General Certificate of Education (A-level)
January 2012

Statistics SS03

(Specification 6380)

Statistics 3

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 students' 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 students' 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 students' 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

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.

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.

SS03

Q	Solution	Marks	Total	Comments
1 (a)				
		M1		Attempt at ranks
		M1		14 correct (can be reversed)
		A1		
				Alternative d = 4, 1,1,2,3,3,4,2,2, 4 $\sum d^2$ =80 B1
	r_s = 0.515 or 17/33 (3 sf from calc)	B3	6	$r_s = 1 - \frac{6\times 80}{10\times 99} = 0.515$ M1, A1
	sc2: 0.51/2 sc1: 0.5 awrt			
(b)	H_o Rank orders of gross takings and body counts are independent. H_1 Rank orders of gross takings and body counts are not independent: a positive association exists 1 tail 10%	B1		or equivalent in words/symbols
	cv = 0.4424	B1		
	test stat r_s = 0.515			
	r_s > cv	M1		comparison ts/cv
	comparison ts/cv			
	Reject H_o Significant evidence at 10% level to suggest a positive association between rank orders of gross takings and body counts. For films with a body count greater than 50, those with higher body counts tend to have higher gross takings.	E1	4	Correct conclusion in context
	Total		**10**	

Film	Ti	Ret	2T	Tr	Ry
x rank	1	2	3	4	5
y rank	5	1	4	2	8
Film	Gl	Sam	BB	Ra	Sol
x rank	6	7	8	9	10
y rank	9	3	10	7	6

SS03 (cont)

Q	Solution	Marks	Total	Comments
2 (a)	H_o μ, η = £81,050 H_1 μ, η < £81,050 1 tail 5%	B1		consistent or equivalent in words
	diffs -17530 -16450 -9050 -22600	M1		For differences
	rank 8 7 3 9	m1		Ranks
	diffs 1150 -14550 5550 -12850 rank 1 6 2 5			
	diffs -11950 rank 4			
	T_+ = 1 + 2 = 3 T_- = 8 + 7 + 3 + 9 + 6 + 5 + 4 = 42	m1 A1		Total of ranks One correct
	Test stat T = 3 n = 9 cv = 8 T < 8	B1 m1		For cv Correct ts identified for cv comparison
	Reject H_o			
	There is significant evidence to suggest that average gross annual salary for consultants and medical specialists in the UK was greater than that for those working in France during 2003	E1	8	In context
(b)	Distribution of <u>differences</u> in gross annual salary for consultants and medical specialists is symmetrical	E1	1	
	Total		**9**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
3(a)(i)	Sign test	B1	1	
(ii)	No measured data just a decision reduced/increased or no change	E1	1	
(b)	H_o η = 0			
	H_1 η < 0 1 tail 5%	B1		Allow H_1 η > 0 if signs consistent
	- + - - - . - - + -	M1		For signs
	ts 7-, 2+	M1 A1		Excluding 'no change'
	Binomial model B (9, 0.5) P ($\geq$ 7-) = P($\leq$2+) = 0.0898 > 0.05 for one tail test	M1 m1		Using B (9, 0.5) Comparison correct prob with 0.05
	Accept H_o .			Condone n=10 ts 7-, 3+ M1M1A1 B(10, 0.5) used M1 0.1719> 0.05 m1 E1 ft
	There is insufficient evidence, at the 5% level, to indicate that the doctor's belief is supported.	E1	7	
(c)	Sample clearly not random . There is no control group.	E1	1	Or other relevant reason
	Total		**10**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
4(a)	H_0 Samples from identical populations H_1 Samples not from identical populations 5% sig level	B1		or symbols/words ref to population average
	Ranks	M1		Attempt at ranks
	Diet A: 1 20; 3 18; 4½ 16½; 6 15; 7 14; 10 11; 16 5; 18 3 Diet B: 2 19; 4½ 16½; 8 13; 9 12; 11 10; 13 8; 14 7 Diet C: 12 9; 15 6; 17 4; 19 2; 20 1	m1		14 or more correct (can be reversed)
	T_A=65½(102½) T_B=61½(85½) T_C=83(22) $n_A = 8$ $n_B = 7$ $n_C = 5$	m1		Totals attempted
	$\sum_{i=1}^{m} \frac{T_i^2}{n_i} = \frac{65\frac{1}{2}^2}{8} + \frac{61\frac{1}{2}^2}{7} + \frac{83^2}{5}$ $= 2454.4$	m1		$\sum_{i=1}^{m} \frac{T_i^2}{n_i}$ attempt
	$H = \frac{12}{20 \times 21} \times 2454.4 - (3 \times 21)$ $= 7.12$(or 7.13)	m1 A1		H attempt 7.0 – 7.2
	Critical value from $\chi^2_2 = 5.99$ $H > 5.99$	B1 M1		
	Sig evidence to reject H_0 and conclude that samples are not from identical populations	A1		
	There is significant evidence of a difference between <u>at least two</u> of the diets in terms of average percentage reduction in body weight for overweight men. (Diet C is most effective).	E1	11	
(b)(i)	There may be concern that diet C caused problems that led to the men not completing the diet.	E1	1	
(ii)	She may wish to check that the illness was not caused or made worse by the diet	E1	1	
	Total		**13**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
5(a)(i)				
	Accident in 09 / **No accident in 09** / **Tot** **17-18 years** 26 **174** **200** **19-50 years** **48** **652** **700** **51 years +** **12** **288** 300 **Total** 86 **1114** 1200	M1 A1	2	Some sensible effort 3 cell frequencies correctly placed
(ii)	H_o Involvement in a car accident is independent of age H_1 Involvement in a car accident is not independent of age 1 tail 1%	B1		
	Expected / **Acc** / **No acc** **17-18 years** 14.33 185.67 **19-50 years** 50.17 649.83 **51 years +** 21.5 278.5	M1 A1		Method for expected frequencies All correct
	ts = $\sum \frac{(O-E)^2}{E}$			
	= $\frac{11.67^2}{14.33}+\frac{11.67^2}{185.67}+.....+\frac{9.5^2}{21.5}+\frac{9.5^2}{278.5}$	m1 m1		Numerator correct Denominator correct
	= 14.85	A1		ts correct (13.0 -16.0)
	cv df = 2 1% cv = 9.21	B1		
	ts > 9.21	M1		
	Reject H_o	A1		
	Sig evidence to suggest that involvement in a car accident is not independent of age	E1ft	10	
(iii)	17-18 year olds are far **more** likely to have had an accident.	E1		
	Those aged 51 years and over are far **less** likely to have had an accident	E1	2	

