A-LEVEL
Statistics

SS05
Mark scheme

6360
June 2016

Version 1.0: Final Mark Scheme

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.

Q	Solution	Mark	Total	Comment
1(a)(i)	$s = 3.90$, $s^2 = 15.2$ or $\sum(x-\bar{x})^2 = 197.9$	B1		Stated or used awfw 3.89 ~3.91 (3.90179…) , awfw 15.1~15.3 9 (15.224…) or awfw 197 ~ 199 (197.912…)
	98% limits for χ^2_{13} = 4.107 , 27.688	B1, B1		accept 3sf or better. If B0, B0 then s.c. B1 for sight of 13 degrees of freedom.
	CI limits for variance: $\frac{13 \times 3.90^2}{27.688}$, $\frac{13 \times 3.90^2}{4.107}$	M1, m1		M1: correct form for at least one limit, condone 14 s^2, ft on χ^2 values. m1: both expressions completely correct.
	Limits 7.15 , 48.19	A1	**6**	Both : awfw 7.1 ~ 7.2 , 48 ~ 48.3
(ii)	5 ml^2 is less than the lower limit of the CI . Jamal's suspicion is verified	B1ft E1dep	**2**	B1 ft comment on their CI E1 dep on A1 in (a)
(b)	Lower limit of 95% CI is greater than that of 98% interval , Kajika will come to the same conclusion as Jamal.	E1 E1dep	**2**	Or 95% CI is narrower than a 98% CI oe E1 dep on first E1 <u>and</u> A1 in (a)
	Total		**10**	

a(ii) B1 ft accept "below CI" but NOT " not within CI" and cand. must be using 5ml^2 for the comparison.

(b) E1 accept" lower limit is increased" or" lower limit is higher .." but NOT "CI is higher"
E1 dep – cand. must be using 5ml^2 in a(ii).

Q	Solution	Mark	Total	Comment
2(a)(i)	Area = 1 $\Rightarrow$ 50k = 1 so $k = \frac{1}{50}$	M1 A1		M1 for sight of 130 – 80 = 50 ; may be by using integration or may be from a graph. note: must have sight of "k = " to award A1 ALT: $k = \frac{1}{b-a} = \frac{1}{130-80}$ M1 $= \frac{1}{50}$ A1
(ii)	P(X = 100) = 0 Impossible to estimate a 1m length exactly	B1 E1	**4**	 o.e.
(b)(i)	Mean = 105	B1		cao
(ii)	S.D. = $\sqrt{\frac{2500}{12}} = \sqrt{208.33..} = 14.4$	B1	**2**	awrt (14.4337…)
(c)(i)	P(score < 5) = P (95 < x < 105) $= \frac{10}{50}$ = 0.2	M1 A1		sc B1 for $\frac{5}{50} = 0.1$
(ii)	$X \sim B(25, 0.2)$ P($X \geq 6$) = 1 – P($X \leq 5$) = 1 – 0.6167 = 0.3833	M1 A1	**4**	Using B(25, 0.2 or 0.1) PI awfw 0.38~ 0.39 note:P(X≥6) when $X \sim B(25, 0.1)$ =1 - 0.9666 = 0.0334
	Total		**10**	

Q	Solution	Mark	Total	Comment
3(a)	σ_w:s.d. for distances driven with *Whizzer* balls σ_s: s.d. for distances driven with *Screamer* balls $H_0: \sigma_w^2 = \sigma_s^2$ or $H_0: \sigma_w = \sigma_s$ $H_1: \sigma_w^2 > \sigma_s^2$ or $H_1: \sigma_w = \sigma_s$	B1		Both hypotheses; other suffices must be clearly identified. Must use σ or population s.d./variance
	$s_w = 23.261$ or $s_w^2 = 541.06$ $s_s = 13.03$ or $s_s^2 = 169.8$ $(n_w = 16$ $n_s = 12)$	B1		Either ; awfw 23.2 ~ 23.3 or 540~ 542 ; 13.0 ~ 13.1or 169 ~ 170
	t.s. $F = \frac{23.26^2}{13.03^2} = 3.19$	M1 A1		awfw 3.135 ~ 3.215 (3.18672…)
	c.v. $F_{15,11} = 2.719$	B1,B1		B1 : df – in order 15,11 PI by correct c.v. B1 c.v. accept 3 sf or better Do not condone ± Ignore extra value; $\frac{1}{2.507} =$ 0.399
	3.19 > 2.719 or p = 0.029 < 0.05; reject H_0	A1 dep		Dep A1 for ts and B1 for cv p = 0.031 ~ 0.028 (0.02921…)
	Evidence, at the 5% level, that Lucy's belief is supported.	E1 dep	**8**	Correct conclusion in context dep previous A1dep. Must indicate some level of uncertainty.
(b)	Samples random and/or independent	E1		Either Do not allow eg "random" or "normal"
	Distances driven with each type of ball must be normally distributed	E1	**2**	
	Total		**10**	

