Version 1.0

General Certificate of Education (A-level)
June 2011

Statistics SS02

(Specification 6380)

Statistics 2

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 candidates' 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 candidates' 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 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 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.

Q	Solution	Mark	Total	Comments
1(a)(i)	On graph	B1		B1 accurate plot – by eye
		B1	2	B1 trend line – must be a line
(ii)	Riz effect (49.3 + 66.3 +70.7)/3 = 62.1	M1		M1 attempt to find Riz deviations from trend line or moving average
		m1		m1 method, ignore sign
		A1	3	A1 62.1 (60 ~ 64)
(iii)	555 + 62.1 = 620	B1		B1 trend 555 (545 ~ 565)
		M1		M1 'their' trend + 'their' seasonal effect
		A1	3	A1 620 (605 ~ 630)
(b)	Week 18 – trend line below 640 for weeks 16 and 17. Attendances for Ed and Ja not likely to be above trend line.	B1		B1 week 18
		E1		E1 justification for any of weeks 16, 17, 18
	Week 18 trend line about 600 and Riz likely to be more than 40 above trend line.	E1	3	E1 full explanation
	Total		**11**	
2(a)(i)	$E(X) = 100 \times 0.22 + 200 \times 0.31 + 300 \times 0.21 + 400 \times 0.12 + 600 \times 0.14 = 279$	M1		M1 method
		A1		A1 279 CAO AG
(ii)	$E(X^2) = 100^2 \times 0.22 + 200^2 \times 0.31 + 300^2 \times 0.21 + 400^2 \times 0.12 + 600^2 \times 0.14 = 103100$			
	$V(X) = 103100 - 279^2 = 25259$	M1		
	s.d. $= \sqrt{25259} = 158.9$	A1	4	B2 159 (158.5 ~ 159.5) or M1A1
				SC: allow B1 for variance = 25259
(b)	Standard deviation would increase as distribution would be more spread out	B1		B1 increase
		E1	2	E1 reason
(c)	Standard deviation would be less than for X. Nearly all cars have parked for free so there is little variability in the distribution.	B1		B1 less than X
		E1	2	E1 reason
	Total		**8**	

Q	Solution	Mark	Total	Comments
3(a)	3 509 000	B2,1	2	B2 3 509 000 allow B1 for 3509
(b)	males < 1 + females < 1 = 471 + 466 = 937 thousand total < 1 = 938 thousand			
	This is consistent with rounding error eg if 471 400 males and 466 400 females, total = 937 800 which rounds to 938 thousand	E1 E1	2	E1 no, E1 could be due to rounding error
(c)(i)	In each census there are more males than females under 1 enumerated. This suggests that the probability of a baby being female is less than 0.5.	B1		B1 not supported
	The 'fact' is not supported.	E1	2	E1 explanation SC: allow B1 supported because proportion close to 0.5
(ii)	In each census there are many more females than males aged 75 and over. This supports the 'fact' that on average females live longer than males.	B1 E1	2	B1 supported E1 explanation
(d)(i)	<table><tr><th>Males aged</th><th>Thousands</th></tr><tr><td>under 15</td><td>5689</td></tr><tr><td>15–29</td><td>5623</td></tr><tr><td>30–44</td><td>6645</td></tr><tr><td>45–59</td><td>5534</td></tr><tr><td>60–74</td><td>3720</td></tr><tr><td>over 74</td><td>1620</td></tr></table>	M1 A1 A1	3	M1 reasonable attempt – allow wrong year, wrong units, wrong section A1 three correct 3sf A1 all correct 3sf
(ii)	On graph	B1 B1 B1	3	B1 vertical axis correctly labelled starting at 0 B1 accurately plotted points B1 vertical line starting at 0
(iii)	Up to age 60 similar numbers in each age group except for 30–44 which has more than the others. After age 60 there are less in the groups as there will be more deaths in these groups.	E1 E1	2	E1 age groups similar E1 more in 30–44 or any sensible comment
(iv)	Purpose of diagram is to assist interpretation of data. Drawing a line diagram with unequal age groupings would make interpretation extremely difficult.	E1	1	E1 unequal age groups
	Total		**17**	

