A-LEVEL
Statistics
SS05
Mark scheme

6380
June 2018

Version/Stage: 1.0 Final

Mark schemes are prepared by the Lead Assessment Writer 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 associates 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 associate understands and applies it in the same correct way. As preparation for standardisation each associate 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, associates encounter unusual answers which have not been raised they are required to refer these to the Lead Assessment Writer.

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

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.

Q1	Solution	Mark	Total	Comment
(a)	$\mu = 0$,	B1		
	$\sigma^2 = \frac{1}{12}(0.05 - -0.05)^2$	M1		OE PI var 0.0008333
	$\sigma = 0.0289$	A1	3	AWRT 0.0288 or 0.0289 (0.02886751)
(b) (i)	0	B1	1	
(ii)	p(underestimated by at least $10g$) = 0.4	B1		PI
	$(0.05 - 0.035 =)\ 0.015$	M1		PI Correct subtraction to consistent scale e.g. $(50 - 35 =)\ 15$ or $(0.5 - 0.35 =)\ 0.15$ (B1M1 may be seen on a diagram.)
	$0.015/0.04$ (or $0.15/0.4$)	m1		OE PI Correct ratio
	$= 0.375$	A1	4	OE eg 15/40 3/8 $\left(\text{SC2}\frac{0.37}{\frac{0.38}{\text{NMS}}}\right)$
(c)	Use Normal distribution with $\mu = 0$ and			
	$\sigma = \frac{0.0289}{\sqrt{46}} (= 0.00426)$	M1		PI Use of "their" $\sigma/\sqrt{46}$ or $\sigma^2/46$
	$z = \frac{0.01 - 0}{\left(0.0289 \middle/ \sqrt{46}\right)} = 2.35$	m1 A1		PI standardisation PI correct z-value (2.34946803)
	$P(\lvert\bar{X}\rvert < 0.01) = P(\lvert Z\rvert < 2.35) = 0.981$	A1	4	AWRT (0.98122658) SC3 NMS p = 0.991 (0.99059987) correct to at least 3 s.f.
	Total		**12**	

Q2	Solution	Mark	Total	Comment
(a) (i)	$\mu = \frac{1}{\lambda} = \frac{1}{0.36} = 2.78$	B1	1	Or 2.77
(ii)	$1 - e^{-0.36\times3}$ $= 0.660$ or 0.66	M1 A1	 2	PI Correct use of formula AWRT (0.6604044744)
(b)	(*Use* of memoryless property) x = 2 $1-(1-e^{-0.36\times2})$ or $e^{-0.36\times2}$ $= 0.487$ or 0.486	M1 A1	 2	*Not* using both $1-e^{-0.36\times2} = 0.513$ x = 2 and x = 5 Allow correct use of conditional probability with x = 5 and x = 7. AWRT (0.486752256)
(c)	***Either*** The locations of platypus <u>burrows</u> might not be independent (of each other). ***or*** Platypus burrows might not occur at a constant average rate/interval along the river	B1 E1 (B1) (E1)	 2	PI not independent or equivalent random clear and correct context (Might be stated as "platypus burrows may be grouped together" or "platypus burrows might be at a fixed distance apart".) "platypuses live in colonies" not constant average rate/interval clear and correct context For E1 must have mentioned burrows/homes not just platypuses
	Total		**7**	

Q3	Solution	Mark	Total	Comment
(a) (i)	$s_A = 0.9615$ or $s_A^2 = 0.9244$ or $S_{AA} = 8.32$	B1		Any correct, awrt $s_A = 0.961$~0.962 awrt $s_A = 0.924$, awrt $S_{AA} = 8.32$
	df = 9	B1		PI df
	$\chi^2 = 2.700$ & 19.023	B1		Both, do not allow ±
	$\frac{9\times0.9615^2}{19.023} < \sigma^2 < \frac{9\times0.9615^2}{2.7}$	M1ft		Formula correct for *their* s_A Allow small slip on χ^2 values.
	$0.4374 < \sigma^2 < 3.0815$	A1		Formula all correct
	95% CI for σ_A is (0.66, 1.76)	A1	6	Lower limit CSO (correct to 2 d.p.) Upper limit AWRT 1.75~1.76
(ii)	(df = 11) $\chi^2 = 3.816$ & 21.92	B1		Both, do not allow ±
	$\frac{11\times1.21^2}{21.92} < \sigma^2 < \frac{11\times1.21^2}{3.816}$ $0.7347 < \sigma^2 < 4.2204$	M1		Formula correct for *their* s_A with *correct* χ^2 values.
	95% CI for σ_B is (0.86, 2.05)	A1	3	Lower limit AWRT 0.85 or 0.86 Upper limit CSO (correct to 2 d.p.)
(iii)	The two confidence intervals overlap considerably. Or comparing <u>both</u> limits of the CIs	B1dep		OE
	Therefore there is no significant evidence of a difference in the variability of the <u>lengths</u> of rolls of Brand A and Brand B.	E1dep		Both marks dependent on correct CI in (a)(i) *and* (a)(ii). Allow small arithmetic slip.
	some evidence that lengths of Brand B are more variable than those of Brand A.		2	Must express element of doubt.

