A-level

STATISTICS

SS04 Statistics

Mark scheme

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)	$\bar{x} = 19.14$ $s_{n-1}^2 = 29.1(2125)$	B1		For both. PI Allow for $s_{n-1} = 5.396(411)$ or $s_n^2 = 27.9564$ or $s_n = 5.287381$ All answers truncated or rounded to 3sf.
	$19.14 \pm t_{24;0.005} \frac{5.396}{\sqrt{25}}$	M1		Correct form including use of *t* distribution and **either** use of $\sqrt{25}$ with s_{n-1} **or** use of $\sqrt{24}$ with s_n PI
	with *t* = 2.797	B1		Allow 2.8
		A1		Completely correct numerical expression
	= 19.14 ± 3.01875		5	
	giving limits (16.121, 22.159)	A1		Either form 19.1 ± AWRT(3.02) OR 16.05 ~ 16.2 and 22.1 ~ 22.2
Note: Use of z (=2.5758) gives (16.36,21.92) for first B1 only				
(b)	Because, with new sample values, the sample SD might increase to such an extent that $t \times \frac{SD}{\sqrt{n}}$ is larger, making the interval wider.	E1		Some consideration of SD/variance/variation.
		E1	2	Needs recognition that SD may increase **and** clear acknowledgement that it may more than counteract the effect of increasing n (and decreasing t).
	Total		**7**	

Q2	Solution	Mark	Total	Comment
(a)	$(0.01 \times 0.999) + (0.99 \times 0.001)$ OR $1 - (0.01 \times 0.001) - (0.99 \times 0.999)$ = 0.01098	 M1 A1	 2	 Either expression. PI 0.0109 ~ 0.011(0)
(b)	From B(40, 0.01) P(X > 1) = 1 – 0.9393 = 0.0607	M1 A1	 2	PI by 0.9393 or 0.9925 or 0.0075 or 0.669 or 0.331 seen Allow 0.061
(c)	Use Poisson approximation with mean 400×0.001 = 0.4 $P(1 \leq Y \leq 3) = P(Y \leq 3) - P(Y \leq 0)$ = 0.9992 – 0.6703 =0.3289	B1 M1 M1 A1	 4	Poisson stated PI. Condone slip in mean (0.9992 or 0.9999) MINUS (0.6703 or 0.9384) Allow 3dp accuracy for M1M1 CAO to 4dp
Notes	*(i) More accurate Poisson probs on calculator are 0.9992224 – 0.67032 = 0.3289(04) Max 4/4* *(ii) SC Exact binomial probs are 0.999233 – 0.670186 = 0.3290(43) allow max M1 for 0.999(2) and/or 0.670(2)* *(iii) Use of Normal approximation 0/4*			
(d)	P(double yolk) is not the same for each egg. **or** Method of selecting eggs is not given (maybe not a random sample). This may mean the event of having a double yolk may not be independent between eggs. So it is not reasonable model.	E1 E1dep	 2	Needs **some** context. "p not constant" or "trials not independent" or "sample not random" is not enough. Dependent on correct justification.
	Total		**10**	

Q3	Solution	Mark	Total	Comment
(a)(i)	$H_0 : p = 0.2$ $H_1 : p \neq 0.2$	B1		For both (or π or in words)
	Under H_0 , the number who have not seen the first film ~ B(20, 0.2)	M1		Use of correct distribution PI
	Then P(X ≥ 8) = 1 – 0.9679	m1		For finding P(X ≥ 8) or P(X > 8) from binomial distribution. PI. Allow for P(X > 8) = 0.01 or 0.99 or 0.9679
	= 0.0321	A1		(0.032 ~ 0.033)
	0.0321 > 0.025 Cannot reject H_0 at the 5% level	A1		Comparison with 0.025 OE and accept H_0
	No evidence of **difference** from **20** per cent for **customers/martian zombies**.	A1dep	6	OE. Correct conclusion in context, not too definite. Depends on all previous marks except first B1
Note	*Using P(X > 8) gives P=0.01then reject H_0 for max 4/6 B1B1M1A0AF1A0*			
(ii)	Using the same distribution as in (a) to find P(X ≥ 9)	M1		For finding P(X ≥ 9) or P(X > 9) PI by 0.01 or 0.99 or 0.0026 or 0.9974
	= 0.0100 from tables (0.00998179from calculator)	A1		cao
	Hence, required sig level = 0.02 or 2%	A1	3	cwo 0.0199 ~ 0.02 or 1.99% ~ 2% (1.99636%)
(b)	People usually go to cinema in couples or groups so if one has already seen the first film, it's more likely that others in the same group will also have seen it.	E1		Allow some contextualised reason why the sample is not random (eg "all from the same screening" or "only one group of viewers").
	Not reasonable because not likely to be independent trials.	E1dep	2	For "No" or "Not reasonable" **and** previous E1. **SC** If "independent trials" not put in context, max E1 **SC** "Sample large so would use normal approx. rather than exact binomial", max E1.
	Total		**11**	

