Version 1.0: 0110

General Certificate of Education

Statistics 6380

SS03 Statistics 3

Mark Scheme

2010 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 © 2010 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 *Dr Michael Cresswell Director Genera*

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(a)	H_0 pop median/η = 11250 H_1 pop median/η $\neq$ 11250	B1		Or words referring to average price
	2 tail 10%			
	signs $-+$ $+-$ $-+$ $++$ $.+$	M1		signs
	$n = 9$ test stat $= 6^+ / 3^-$ Model B(9 , 0.5)	A1 M1		test stat correct Bin model seen to be used Or cr $\{0,1\}\{8,9\}$ with probs
	$P(\leq 3^-) = P(\geq 6^+) = 0.254 > 0.05$	M1		Comparison of correct B(9, 0.5) prob with 0.05 or use of identified cv with probability (or 0.508/0.10)
	Accept H_0			
	There is no significant evidence to doubt that the median asking price is £11250.	A1	6	
(b)	A Type II error occurs when an incorrect null hypothesis is accepted.	B1		Type II correctly identified.
	In this case, it would mean that we concluded that the population median asking price was £11250 but, in fact, the median asking price was not equal to £11250.	E1	2	Context
	Total		**8**	

SS03(cont)

Q	Solution	Marks	Total	Comments
2(a)(i)	From calculator $r = 0.891$			Alternative $n = 7$
	or $r = \dfrac{29495 - \dfrac{2885 \times 69}{7}}{\sqrt{14242.86} \times \sqrt{98.86}}$			$\sum y = 69$ $\sum x = 2885$ $\sum y^2 = 779$
	$= \dfrac{1057.14}{119.34 \times 9.94}$			$\sum x^2 = 1203275$
	$= 0.891$			$\sum xy = 29495$ M1
		M1		
		m1		m1 formula in (i) or (ii)
		A1		0.885 to 0.905 A1 (3sf)
(ii)	From calculator $r = 0.658$			Alternative $n = 7$
	or $r = \dfrac{34021 - \dfrac{2885 \times 81.8}{7}}{\sqrt{14242.86} \times \sqrt{15.35}}$			$\sum z = 81.8$ $\sum z^2 = 971.24$
	$= \dfrac{307.71}{119.34 \times 3.92}$			$\sum xz = 34021$ M1
	$= 0.658$	M1A1	5	0.650 to 0.665 A1
(b)	$r_{xy} = 0.891$ $r_{xz} = 0.658$			
	H_0 $\rho = 0$			
	H_1 $\rho > 0$ 1 tail 5 % sig level	B1		For hypotheses stated correctly once
	Need only be stated once			
	test stat $r_{xy} = 0.891$			
	cv = 0.6694 $n = 7$			For cv and comparison
	since $t > 0.6694$			
	Reject H_0	M1		For Reject H_0; ft
	test stat $r_{xz} = 0.658$	A1√		
	cv = 0.6694 $n = 7$			
	since $t < 0.6694$			
	Accept H_0	A1	4	For Accept H_0
(c)	There is significant evidence to suggest a positive correlation between the calories and the fat content of milkshakes: the higher the fat content, the higher the calories.	E1		
	There is no significant evidence to suggest a positive correlation between the calories and the volume of the milkshakes.	E1	2	Need to refer to part (b)
	Total		**11**	

SS03(cont)

Q	Solution	Marks	Total	Comments
3(a)	500 – 150 –100 –80 = 170 for West 500 – 105 = 395 rejected	M1 M1		Seen or used
		A1 A1	4	For one unknown 'select' correct All correct
(b)	H_0 Selection independent of home region H_1 Selection not independent of Home region 1 tail 1%	B1		
	Expected frequencies			
		M1		E method for 3 correct; ft
		A1√		For all E correct
	ts = $\sum \frac{(O-E)^2}{E}$	m1		dep sensible effort for E Correct denominator ft
	$= \frac{7.5^2}{31.5} + \frac{7.5^2}{118.5} + \ldots\ldots\ldots\ldots + \frac{21.3^2}{134.3}$	m1		Correct effort at ts ft
	= 24.97	A1		24.0 to 26.0 (or p = 0.0000157)
	df = 3 1% cv = 11.345	B1		3 df
	ts > 11.345	B1		for cv and comparison
	Reject H_0	A1	9	
(c)	There is significant evidence to suggest that selection is not independent of home region.	E1		General conclusion in context (could be in part (b))
	Artists from the south seem **less** likely to be selected (expected higher than observed) and those from the west seem **much more** likely to be selected (expected lower than observed).	E1	2	More detailed identification
	Total		**15**	

	Select	Reject	Total
N	**24**	**126**	**150**
E	**12**	**88**	**100**
S	**12**	**68**	**80**
W	**57**	**113**	**170**
Total	**105**	**395**	**500**

	Select	Reject
N	**31.5**	**118.5**
S	**21**	**79**
E	**16.8**	**63.2**
W	**35.7**	**134.3**

SS03(cont)