Q	Solution	Mark	Total	Comments
4(a)(i)	P(0) = 0.301	B1		B1 0.301 (0.3 ~ 0.3015)
(ii)	P(1) = 0.6626 – 0.3012 = 0.361	M1 A1		M1 method A1 0.361 (0.36 ~ 0.362)
(iii)	P(5 or more) = 1 – 0.9923 = 0.0077	M1 A1	 5	M1 method A1 0.0077 (0.0077 ~ 0.0078)
(b)	$\bar{x} = 1.2 \quad s^2 = 2.436 \ (\sigma^2 = 2.382)$	B1 B1	 2	B1 1.2 CAO B1 2.44 (2.43 ~ 2.44) or 2.38 (2.37 ~ 2.39)
(c)(i)	proportion 0s observed = 0.489 not similar to 0.301	M1 A1	 2	M1 attempt to compare observed proportion of 0s with probability calculated in (a)(i) A1 correct comparison
(ii)	Mean and variance of observed values not similar.	E1	1	E1 reason – even if based on incorrect answers to (b)
(iii)	Most credit cards lost in the evening may be reported in first hour / cards lost over the weekend likely to be reported on Monday etc.	E1 E1	 2	E1 reference to valid property of Poisson distribution E1 plausible context
(d)	No. Since distribution appears not to be Poisson no evidence to support view that reports arrive at random times.	E1 E1	 2	E1 No E1 not Poisson → not random or explanation of why not random
	Total		**14**	

Q	Solution	Mark	Total	Comments
5(a)	H_0: $\mu = 700$ H_1: $\mu \neq 700$	B1		B1 one correct hypothesis
	$\bar{x} = 699.6$	B1		B1 699.6 CAO
	$z = (699.6 - 700)/(2.1/\sqrt{6}) = -0.467$	M1 m1 A1		M1 use of $2.1/\sqrt{6}$ m1 method for z – ignore sign A1 –0.467 (–0.46 ~ –0.47)
	c.v. ± 1.96	B1		B1 ± 1.96 – ignore sign
	Accept H_0	A1√		A1√ conclusion – allow even if contradicted later; also allow for valid comparison with +ve test statistic
	Conclude that there is no significant evidence to doubt that the mean is 700mm	A1	8	A1 in context – needs previous A1√ plus something additional to mean = 700; disallow for +ve test statistic
(b)	H_0: $\mu = 700$ H_1: $\mu \neq 700$	B1		B1 both hypotheses correct – needs μ or 'population'
	$z = (701.34 - 700)/(2.1/\sqrt{40}) = 4.04$	M1 A1		M1 method for z – ignore sign A1 4.04 (4 ~ 4.04)
	c.v. ± 1.96			
	Reject H_0	A1√		A1√ conclusion – allow even if contradicted later; also allow for valid comparison with + ve test statistic
	Conclude that there is significant evidence to conclude that the mean is not equal to (greater than) 700mm	A1	5	A1 in context – needs previous A1√ plus something additional to mean ≠ 700; disallow for +ve test statistic
(c)(i)	Neither. Risk of Type I error is 5% regardless of sample size (or zero if H_0 untrue).	B1 E1		B1 neither E1 explanation
(ii)	Larger sample would lead to a smaller risk of Type II error because s.d./$\sqrt{n}$ is smaller and so more likely to detect a deviation from 700 (or the same if H_0 true).	B1 E1	4	B1 sample of 40 E1 explanation – allow for anything which implies a correct definition of Type II error
	Total		**17**	

Q	Solution	Mark	Total	Comments
6(a)	Number pupils 000 to 307. Choose 3-digit random numbers from Table 13. Ignore > 307 and repeats. Continue until 12 numbers obtained and choose corresponding pupils.	E1 E1 E1 E1	4	E1 number pupils 000–307 E1 choose 3-digit random numbers E1 ignore > 307 – consistent with their numbering E1 ignore repeats
(b)	Pupil numbered 000 will be selected if 000, 308, 616 or 924 are selected – probability 0.004. Pupil numbered 307 will be chosen if numbers 307, 615 or 923 are selected – probability 0.003. Since probabilities not equal sample cannot be random.	E1 E1	2	E1 different probabilities of being selected E1 explanation
(c)	Pupil 000 would only be selected if, after multiplying by 308, the answer was between 0.0 and 0.5. All other pupils would have a range with double the width, eg pupil 001 would be selected if the answer was between 0.5 and 1.5 **or** Pupil 000 has 2 chances 0.000 and 0.001 all others have 3 or 4 chances eg Pupil 001 0.002/3/4 Pupil 002 0.005/6/7/8	E2,1	2	E2,1 explanation
	Total		**8**	
	TOTAL		**75**	