Version 1.0:01.08



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.

Further copies of this Mark Scheme are available to download from the AQA Website: www.aqa.org.uk

Copyright © 2008 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.

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

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								
В	mark is independent of M or m marks and is for method and accuracy								
Е	mark is for explanation								
$\sqrt{\text{or ft or F}}$	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_0 pop median = £5.60 H_1 pop median \neq £5.60 2 tail 10%	B1		
	Signs - + + + + + + + + + + + + + + + + + +	M1		
	test statistic $2^{-}/10^{+}$ $n = 12$	A1		ts correct
	$P(\le 2^-) = 0.0193 \text{ or } P(\ge 10^+) = 0.0193$	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	

2 H _o Samples are taken from identical populations H _I Samples are not taken from identical populations (rugby players have higher average scores) 1 tail 5%	Q	Solu	ıtion	Marks	Total	Comments
populations (rugby players have higher average scores) 1 tail 5% Golfers ranks Rugby players ranks M1 1	2	populations		B1		*
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		populations (rugby p		B1		
$\begin{array}{ c c c c c c }\hline & ranks \\ \hline 1 & 4 & \\ \hline 2 & 6 & \\ \hline 3 & 7 \frac{1}{2} & \\ \hline 5 & 10 & \\ \hline 7 \frac{1}{2} & 13 & \\ \hline 9 & 14 & \\ \hline 11 & 15 & \\ \hline 12 & \\ \hline \\ T_G = 1 + 2 + + 12 = 50.5 & \\ T_R = 4 + 6 + + 15 = 69.5 & \\ \hline U_G = 50.5 - \frac{8 \times 9}{2} = 14.5 & \\ U_R = 69.5 - \frac{7 \times 8}{2} = 41.5 & \\ \hline Test statistic U = 14.5 & \\ \hline n = 8 \ , m = 7 \ , cv = 13 & \\ \hline u = 14.5 > 13 & \\ \hline Al & \\ \hline Accept H_o & \\ No significant evidence at the 5\% level to & E1 & 10 & In context \\ \hline \end{array}$		_				
$\begin{array}{ c c c c c c }\hline 2 & 6 & \\\hline 3 & 7\frac{1}{2} \\\hline 5 & 10 \\\hline 7\frac{1}{2} & 13 \\\hline 9 & 14 \\\hline 11 & 15 \\\hline 12 \\\hline \\\hline \\\hline T_G = 1 + 2 + + 12 = 50.5 \\\hline T_R = 4 + 6 + + 15 = 69.5 \\\hline \\\hline U_G = 50.5 - \frac{8 \times 9}{2} = 14.5 \\\hline U_R = 69.5 - \frac{7 \times 8}{2} = 41.5 \\\hline \\\hline Test statistic U = 14.5 \\\hline \\\hline N = 8, m = 7, cv = 13 \\\hline \\\hline U = 14.5 > 13 \\\hline \\\hline A1 \\\hline \\\hline A2 \\\hline \\\hline A2 \\\hline \\\hline A3 & 7\frac{1}{2} \\\hline \\\hline A1 \\\hline \\\hline \\\hline A2 \\\hline \\\hline \\\hline A3 & 7\frac{1}{2} \\\hline \\\hline \\\hline B1 \\\hline $		Golfers ranks				
$\begin{array}{ c c c c c c c c c }\hline 3 & 7\frac{1}{2} \\ \hline 5 & 10 \\ \hline 7\frac{1}{2} & 13 \\ \hline 9 & 14 \\ \hline 11 & 15 \\ \hline 12 \\ \hline \\ T_G = 1 + 2 + + 12 = 50.5 \\ T_R = 4 + 6 + + 15 = 69.5 \\ \hline \\ U_G = 50.5 - \frac{8 \times 9}{2} = 14.5 \\ \hline U_R = 69.5 - \frac{7 \times 8}{2} = 41.5 \\ \hline \\ Test statistic U = 14.5 \\ \hline \\ Test statistic U = 14.5 \\ \hline \\ M1 & Correct/relevant cv used \\ \hline \\ n = 8 \ , m = 7 \ , cv = 13 \\ \hline \\ U = 14.5 > 13 \\ \hline \\ Accept H_o \\ No significant evidence at the 5\% level to \\ \hline E1 & 10 & In context \\ \hline \end{array}$		1	4	M1		Attempt at Mann-Whitney; ranks as one
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		2				group
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		3	$7\frac{1}{2}$			
$\begin{array}{ c c c c c c c c c c c c c c c c c c c$			10			
