A-LEVEL

STATISTICS

Statistics 5 – SS05

Mark scheme

6380

June 2014

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

Q	Solution	Marks	Total	Comments
1	*s* = 38.2 or s^2 = 1460	B1		awfw 38.2~38.3 or 1450 ~ 1460 accept $\sum(x-\bar{x})^2$=16100 (3sf)
	H$_0$: σ = 30 H$_1$: σ ≠ 30	B1		Both hypotheses o.e.; must use σ
	$\frac{(n-1)S^2}{\sigma^2} = \frac{11\times 38.2^2}{30^2}$	M1		M1 "their" $\frac{11s^2}{\sigma^2}$.
	$= \frac{16065.66}{30^2}$			
	= 17.85	A1		awfw 17.8 ~17.9
		B1		B1 for 11 df; *p* value 0.170 ~ 0.175
	cv's χ^2_{11}= 3.816, 21.92	B1,B1		B1 for each χ^2 value; awfw 3.81~3.82 and 21.9~22.0
	3.816 < 17.8 < 21.92 accept H$_0$ Insufficient evidence at the 5% level to doubt that the standard deviation of volumes of milk delivered by the machine is 30ml.	A1		All correct and statement in context. Comparison with both ends.
			8	

Q	Solution	Marks	Total	Comments
2(a)	$H_0: \mu_{females} - \mu_{males} = 1.5cm$ $H_1: \mu_{females} - \mu_{males} > 1.5cm$	B1 B1	**2**	B1 : an inequality <u>and</u> 1.5 B1 : both correct
(b)	t.s. = $\frac{(8.54-6.28-1.5)}{\sqrt{\left(\frac{0.6^2}{10}+\frac{0.6^2}{8}\right)}}$ = 2.67 z = 2.3263 2.67 > 2.3263 - reject H_0 Sufficient evidence at 1% level to suggest that the <u>mean length of female toads</u> is more than 1.5cm greater than the <u>mean length of male toads.</u>	M1 M1 A1 B1 A1 E1dep	**6**	numerator , allow "8.54 – 6.28" seen denominator awfw 2.67~2.675 Ignore sign , *p* - value : 0.00378 ~ 0.00379 comparison of correct cv with correct t.s; both + ve or both - ve. Statement in context – dependent on previous A1.
			8	
	Notes: Hypotheses: must use μ or population mean. Allow *f* and *m* as suffices but other suffices only if clearly assigned.			

Q	Solution	Marks	Total	Comments
3(a) (i)	mean = 0	B1		cao
	s.d $=\sqrt{\frac{(0.5-(0.5))^2}{12}}$	M1		or $\sqrt{\frac{1}{12}}$; must have square root
	$=\frac{1}{\sqrt{12}}$ = 0.289	A1		$\frac{1}{\sqrt{12}}$ or awfw 0.288 ~0.289
				s.c. B1 for $\frac{1}{12}$ identified as variance
			3	
(ii)	1 – (0.6) or 0.2 + 0.2	M1		B1 for 0.2 or 0.6 seen
	= 0.4	A1		
			2	
(b)	Error more than 0.3	M1		Must be using rectangular distribution.
	P(X > +0.3) = 0.2	A1	**2**	or 0.4 ÷2
			7	

Q	Solution	Marks	Total	Comments
4	$\overline{x}$=511.5	B1		cao
	s = 11.7 , s^2 = 137 or $\sum(x-\overline{x})^2 = 1230$	B1		awfw 11.69~11.71, 136~ 137 or 1230~1231
	$v = 9$	B1		B1 for 9 df
	$t_9 = 2.262$	B1		
	CI: $511.5 \pm 2.262 \times \frac{11.7}{\sqrt{10}}$	M1		M1 : Use of their $\frac{s}{\sqrt{10}}$
		m1		m1 : method for interval ; allow incorrect t_9 value
	[503,520] or 511.5 ± 8.36	A1	**7**	awfw (503.0~503.2, 519.8 ~520 or 511.5 ± 8.362 ~ 8.365)
(b)	$\chi_9^2 = 3.325, 16.919$ CI limits for variance:	B1		either
	$\frac{9 \times 11.7^2}{16.919}, \frac{9 \times 11.7^2}{3.325}$	M1 m1		M1: method for interval – condone one small slip eg $10s^2$ m1: correct expression - condone use of 4.168 and 14.584 for (incorrect) χ_9^2 values
	CI for sd:	m1		m1: method for sd
	[8.53,19.2]	A1	**5**	A1both : awfw 8.52~8.54,19.2~19.4
(c)	There is no reason to doubt that Ahmed is filling his punnets with a mean of 515 grams of strawberries as 515grams is within Ahmed's confidence interval	B1 ft E1dep		B1ft: comment about "their" CI in (a) E1 dep comment in context dep. A1 in (a) accept eg: "Ahmed is filling his punnets with a sufficient <u>mean</u> weight of strawberries" or "Ahmed's <u>mean</u> packed weight is OK/good/suitable"
	The weights of strawberries in Ahmed's punnets have too much variability as 5g is below the lower limit of the confidence interval for Ahmed's sd.	E1dep E1dep	**4**	E1dep : below CI dep A1 in (b) E1dep : comment on CI in (b) in context dep A1 in (b) ; accept eg Ahmed's packing is too <u>variable</u> in weight or "Ahmed's punnets need to be more <u>consistent</u> in weight "
			16	

