6	**Or** $d = 1, 3, 1, 2, 2, 4, 4, 3, \frac{1}{2}, 8\frac{1}{2}, 3$ $\sum d^2 = 141\frac{1}{2}$ B1 $r_s = 1 - \frac{6\times 141\frac{1}{2}}{11\times 120} = 0.357$ M1A1 SC4 0.36 SC4 0.318
(b)	H_0 Rank orders of gate receipts and player costs are independent H_1 Rank orders of gate receipts and player costs are not independent – there is a positive association 1 tail 10%	B1		
	cv = 0.4182	B1		For cv
	ts $r_s =$ 0.355	M1		For comparison ts/cv
	r_s < 0.4182	A1		$r_s =$ 0.355 or 0.357
	Accept H_0 No significant evidence at 10% level to suggest a positive association between rank orders of gate receipts and player costs	E1	5	In context
	Total		**11**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
6	H_o Samples are taken from identical populations	B1		or H_0 $\eta_{Normal} = \eta_{Depres} = \eta_{MildAlz}$
	H_1 Samples are not taken from identical populations – population average recall scores differ	B1		H_1 at least two of $\eta_{Normal}, \eta_{Depres}, \eta_{MildAlz}$ differ
	1% 1 tail			
	Ranks			
	Normal: 8, 14, 15, 16, 17, 18, 19 / Depression: 5, 9, 10, 11, 12, 13 / Mild Alzheimer's: 1, 2, 3, 4, 6, 7	M1		Ranks
		m1		At least 12 correct
	$T_{Normal} = 107$ $T_{Depres} = 60$ $T_{MildAlz} = 23$	m1		Totals
	$n_{Normal} = 7$ $n_{Depres} = 6$ $n_{MildAlz} = 6$	A1		Any one correct
	$\sum_{i=1}^{m} \frac{T_i^2}{n_i} = \frac{107^2}{7} + \frac{60^2}{6} + \frac{23^2}{6} = 2323.74$	m1		
	$H = \frac{12}{19 \times 20} \times 2323.74 - (3 \times 20) = 13.38$	A1		ts correct 13.0 – 13.8
	Critical value from $\chi^2_2 = 9.210$	B1		
	$H > 9.210$	M1		
	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 recall scores differs for the three categories of adults: at least two of the averages differ. It appears that those adults with Mild Alzheimer's disease have a significantly lower average recall score than those who have normal memory function	E1	12	Difference in context Mention of 'at least two' or a significant difference between scores for Mild Alzheimer's and those with normal memory function
	Total		**12**	
	TOTAL		**75**	