Q3	Solution	Mark	Total	Comment
(b) (i)	$S_p^2 = \dfrac{9\times 0.9615^2 + 11\times 1.21^2}{20} = 1.22$	M1		Numerator of formula correct using "*their*" s_A or s_A^2 (=0.9244)
	AG	A1	2	CSO with full correct working
(ii)	H_0: $\mu_B = \mu_A + 0.2$ H_1: $\mu_B > \mu_A + 0.2$	B1		OE Other suffices must be clearly defined.
	$\bar{x}_A = 30$	B1		Or 3000
	t.s. $= \dfrac{30.9 - 30 - 0.2}{\sqrt{1.22\left(\frac{1}{10}+\frac{1}{12}\right)}} = 1.48$	M1 M1 A1		Numerator correct using *their* $\bar{x}_A$ Denominator correct AWRT 1.48 (allow rescaling to cm for M1M1A1)
	c.v. $t_{20} = 1.325$ or $p = 0.0773 < 0.1$	B1		correct c.v. or p-value correct, compared to 0.1
	$(1.48 > 1.325)$ Reject H_0.	A1dep		Correct conclusion dependent on all values correct.
	There is significant evidence (at the 10% level) that rolls of baking parchment of Brand B are <u>more than</u> 20 cm <u>longer</u> on <u>average</u> than those of Brand A.	E1dep	8	In context
(c)	Brand <u>B</u> rolls are significantly <u>longer</u> (on average) than brand <u>A</u> rolls (by more than 20 cm).	E1dep		Dependent on second A mark in part (b)(ii).
	Brand <u>B</u> rolls are not significantly more <u>varied</u> in length than brand <u>A</u> rolls.	E1dep		No difference in variability of lengths. Dependent on correct CI in (a)(i) *and* (a)(ii). Allow small arithmetic slip. Disregard answer to (a)(iii).
	Brand B should be bought…	E1		Recommendation of brand B.
	…(unless Brand A is of) better quality or …much cheaper.	E1	4	One other practical factor such as quality, price, availability etc.
	Total		**25**	

<table>
<tr><th>Q4</th><th>Solution</th><th>Mark</th><th>Total</th><th>Comment</th></tr>
<tr><td>(a)</td><td>H_0: The rectangular distribution is an adequate model
H_1: The rectangular distribution is not an adequate model

<table><tr><th>Interval</th><th>O</th><th>p</th><th>E</th><th>$(O-E)^2/E$</th></tr><tr><td>6 to 7</td><td>71</td><td>0.128</td><td>76.8</td><td>0.438</td></tr><tr><td>7 to 8</td><td>84</td><td>0.152</td><td>91.2</td><td>0.568</td></tr><tr><td>8 to 9</td><td>110</td><td>0.204</td><td>122.4</td><td>1.256</td></tr><tr><td>9 to 10</td><td>142</td><td>0.188</td><td>112.8</td><td>7.559</td></tr><tr><td>10 to 11</td><td>102</td><td>0.192</td><td>115.2</td><td>1.513</td></tr><tr><td>11 to 12</td><td>91</td><td>0.136</td><td>81.6</td><td>1.083</td></tr><tr><td>Σ</td><td>600</td><td></td><td></td><td>12.417</td></tr></table>
$\chi^2 = \frac{5.8^2}{76.8} + \frac{7.2^2}{91.2} + \frac{12.4^2}{122.4} + \frac{29.2^2}{112.8} + \frac{13.2^2}{115.2} + \frac{9.4^2}{81.6}$

t.s. = 12.4

c.v. χ^2_5 is 9.236 or p = 0.0295< 0.1

(12.4 > 9.236) Reject H_0

There is significant evidence that the rectangular distribution does <u>not</u> adequately <u>model</u> the distribution of items of <u>debris</u> on the motorway verges.</td><td>B1

M1

m1

A1

M1ft

A1

B1

A1dep

E1dep</td><td>9</td><td>OE both

p = interval length/25
OR mean = 600/25 = 24

E-values = 600×p OR 24×distance at least four correct.