Q4	Solution	Mark	Total	Comment
(a)(i)	U ~ N(2.20, 0.08^2) V ~ N(5.26, 0.13^2) $\frac{2-2.20}{0.08}$ = –2.5 $P(Z > -2.5)$ = 0.99379 $\frac{5-5.26}{0.13} = -2$ $P(Z > -2)$ = 0.97725	M1 m1 A1		Standardisation seen once PI Correct tail (both probabilities > 0.5) Both answers to 3sf
	Answer = 0.99379×0.97725	M1		Multiply their two probs. Independent of first M1
	= 0.971(181)	A1dep	5	0.971 ~ 0.972. Dep previous 4 marks.
(ii)	Mean = 2.20 + 5.26 = 7.46 Standard deviation = $\sqrt{0.08^2 + 0.13^2}$	B1		7.46 CAO
	= 0.1526(43)	B1	2	0.1515 ~ 0.1535 Allow also $\sqrt{AWRT\ 0.023}$ seen. Not just 0.023
(b)	Use of: W = (U_1+ U_2+ U_3+ U_4+ U_5) – (V_1+V_2) E(W) = 5×2.20 – 2×5.26 = 0.48 Var(W) = 5 × 0.08^2 + 2 × 0.13^2 = 0.0658	M1 A1 A1		PI CAO or SD 0.256 ~ 0.257
	$P(W > 0) = P\left(Z > \frac{-0.48}{\sqrt{0.0658}}\right)$	m1		Ignore sign. Method for standardisation PI and correct tail (prob > 0.5)
	$= P(Z > -1.87(12))$ = 0.969(26) from tables	A1	5	0.969 ~ 0.97(0) From calculator, prob = 0.96934

Note: Use of 5^2 and 2^2 for 5 and 2 in Var(W) gives Var=0.2276, z = -1.01 and prob 0.8428 for max M1A1A0m1A0 3/5

(c)(i)	0.5	B1	1	CAO
(ii)	Let X be the total weight of 8 potatoes Then price (£) of 8 potatoes Y = 0.90X E(X) = 0.25 × 8 = 2 Var(X) = 8 × 0.015^2 = 0.0018 E(Y) = 1.8 Var(Y) = 0.9^2 × Var(X) = 0.001458 Require price less than 1.85 $P(Y < 1.85) = P(Z < 1.31)$ $= 0.90490$ from tables *Alternative* Maximum weight for 8 individual potatoes to be cheaper is: 1.85/0.9 = 2.055… Then P(X < 2.0555…) from the N(2, 0.0018) distribution = 0.90490	 M1 A1 m1 m1 A1 (M1) (B1) (m1) (m1) (A1)	 5	 Working with correct rv. PI. (SD = 0.04243) CAO Their var(X) × 0.9^2. (SD = 0.03818) Attempt at this normal probability. PI. Requires correct tail (answer > 0.5) 0.904 ~ 0.905 From calculator, prob = 0.90481 Method for max weight. PI. 2.05 ~ 2.06 Attempt at this probability From the correct normal distribution. 0.904 ~ 0.905
Notes (i) M0 if working with some linear combination of rv's (ii) May work with excess price 0.90X – 1.85 = +/- 0.05. Then prob(excess < 0) = 0.9049				
	Total		**18**	