(a) Alt 1**:** t.s. $F = \frac{13.03^2}{23.26^2} = 0.314$ M1 A1 (awfw 0.311~ 0.319) ; c.v. $\frac{1}{F_{11,15}} = \frac{1}{2.507} = 0.399$ B1df B1 cv (awfw 0.398~ 0.40);
0.314 < 0.399 ; reject H0 A1dep; E1 dep as on MS

Alt 2: Use of p value and if no intermediate evidence seen:
B1 (hypotheses); p = 0.031 ~ 0.028 (0.02921…..) implies B1 (for variances) M1 A1 for ts (outside this range and they lose all 3 marks) ;comparing 0.029..< 0.05 and rejecting H_0 implies B1B1 for cv and A1 dep; E1 conclusion as on MS

(a) alternative conclusion: There is sufficient evidence to suggest that Lucy's driving distances are more variable when she uses a Whizzer ball than when she uses a Screamer ball.

Q	Solution	Mark	Total	Comment
4a(i)	P(X ≤ 1.5) = P(Z < $\frac{1.5-12.31}{7.40}$) $= P(Z < -1.46)$ = 1 – 0.928 = 0.072 P(X > 29.5) = P(Z > 2.32) = 1- 0.99 = 0.01	M1		M1 :attempt at any appropriate normal probability eg P(*X* < 1 or 1.5), P(*X* > 29 or 29.5) or P(X < 6.5 or 7) or P(X < 9 or 9.5) [P(X < 1) = 0.0632; P(X > 29) = 0.01205; P(X <7) = 0.2365; P (X < 9) = 0.32733]
	w = 0.072 x 200 = 14.4	M1A1		M1: multiplying an appropriate normal probability by 200 A1: *w* 14.4 ~14.6 (14.4067……)
	P(6.5 < X < 9.5) = P(-0.785 < Z < -0.380) = 0.785 - 0.648 = 0.137	M1		M1 subtracting : Allow for: P(X < 9 or < 9.5) – P(X < 7 or 6.5) [0.3273 – 0.2365 = 0.0908] or 200 – 156.39 – *w* – *z* oe or 1 – sum of probabilities
	y = 0.137 x 200 = 27.4	A1		A1: *y* 26.6 ~ 27.8 (27.177…)
	z = 0.01 × 200 = 2.0	A1	**6**	A1: *z* 1.51 ~ 2.41 (2.018…)
4(a)(ii)	H_0: Data follows a normal distribution H_1: Data does not follow a Normal distribution	B1		B1: at least H_0 o.e. eg The normal distribution is a suitable model No parameters included
	(table below)	M1		M1 : attempt at $(O - E)^2/E$ allow for at least **4** values correct to 1dp
		M1		M1: combining last 2 classes If more than the final 2 classes are combined – M0
		m1		m1: adding , dep on both previous M's
		A1		awrt 14.2 ~ 15.6 (14.8361…)

classes	O	E	(O - E)	(O-E)^2	(O - E)^2 /E
≤1.5	8	14.410	-6.410	41.088	2.851
2 - 3	13	8.980	4.020	16.160	1.800
4 - 6	27	19.850	7.150	51.123	2.575
7 - 9	37	27.170	9.830	96.629	3.556
10 - 14	39	52.860	-13.860	192.10	3.634
15 - 19	45	43.600	1.400	1.960	0.045
20 - 24	23	23.170	-0.170	0.029	0.001
>24.5	8	9.960	-1.960	3.842	0.386
	200				14.849

Q	Solution	Mark	Total	Comment
	df = 8 – 1 – 2 = 5 cv = 11.07	B1ft B1		Their number of classes - 3 Do not condone ±
	14.84 > 11.07 or *p* = 0.011 < 0.05 ; reject H_0			*p* = 0.009 ~ 0.013 (0.011027..)
	Sufficient evidence to say that the data does not follow a Normal distribution.	A1dep	**8**	Correct conclusion in context dep A1 for ts and B1 for cv.

4(b)	Large sample ; CLT applies Mike's claim correct	E1 E1dep	**2**	E1 : mention of CLT o.e. E1 : dep on previous E1
	Total		**16**	

(a) (i) NMS each value in range B2
(a) (ii) NMS B1 hypotheses; ts in range M1M1m1A1 ; $p = 0.011(0.009 \sim 0.012) < 0.05$; reject H_0 B1ft B1 ;A1 dep as on MS

Alt: using $\sum \frac{O^2}{E} - N$: M1 attempt at O^2/E (at least 4 values correct to 1sf) ; M1 combining last 2 classes ; m1 summing and subtracting 200– must be positive answer , A1 answers 14.5 ~ 15.5 (14.849).