SS03 (cont)

Q	Solution	Marks	Total	Comments
5 cont. (b)(i)	**Expected** / **£0-£2000** / **£2001-£4000** / **Over £4000** **17-30 years** 31.40 15.70 6.90 **31 years +** 18.60 9.30 4.10	M1 A1	2	For 3 correct All correct to 1 dp These marks may be gained in part (ii) Allow already pooled
(ii)	Pooled expected values **Expected** / **£0-2000** / **Over £2000** **17-30 years** 31.40 22.60 **31 years +** 18.60 13.40	M1		Last 2 columns pooled
	H_o Size of claim is independent of age H_1 Size of claim is not independent of age 1 tail 1%	B1		
	ts = $\sum \frac{(\lvert O-E\rvert - 0.5)^2}{E}$	M1		ts
	$= \frac{4.9^2}{31.4} + \frac{4.9^2}{22.6} + \frac{4.9^2}{18.6} + \frac{4.9^2}{13.4}$	m1		Yates used
	= 4.91	A1		4.7 – 5.2 **Alt** for non pooling **sc 5** B1 hypotheses M1 test stat m0 no Yates A1 ft 6.06 (5.9 – 6.2) B1 ft df=2 cv = 9.210 E1 ft
	cv df = 1 1% cv = 6.635 ts < 6.635 Accept H_o	B1		Pooled but no Yates **sc 6** M1B1M1m0A1B1M1
	No sig evidence to suggest that size of claim is associated with age	E1	7	
	Total		**23**	

SS03 (cont)

Q	Solution	Marks	Total	Comments
6	H_0 Samples from identical populations H_1 Samples not from identical populations 5% sig level 2 tail Ranks **Men who have taken caffeine** / **Men who have not taken caffeine** 2 13 — 1 14 6 9 — 3 12 8 7 — 4 11 10½ 4½ — 5 10 12 3 — 7 8 13 2 — 9 6 14 1 — 10½ 4½ T_{Caf} = 65½(39½) T_{nocaf} = 39½(65½) n_{Caf} = 7 n_{nocaf} = 7 U_{Caf} = 65½ - $\frac{7\times8}{2}$ = 37½ U_C = 39½ - $\frac{7\times8}{2}$ = 11½ U = 11½ cv = 9 for n= 7, m = 7 2 tail 5% U > 9 Accept H_o No significant evidence of any difference between average RER for the men who took caffeine and those who did not.	B1 M1 A1 m1 m1 A1 B1 M1 A1 E1ft	10	Or symbols/words ref to population average Attempt at ranks as one group For 12 correct ranks Totals attempted *U* attempt Either *U* correct For cv sc: cv=11 B0M1A0 Correct comparison
	Total		**10**	
	TOTAL		**75**	