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		$7\frac{1}{2}$	13	m1		For ties
$T_{G} = 1 + 2 + + 12 = 50.5$ $T_{R} = 4 + 6 + + 15 = 69.5$ $U_{G} = 50.5 - \frac{8 \times 9}{2} = 14.5$ $U_{R} = 69.5 - \frac{7 \times 8}{2} = 41.5$ $Test statistic U = 14.5$ $n = 8, m = 7, cv = 13$ $U = 14.5 > 13$ $Accept H_{o}$ No significant evidence at the 5% level to E1 10 In context						
$T_{\rm G}=1+2++12=50.5$ $T_{\rm R}=4+6++15=69.5$ $U_{\rm G}=50.5-\frac{8\times 9}{2}=14.5$ $U_{\rm R}=69.5-\frac{7\times 8}{2}=41.5$ $Test statistic U=14.5$ $n=8, m=7, {\rm cv}=13$ $U=14.5>13$ $Accept H_{\rm o}$ No significant evidence at the 5% level to E1 M1 For total attempt For U formula correct Correct/relevant cv used			15			
$T_{R} = 4 + 6 + + 15 = 69.5$ $U_{G} = 50.5 - \frac{8 \times 9}{2} = 14.5$ $U_{R} = 69.5 - \frac{7 \times 8}{2} = 41.5$ $Test statistic U = 14.5$ $n = 8, m = 7, cv = 13$ $U = 14.5 > 13$ $Accept H_{o}$ No significant evidence at the 5% level to $E1$ $In context$		12				
$T_{R} = 4 + 6 + + 15 = 69.5$ $U_{G} = 50.5 - \frac{8 \times 9}{2} = 14.5$ $U_{R} = 69.5 - \frac{7 \times 8}{2} = 41.5$ $Test statistic U = 14.5$ $n = 8, m = 7, cv = 13$ $U = 14.5 > 13$ $Accept H_{o}$ No significant evidence at the 5% level to $E1$ $In context$		$T_{\alpha} = 1 + 2 + + 12 =$	= 50.5	M1		For total attempt
$U_{\rm G} = 50.5 - \frac{8 \times 9}{2} = 14.5$ $U_{\rm R} = 69.5 - \frac{7 \times 8}{2} = 41.5$ $Test statistic U = 14.5$ $n = 8, m = 7, cv = 13$ $U = 14.5 > 13$ $Accept H_{\rm o}$ No significant evidence at the 5% level to $E1$ $M1$ $Correct/relevant cv used$ $A1$				1411		Tor total attempt
$U_{R} = 69.5 - \frac{7 \times 8}{2} = 41.5$ $Test statistic U = 14.5$ $n = 8, m = 7, cv = 13$ $U = 14.5 > 13$ $Accept H_{o}$ No significant evidence at the 5% level to E1 10 In context		R	07.0			
$U_{R} = 69.5 - \frac{7 \times 8}{2} = 41.5$ $Test statistic U = 14.5$ $n = 8, m = 7, cv = 13$ $U = 14.5 > 13$ $Accept H_{o}$ No significant evidence at the 5% level to E1 10 In context		8×9	14.5	3.41		
Test statistic $U = 14.5$ $n = 8, m = 7, cv = 13$ $U = 14.5 > 13$ Accept H_o No significant evidence at the 5% level to $E1$ Correct/relevant cv used $A1$ $A1$ $E1$ In context		$U_{\rm G} = 50.5 - \frac{1}{2} = \frac{1}{2}$	M1		For U formula correct	
Test statistic $U = 14.5$ $n = 8, m = 7, cv = 13$ $U = 14.5 > 13$ Accept H_0 No significant evidence at the 5% level to $E1$ $D = 14.5 > 13$ $E1$ $E1$ $E1$ $E1$ $E1$ $E1$ $E1$ $E1$		11 - 60 5 7×8 -				
n=8, $m=7$, $cv=13$ M1 $U=14.5>13$ A1 Accept H _o No significant evidence at the 5% level to E1 10 In context		$O_R = 09.3 = {2}$	41.5			
n=8, $m=7$, $cv=13$ M1 $U=14.5>13$ A1 Accept H _o No significant evidence at the 5% level to E1 10 In context			_	.		
$U = 14.5 > 13$ A1 $Accept H_o$ No significant evidence at the 5% level to $E1 \qquad 10 \qquad \text{In context}$		Test statistic $U = 14$.5	B1		Correct/relevant cv used
$U = 14.5 > 13$ A1 $Accept H_o$ No significant evidence at the 5% level to $E1 \qquad 10 \qquad \text{In context}$		n = 8 m = 7 cv = 1	3	M1		
Accept H_o No significant evidence at the 5% level to $E1$ 10 In context		n = 0, $m = 1$, $cv = 1$	S	1411		
No significant evidence at the 5% level to E1 10 In context		U = 14.5 > 13		A1		
No significant evidence at the 5% level to E1 10 In context						
				F.1	10	
NITO OPEN THAT THE AVELAGE TEXT SCHIETS				El	10	In context
higher for rugby players		higher for rughy play	ige test score is vers			
Total 10		inglier for ragely play			10	