All E-values correct to 3 s.f. at least

Attempt at $\sum \frac{(O-E)^2}{E}$
ft incorrect Es

AWRT 12 to 12.8 (12.416859)
(SC1 χ^2 = 30.4 to 30.5 NMS for all E=100)

correct c.v.
or p-value correct, compared to 0.1

Correct conclusion dependent on all values correct.

In context</td></tr>
<tr><td>(b)</td><td>There was much more debris observed (142) between junctions 9 and 10 than expected (113).

Prudence should concentrate her investigation on the motorway between junctions 9 and 10.</td><td>M1

E1dep</td><td>2</td><td>Correctly comparing observed and expected or reference to the contribution to χ^2 (of 7.559) between junctions 9 and 10 only.

Advising Prudence in context. Dependent on M1.</td></tr>
<tr><td></td><td>Total</td><td></td><td>11</td><td></td></tr>
</table>

Q5	Solution	Mark	Total	Comment
(a)	H_0: $\sigma_P = \sigma_G$ H_1: $\sigma_P > \sigma_G$	B1		OE Both Other suffices must be clearly defined.
	$s_P = 160.6058$ or $s_P^2 = 25794.2$ $s_G = 86.8525$ or $s_G^2 = 7543.4$ ($n_P = 7$ $n_G = 8$)	B1		PI Any one sd or var correct to 3 s.f.
	t.s. $F = \frac{160.61^2}{86.85^2} = 3.42$ (or 0.292)	M1ft A1		Method for calculating F ft "*their*" sd or var AWRT 3.4 or 2.9
	c.v. $F_{6,7} = 3.866$ (or 0.2377) or p = 0.0663	B1 B1		PI d.f. both correct matching c.v. or p-value correct & compared to 0.05
	(3.42<3.866 or 0.292>0.2377) Accept H_0. There is not sufficient evidence at the 5% level to support Syed's suspicion.	A1dep E1dep	8	Correct conclusion dependent on all values correct. In context
(b)	H_0: $\sigma^2 = 7100$ (or $\sigma^2 \leq 7100$) H_1: $\sigma^2 > 7100$	B1		OE Both, disregard suffices. ($\sqrt{7100} = 84.2615$)
	t.s. $\chi^2 = \frac{7 \times 86.85^2}{7100} = 7.44$	M1ft A1		or $(7 \times 7543.4)/7100$ ft "their" σ AWFW 7.42 ~ 7.45
	c.v $\chi^2_7 = 14.067$ or p = 0.385 > 0.05	B1 B1		PI d.f. (implied by $\chi^2 = 2.167$) correct cv or p-value correct, compared to 0.05
	(7.44 < 14.067 or 0.385 > 0.05) Accept H_0. There is no significant evidence to doubt the manufacturer's claim.	A1dep E1dep	7	Correct conclusion dependent on all values correct. In context

Q5	Solution	Mark	Total	Comment
(c)	Random samples of (patients)	B1		Patients were independent of each other Patients selected at random
	The (patch and gum) samples were independent	B1		Patients were randomly allocated to the treatment groups Context not required for first two B marks
	The populations of <u>blood nicotine levels</u> both for patients using nicotine gum and for those using nicotine patches should be <u>normal</u>ly distributed.	B1		Both populations are normally distributed, in context. "Populations" can be implied as long as "samples" not stated. B marks may be gained by clear statements in parts (a) and (b).
	E.g. Patients were all from the same clinic/doctors' surgery so not a random sample.	E1		Any **two** sensible comments in context regarding random or independent samples or normally distributed populations.
	Patients were all from the same clinic/doctors' surgery so might be related so the samples might not be independent.	E1		
	Patients chose their own therapy so the samples are not random.	(E1)		
	The samples contain only patients who found the therapy successful so they are not random.	(E1)		
	We have no evidence that the populations of blood nicotine levels are normally distributed.	(E1)	5	
	Total		**20**	