Q	Solution	Marks	Total	Comments
5(a)	$s_k = 0.680$ $s_a = 0.646$ ($s_k^2 = 0.462$ $s_a^2 = 0.418$)	B1		awfw 0.679 ~ 0.680, 0.646 ~ 0.647 (0.462 ~ 0.463,0 .417~ 0 .418)
	$H_0 : \sigma_k^2 = \sigma_a^2$ $H_1: \sigma_k^2 \neq \sigma_a^2$	B1		Both hypotheses. Use of other suffices must be clearly defined.
	t.s. $F = \frac{0.680^2}{0.646^2} = 1.107$	M1 A1		M1 : using their s^2 A1 : awfw 1.10 ~ 1.11
	df 6,8 c.v. = 4.652	B1,B1		*p*- value 0.868~ 0.869
	Accept H_0 , there is no significant evidence that the samples come from populations with different variances.	E1dep		E1: conclusion in context, dependent on A1 for t.s. and B1 for cv
			7	
(b)	$\bar{x}_k = 9.77$ $\bar{x}_a = 9.14$	B1		Both means awfw (9.77~ 9.78 , 9.14~ 9.15)
	$H_0: \mu_k = \mu_a$ $H_1: \mu_k > \mu_a$	B1		Both hypotheses. Use of other suffices must be clearly defined
	Pooled variance: $s_p^2 = \dfrac{6 \times 0.680^2 + 8 \times 0.646^2}{7 + 9 - 2} = 0.437$	M1A1		awfw 0.436 ~ 0.437 (s_p = 0.660 ~ 0.661)
		M1		M1 (numerator)(accept 9.14 - 9.77)
	$t = \dfrac{9.77-9.14}{\sqrt{0.437(\frac{1}{7}+\frac{1}{9})}} = 1.88$	m1		m1 (denominator – ft on their S_p^2 but must have 1/7 and 1/9.)
		A1		A1 :awfw 1.88 ~ 1.89, ignore sign ignore sign
	cv $t_{14} = 1.345$	B1		*p*-value : 0.0403~0.0404
	1.88 > 1.345 reject H_0	A1dep		A1dep; comparison and conclusion; dependent on A1 for t.s. and B1 for cv. (signs must be consistent with the alternative hypothesis.)
	evidence at 10% level that Kanwar's car travels, **on average**, more miles to the litre than Ashok's car	E1dep		E1 conclusion in context; dependent on previous A1
			10	
(c)	Not reasonable: miles travelled per litre of petrol does not depend on distance travelled.	E1		
			1	
			18	

Q	Solution	Marks	Total	Comments
6(a)(i)	P(X>2) = 1 - 0.6767	B1		B1: λ= 2 used
		M1	**3**	M1 : 1 - 0.6767 (allow 1 - 0.4060)
	= 0.323			
		A1		awfw (0.323~0.324)
(ii)	Emails must be : independent	B1		B1: Any 2 correct assumptions
	occur at a constant rate	E1	**2**	E1: assumptions in context
	occur at random			
(b)	H_0: data can be modelled by an exponential distribution with parameter $\lambda = \frac{2}{15}$	B1		B1 both hypotheses (accept mean = 7.5)
	H_1: data cannot be modelled by an exponential distribution with parameter $\lambda = \frac{2}{15}$	B1		B1: $\lambda = \frac{2}{15}$ seen or used: accept $\frac{1}{7.5}$ or 0.133 or better

x	O	P(x)	E	$(O - E)^2$	$(O-E)^2/E$
0 -	124	0.3297	105.50	342.3394	3.2450
3 -	82	0.2210	70.72	127.3028	1.8002
6 -	51	0.1857	59.43	71.1355	1.1969
10 -	31	0.1283	41.04	100.8778	2.4578
15 -	16	0.0659	21.07	25.7311	1.2211
20 -	11	0.0512	16.37	28.8767	1.7636
>30	5	0.0183	5.86	0.7413	0.1265
			Total		**11.8110**

Q	Solution	Marks	Total	Comments
		M1		M1: method for probabilities ; at least one value correct.
		m1		m1 expected values ft “p” × 320
		m1		m1 combining classes “30 and over”
		m1		m1 attempt at $(O - E)^2$
		m1		m1 dividing $(O - E)^2$ by E and summing – at least 2 values seen.
		A1		All correct and awfw 11.8 ~12.0
	$\nu = 7 - 1 = 6$	B1ft		df – ft their number of classes - 1
	$\chi^2_6 = 10.645$	B1		awfw 10.6 ~ 10.7
				p-value 0.0663~ 0.0666
	11.811 > 10.645 ; reject H_0 ; data do not fit an exponential distribution with parameter $\lambda = \frac{2}{15}$	E1dep	**11**	E1 conclusion in context dependent on A1 for ts and B1 for cv.
c)	Data cannot be modelled adequately by an exponential distribution and hence the Poisson model is unlikely	E1dep		E1 : accept either Poisson model is unlikely or data lacks fit to Exp(2/15) dependent on a χ^2test in (b).
	The answer in a(i) is not likely to be valid.	E1dep	**2**	E1 dep on first E1 - invalid
			18	