Q	Solution	Marks	Total	Comments
4(a)(i)	Ranks			
		M1		attempt at ranks (can be reversed)
		M1		for 12 correct
		A1		all correct
	$r_s =$ 0.766(3 sf from calc)	B3	6	alternative $d = 0,0, 3.5, 2.5, 0, 2, 4, 0,0,0$ $\sum d^2 = 38.5$ B1 $r_s = 1 - \frac{6\times38.5}{10\times99} = 0.767$ M1, A1ft small slip
(ii)	H_0 Rank orders of unleaded petrol excise duty and diesel excise duty are independent. H_1 Rank orders of unleaded petrol excise duty and diesel excise duty are not independent – there is an association 2 tail 5%	B1		or alternatives indicating H_0 No association H_1 Association
	cv = ±0.6485 $n = 10$ 2 tail 5% test stat $r_s =$ 0.766	B1		For cv
	$\lvert r_s \rvert > 0.6485$	M1		For comparison ts/cv; ft
	Reject H_0 Significant evidence at 5% level to suggest an association between unleaded petrol excise duty and diesel excise duty for countries in Europe.	E1	4	For correct conclusion in context [Allow 1 tail H_1 and consistent cv]

	Unleaded	Diesel
Cyprus	**1**	**1**
Romania	**2**	**2**
Sweden	**3**	**6.5**
Slovakia	**4**	**6.5**
Austria	**5**	**5**
Malta	**6**	**4**
Finland	**7**	**3**
France	**8**	**8**
Germany	**9**	**9**
UK	**10**	**10**

SS03(cont)

Q	Solution	Marks	Total	Comments
4(b)	H_0 pop median/mean diff $\eta_d = 0$	B1		
	H_1 pop median/mean diff $\eta_d > 0$			
	1 tail 1% (d is unleaded – diesel)	B1		Consistent with differences
	diff 4 5 –3 1 8 12 22 14 15 0	M1		For differences UL – Diesel or Diesel – UL
	rank 3 4 2 1 5 6 9 7 8 exclude	M1		For ranks
	$T_+ = 3 + \ldots\ldots + 8 = 43$	m1		For total of ranks
	$T_- = 2$	A1		For one correct total or ts = 2 if method seen
	Test stat $T = 2$			
	$n = 9$ cr ≤ 3	B1		For cv
	$T < 3$	M1		Comparison correct cv/ts
	Significant evidence at 1% level to reject H_0 and conclude that average excise duty for diesel is less than that for unleaded petrol in European countries	E1	9	In context
	Total		**19**	

SS03(cont)

Q	Solution	Marks	Total	Comments
5(a)				
		M1		Effort to put into 3 categories
		A1		6 correctly placed (can be implied by totals later)
	Ranks			
		M1		Ranks as one group
		A1		At least 10 correct
	$T_C = 50\frac{1}{2}$ $T_D = 33\frac{1}{2}$ $T_E = 36$	m1		
	$n_C = 5$ $n_D = 5$ $n_E = 5$	B1		Totals of ranks
	H_o Samples are taken from identical populations			or
	H_1 Samples are not taken from identical populations – at least two population average fuel usages differ 10% 1 tail	B1		H_0 $\eta_C = \eta_D = \eta_E$ H_1 at least two of η_C, η_D, η_E do differ
	$\sum_{i=1}^{m} \frac{T_i^2}{n_i} = \frac{50.5^2}{5} + \frac{33.5^2}{5} + \frac{36^2}{5} = 993.7$	m1		for $\sum_{i=1}^{m} \frac{T_i^2}{n_i}$
	$H = \frac{12}{15 \times 16} \times 993.7 - (3 \times 16) = 1.685$	A1		test stat correct 1.6 to 1.8
	Critical value from $\chi_2^2 = 4.605$	B1		
	$H < 4.605$			
	No sig evidence to reject H_0 Conclude that samples are from identical populations. Population average fuel usages between models do not differ	M1 A1	12	

C	D	E
14.4	**14.1**	**13.9**
14.5	**14.3**	**14.2**
14.7	**14.4**	**14.6**
15.2	**14.8**	**14.9**
15.4	**15.0**	**15.1**

C	D	E
5½	**2**	**1**
7	**4**	**3**
9	**5½**	**8**
14	**10**	**11**
15	**12**	**13**

SS03(cont)

Q	Solution	Marks	Total	Comments
5(b)	H_0 Samples are taken from identical populations			
	H_1 Samples are not taken from identical populations – pop average miles per gallon greater for compact cars. 1 tail 5%	B1		Hypotheses referring to population averages also acceptable
	Compact ranks / **Midsize rank**: **6** / **3**; **13** / **4**; **9** / **10**; **12** / **1**; **14** / **2**; **8** / **5**; / **7**; / **11**	M1		Attempt at M–Whitney – ranks as one group
		m1		for 12 correct
		m1		for total attempt (any ranks)
	$T_C = 6 + ... + 8 = 62$ $T_M = 3 + ... + 11 = 43$	m1		for U
	$U_C = 62 - \frac{6\times7}{2} = 41$ $U_M = 43 - \frac{8\times9}{2} = 7$	A1		one U correct
	Test stat $U = 7$ $n = 6$, $m = 8$ cr ≤ 11	B1		for cv
	$U = 7 < 11$	M1		correct comparison cv/U
	Reject H_0 Significant evidence at the 5% level to suggest that the average city miles per gallon is greater for compact cars.	A1 E1√	10	reject H_0 Conclusion in context
	Total		**22**	
	TOTAL		**75**	