Use of p value and if no intermediate evidence seen:
B1 hypotheses, $p = 0.011$ (0.009~0.013) implies M1 M1m1A1B1ft ; comparing $0.011 < 0.05$ B1 and correct conclusion in context E1dep (dependent on previous A1 and B1) ;

Q	Solution	Mark	Total	Comment
5a	Mean = $\frac{1}{\lambda} = 40$; variance = $\left(\frac{1}{\lambda}\right)^2 = 1600$	B1,B1	2	Cao both
5b(i)	P(T > 30) = $e^{-0.025\times30}$ = 0.4724	M1 A1		or 1 - (1- $e^{-0.025\times30}$) = 1 – 0.528 awfw 0.472 ~ 0.473 (0.472366…)
b(ii)	On 2 occasions : prob = 0.4724^2 = 0.2231	B1ft	3	awrt 0.223 ~ 0.224: f.t. on their b (i)
(c)	$P(\overline{T} > 35) = P(Z > \frac{35-40}{\sqrt{\frac{1600}{75}}})$	M1		Standardising with 35 and 40 ; condone $\sqrt{40}$ or $\frac{1600}{75}$ as denominator.
		B1		$\sigma = \sqrt{\frac{1600}{75}}$ or $\sigma^2 = \frac{1600}{75}$ seen or implied by correct probability. [awrt 4.62 (4.6188..)]
	= P(Z > –1.08..)	A1		awfw –1.08 ~ – 1.09
	= 0.860	A1	4	0.859 ~ 0.863 (0.86049…) NMS 4/4 for a probability in correct range.
	Total		**9**	

Q	Solution	Mark	Total	Comment
6(a)(i)	$H_0: \sigma^2 = 16$ $H_1: \sigma^2 \neq 16$	B1		Both hypotheses o.e.
	$s = 1.28$ or $s^2 = 1.63$	B1		awrt 1.27 ~ 1.29 (1.277…) or 1.61 ~ 1.65 (1.631…).
	t.s. $X^2 = \frac{14\times1.28^2}{16} = 1.43$	M1 A1		awrt 1.42 ~ 1.43 (1.427…)
	cv χ^2_{14}= 4.075 and 31.319	B1, B1		B1: df clearly identified, PI by correct cv. B1: at least the lower value – do not condone ±
	1.43 < 4.075 reject H_0 Evidence at the 1% level that the sample does not come from a population with a standard deviation of 4 kg/m^2	A1 dep E1 dep	**8**	A1: dep on B1 for CV and A1 for t.s. Conclusion in context dep previous A1 dep.
6(a)(ii)	The sample is not random because it is taken from Lian's class.	E1	**1**	o.e. eg data is taken from Lian's class so may be biased or unrepresentative.
6(b)	$H_0: \mu_{car} = \mu_{alt\ means} + 1$ $H_1: \mu_{car} > \mu_{alt\ means} + 1$	B1 B1		B1: an inequality in μ and "1" B1: all correct; other suffices must be clearly identified – allow "c" and "a".
	$\bar{x}_{car} = 26.1$ $\bar{x}_{alt.\ means} = 24.225$ $S_{car} = 2.09$ $s_{alt\ means} = 1.74$	B1 B1		B1; either mean (cao, 24.2 ~24.3) B1: either s (2.09 ~2.10, 1.73~ 1.74)
	$s_p^2 = \frac{10\times2.09^2 + 7\times1.74^2}{17} = 3.82$	M1		s_p^2 = 3.81617… s_p = 1.9535… NMS award M1 for value used in range 3.80 ~ 3.84
	$t = \frac{26.1 - 24.225\ (-1)}{\sqrt{3.816\times(\frac{1}{11}+\frac{1}{8})}} = 0.964$	M1 M1 A1		M1: numerator M1: denominatorA1: awfw 0.9 ~ 1.0 ; must have gained all M's.
	cv t_{17} = 1.740	B1		cao, accept 1.74 , condone ±
	0.964 < 1.740 or p = 0.174 > 0.05; accept H_0	A1dep		Dep A1 for t.s. and B1 for positive c.v. p: awfw 1.66 ~ 1.90 (0.1743..)
	No significant evidence at the 5% level that male students who travel to college by car have a BMI which is, on average, more than 1 kg/m^2 greater than those who travel to college by alternative means of transport.	E1 dep	**11**	o.e. conclusion in context dep previous A1
	Total		**20**	

6(b) Alt 1 for lower tail

ts $= \frac{24.225 - 26.1\ (+1)}{\sqrt{3.816\times(\frac{1}{11}+\frac{1}{8})}} = -0.964$ M1 M1 M1A1 ; $t_{17} = \pm1.740$ B1

- 0.964 > -1.74 ; accept H_0 A1 dep (nb signs resolved) ; E1 dep as on MS

Alt 2

Use of p value and if no intermediate evidence seen: B1 B1 hypotheses as on MS ; p = 0.166 ~ 0.190 (0. 1743…) implies B1 B1 M1 M1 M1 A1 (outside this range and they lose all 6 marks) ;comparing 0.174… > 0.05 B1 and accept H_0 A1 (dependent on previous A1 and B1) ; E1 dep as on MS.