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: Individual Business	M1		E method (1dp allowed)
		m1		for 3 correct
	Robbery 126.31 93.69 Burglary 138.94 103.06 Arson 36.75 27.25	m1		for all E correct (SC2 if integers)
	$ts = \sum \frac{(O-E)^2}{E}$			
	$= \frac{(112 - 126.31)^2}{126.31} + \frac{(108 - 93.69)^2}{93.69} + \dots$	m1		ts sum with correct denominators
	126.31 93.69 = 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

Q	Solution	Marks	Total	Comments
3(b)(i)	Expected frequencies:	M1		E method
	Under 25 years 25 years and over	m1		for 3 correct
	Aggravated 3.375 5.625 Simple 11.625 19.375 Intimidation 18 30	m1	3	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)	Assaults – 15 25 years and over 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		
	$ts = \sum \frac{(O - E - 0.5)^2}{E} = \frac{1.5^2}{15} + \frac{1.5^2}{25} + \frac{1.5^2}{18} + \frac{1.5^2}{30}$	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	

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

Q	Solution							Marks	Total	Comments
5(a)								2720223		
	Team	A	В	С	D	Е				
	x rank	1	2	3	4	5		M1		Attempt at ranks (can be reversed)
	y rank	2	5	4	6	7		M1		For 8 correct
	Team	F	G	Н	I	J	К	A1		
	x rank	6	7	8	$9\frac{1}{2}$	$9\frac{1}{2}$	11			
	y rank	10	3	11	9	1	8			
(b)	$r_s = 0.35$ H_0 Rank player co	c orde	ers of	gate :	receip	ots an	d	B3	6	Or $d = 1, 3, 1, 2, 2, 4, 4, 3, \frac{1}{2}, 8\frac{1}{2}, 3$ $\sum d^2 = 141\frac{1}{2} B1$ $r_s = 1 - \frac{6 \times 141\frac{1}{2}}{11 \times 120} = 0.357 M1A1$ $SC4 0.36$ $SC4 0.318$
	H ₁ Ranl player coa positiv	k orde osts a ve asso	ers of re no	gate t inde	receij					
	cv = 0.	4182						B1		For cv
	$ts r_s = 0$).355						M1		For comparison ts/cv
	$r_{\rm s}$ < 0.4	1182						A1		$r_s = 0.355 \text{ or } 0.357$
	Accept I 10% lev associati receipts	el to s	sugge etwee	est a p n ran	ositiv k orde	/e	gate	E1	5	In context
							Total	. [11	

Q	Solution	Marks	Total	Comments
6	H _o Samples are taken from identical	B1		or H_0 n $-n$ $-n$
	populations			or \mathbf{H}_0 $\boldsymbol{\eta}_{Normal} = \boldsymbol{\eta}_{Depres} = \boldsymbol{\eta}_{MildAlz}$
	H ₁ Samples are not taken from identical	B1		H ₁ at least two of
	populations – population average recall			$oldsymbol{\eta_{\scriptscriptstyle Normal}}$, $oldsymbol{\eta_{\scriptscriptstyle Depres}}$, $oldsymbol{\eta_{\scriptscriptstyle MildAlz}}$ differ
	scores differ			Normal Depres Muaniz
	1% 1 tail			
	170 1 tan			
	Ranks			
	Normal Depression Mild			
	Alzheimer's			
	8 5 1	M1		Ranks
	14 9 2			
	15 10 3	m1		At least 12 correct
	16 11 4			
	17 12 6			
	18 13 7			
	19			
	T 107 T 60 T 22	1		Trace 1
	$T_{Normal} = 107$ $T_{Depres} = 60$ $T_{MildAlz} = 23$	m1		Totals
	$n_{Normal} = 7$ $n_{Depres} = 6$ $n_{MildAlz} = 6$	A1		Any one correct
	Normai , Nepepres O NemuaAiz O			
	$\sum_{i=1}^{m} T_{i}^{2} = 107^{2} = 60^{2} = 23^{2}$			
	$\sum_{i=1}^{m} \frac{T_i^2}{n_i} = \frac{107^2}{7} + \frac{60^2}{6} + \frac{23^2}{6} = 2323.74$	m1		
	1-1 1-1			
	12			
	$H = \frac{12}{19 \times 20} \times 2323.74 - (3 \times 20) = 13.38$	A1		ts correct 13.0 – 13.8
	17/120			
	Critical value from $\chi_2^2 = 9.210$	B1		
	λ ₂ γ.2.10			
	H > 9.210	M1		
	Sig evidence to reject H ₀ and conclude	A1		
	that samples are not from identical			
	populations			
	Significant evidence at the 1% level to	E1	12	Difference in context
	suggest that the population average recall	101	12	Mention of 'at least two' or a significant
	scores differs for the three categories of			difference between scores for Mild
	adults: at least two of the averages differ.			Alzheimer's and those with normal
	It appears that those adults with Mild			memory function
	Alzheimer's disease have a significantly			
	lower average recall score than those who			
	have normal memory function		4-	
	Total		12	
	TOTAL		75	