Q5	Solution	Mark	Total	Comment
(a)	$H_0 : \mu = 30$ $H_1 : \mu > 30$	B1		Both. Or population mean for μ. Next 4 marks are PI.
	$\bar{x} = 56.5556 \quad s = 33.3396$	B1		For 56.5 ~ 56.6 and s_{n-1} = 33.3 ~ 33.4 or s_n =31.4 ~ 31.5 (ignore labels)
	$(t =)\dfrac{56.556-30}{33.3396/\sqrt{9}}$	M1		M1 for use of $\dfrac{s_{n-1}}{\sqrt{n}}$ *or* $\dfrac{s_n}{\sqrt{n-1}}$. Condone z=.
		m1		Correct formula, ignore sign for m1. Or $(t =)\dfrac{56.556-30}{31.4329/\sqrt{8}}$
	= 2.39 (2.389552…)	A1		2.38 ~ 2.4(0)
	1% critical value $t_8 = 2.896$	B1		For 2.896 or 2.9(0) **Alternative** for B1B1 *p* = 0.0219 (0.02 ~ 0.022) for B1 Comparison of their *p* with 0.01 B1
	Accept H_0 at 1% level. **No evidence** of a **greater mean** reduction in **PEFR** values (or **30** mentioned)	A1dep	7	In context. Requires correct TS and critical t (both positive) OR correct *p*-value and 0.01 but still requires positive t. Depends on all previous marks except first B1
Notes	*(i) z test gets B1 B1 M1 m1 A1 B0 A0 for max 5/7* *(ii) One sided CI or Decision Int. potentially full marks from 24.3 < 30 OR 62.2 > 56.6 so accept H_0* *(iii) Two-sided test (or CI) gets B0 B1 M1 m1 A1 B0 A0 for max 4/7*			
(b)	Test would have resulted in accept H_0 for type II error to be possible.	E1		H_0 would have been accepted. PI.
	No evidence of a **greater mean** reduction **than 30** litres per minute in **PEFR** values for this breathing disorder.	E1dep	2	Requires context and previous E1. Not too strong. (Allow mention of 30 **or** PEFR)
Notes (i) Simply defining type II error in context gets E0 (ii) Allow "same conclusion as in (a) for 2/3 iff last A1dep in (a) earned				
	Total		**9**	

Q6	Solution	Mark	Total	Comment
(a)(i)	0.369(0)	B1	1	CAO
(ii)	Poisson with mean 6 × 6.5 = 39 Approximate with normal with mean 39 and SD $\sqrt{39}$ = 6.245 $z = \frac{(50.5-39)}{\sqrt{39}}$ = 1.84 (147) Then P(X>50) = 1 – 0.96712 = 0.03288	B1 M1 m1 A1 m1 A1	6	For Poisson and 6 × 6.5 PI Attempt to use normal approx. with SD = $\sqrt{\text{Their mean}}$ Ignore sign. Allow 49.5 or 51.5 (wrong CC), 50 or 51 (no CC) for 50.5 1.84 ~ 1.85 Correct tail of normal distribution (0.032 ~ 0.033) (More accurately, 0.03278)
Notes	*No CC: Use of 50 gives z = 1.76 and p = **0.039**; use of 51 gives z = 1.92 and p = **0.027*** *Wrong CC: Use of 49.5 gives z = 1.68 and p = **0.046**; use of 51.5 gives z =2.00 and p=**0.023*** *These get B1M1m1A0m1A0 for max 4/6*			
(b)(i)	Use of $\lambda \pm z\sqrt{\lambda}$ $= 66 \pm 1.96\sqrt{66}$ $= 66 \pm 15.9(23)$ or (50.1, 81.9)	M1 A1 A1	3	Their λ, z=1.96 or 1.64 or 2.58 only Completely correct expression Either form.AWRT 3sf accuracy.
(ii)	Comparison of like with like ie **either** 6 × 6.5 with the CI in (b)(i) **or** 6.5 with limits in b(i)÷6=(8.35,13.65) The **lower limit** of the CI is greater than 39 or 6.5 (whichever is relevant) So there is evidence that arrival **rate (or mean**) is **greater** in the **evening**	m1 AF1 AFdep 1	3	Requires previous M1. PI (8.3~8.4, 13.6~13.7) Or "the entire CI…" ft their CI in (b)(i) which must be above 39. **Note**: 39 (or 6.5) must be specified. "39 not in interval" OK here but direction needed for next mark. OE. Needs AF1 and a conclusion that is not too definite.
(iii)	Normal distribution used as approx. to Poisson Variance/SD not known and estimated from data (ie SD approximated by $\sqrt{66}$)	E1 E1	2	

(c)	$H_0 : \lambda = 6.5$ $H_1 : \lambda \neq 6.5$	B1		For both. Allow μ or "rate".
	Find $P(X \leq 2)$ from Poisson tables	M1		Attempt to calculate $P(X \leq 2)$ or $P(X < 2)$ (= 0.0113)
	= 0.043(0)	A1		
	This is < 0.05 so reject H_0. There **is** evidence of a **difference** in morning arrival **rates** between the two halls.	M1 A1dep	5	Compare their Poisson prob with 0.05 (OR 2×prob with 0.10) Correct P-value and 0.05, including conclusion in context. Not too definite. Dependent on all previous marks except first B1
	Total		**20**	

	TOTAL MARK FOR PAPER		**75**	