Version 1.0: 0110

General Certificate of Education

Statistics 6380

SS04 Statistics 4

Mark Scheme

2010 examination – January series

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 meeting attended by all examiners and is the scheme which was used by them in this examination. The standardisation meeting 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 the standardisation meeting each examiner analyses a number of candidates' scripts: alternative answers not already covered by the mark scheme are discussed at the meeting and legislated for. If, after this meeting, examiners encounter unusual answers which have not been discussed at the meeting 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.

Set and published by the Assessment and Qualifications Alliance.

Key to mark scheme and abbreviations used in marking

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	MC	mis-copy
CAO	correct answer only	MR	mis-read
CSO	correct solution only	RA	required accuracy
AWFW	anything which falls within	FW	further work
AWRT	anything which rounds to	ISW	ignore subsequent work
ACF	any correct form	FIW	from incorrect work
AG	answer given	BOD	given benefit of doubt
SC	special case	WR	work replaced by candidate
OE	or equivalent	FB	formulae book
A2,1	2 or 1 (or 0) accuracy marks	NOS	not on scheme
$-x$ EE	deduct x marks for each error	G	graph
NMS	no method shown	c	candidate
PI	possibly implied	sf	significant figure(s)
SCA	substantially correct approach	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. However, there are situations in some units where part marks would be appropriate, particularly when similar techniques are involved. Your Principal Examiner will alert you to these and details will be provided on the mark scheme.

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.

SS04

Q	Solution	Marks	Total	Comments
1(a)	Binomial $n = 50$ $p = 34/50 = 0.68$ $\rightarrow$ normal mean 34, s.d. $\sqrt{34 \times 0.32} = 3.298$ 95% confidence interval for p is $0.68 \pm 1.96 \times \sqrt{0.68 \times 0.32/50}$ 0.68 ± 0.129 $0.551 \sim 0.809$	B1 M1 m1 M1 B1 A1	 6	B1 $p =$ 34/50 or 16/50 acf M1 attempt to use normal m1 correct method for mean and s.d. for number or proportion M1 method for confidence interval B1 1.96 A1 0.551(0.55 ~ 0.551) and 0.809 (0.809 ~ 0.81) allow in ± form
(b)	Since 0.9 lies above the interval the cyclist's claim is not supported	B1 E1√	 2	B1 0.9 lies above interval - based on correct method E1√ correct conclusion their interval
		Total	**8**	
2(a)(i)	Binomial $n = 20000$ $p = 0.0001$ $\rightarrow$ Poisson, mean $20000 \times 0.0001 = 2$ P(0) = 0.135	B1B1 M1 A1	 4	B1 binomial B1 20000 and 0.0001 M1 Poisson mean 20000×0.0001- allow slip A1 0.135 (0.135 ~ 0.136)
(ii)	P(>5) = 1 – P(5 or fewer) = 1 – 0.9834 = 0.0166	M1 A1	 2	M1 P(>5) = 1 – P(5 or fewer) A1 0.0166 (0.016 ~ 0.017)
(b)	Binomial $n = 80$ $p = 0.32$ $\rightarrow$ Normal, mean $80 \times 0.32 = 25.6$ variance = $80 \times 0.32 \times 0.68 = 17.408$ s.d. = 4.1723 $z = (20.5 - 25.6)/4.1723 = -1.2223$ P(>20) = 0.889	B1 M1 A1 M1 m1 A1	 6	B1 B(80,0.32) M1 attempt at normal approx A1 mean = 25.6 variance = 17.408 (17.4~17.41) or s.d. = 4.17 (4.17 ~ 4.175) disallow if wrongly used M1 method for z - ignore sign and c.c. m1 correct attempt at c.c.- ignore sign of z A1 0.889 (0.888 ~ 0.89)
		Total	**12**	
3(a)	$\bar{x} = 29.5125$ $s = 8.6331$ H_0: $\mu = 37.5$ H_1: $\mu < 37.5$ $t = (29.5125 - 37.5)/(8.6331/\sqrt{8})$ $= -2.62$ c.v. t_7 -1.895 Reject H_0 There is significant evidence that the mean amount of fuel bought by a customer on each visit is less than 37.5 litres.	B1 B1 M1 m1 A1 B1 B1√ A1√ A1√	 9	B1 29.5 (29.5~29.52) and 8.63 8.63~8.64) B1 both hypotheses M1 use of their s.d./$\sqrt{8}$ m1 method for t - ignore sign A1 –2.62 (2.61 ~ 2.62) B1 7df B1√ 1.895 - their df A1√ conclusion must be compared with lower tail of t and not inconsistent with their H_0. Allow arithmetic errors and incorrect t-values only A1√ in context - requires previous A mark

SS04 (cont)

Q	Solution	Marks	Total	Comments
3(b)	$H_0 : \lambda = 168$ $H_1 : \lambda < 168$	B1		B1 hypotheses - allow $\lambda = 84$
	Poisson mean 168 $\rightarrow$ N(168,168)	M1		M1 attempt at normal approximation
	$z = (142.5 - 168)/\sqrt{168} = -1.97$	M1		M1 method for z –ignore sign
	[or $(142 - 168)/\sqrt{168} = -2.01$	A1		A1 –1.97 (–1.96~ –1.97) or –2.01 (–2.00 ~ –2.01)
	or $(71 - 84)/\sqrt{84/2} = -2.01$]	B1		B1 –2.3263 ignore sign
	c.v. –2.3263	A1√		A1√ conclusion - must be compared with lower tail of z
	Accept H_0, no significant evidence at 1% level to show that mean number of customers has been reduced.	A1√	7	A1√ in context - requires previous A mark
	or $p = 0.024$ (0.024~0.025) or 0.022 (0.022~0.023) compare with 0.01			
(c)	$H_0 : p = 0.20$ $H_1 : p > 0.20$	B1		B1 hypotheses
	Binomial $n = 20$ $p = 0.2$	M1		M1 use of binomial $n = 20$ $p = 0.2$
	P(5 or more) = 1 – 0.6296 = 0.370	A1		A1 0.370 (0.37~0.371)
	>0.1	A1√		A1√ conclusion - requires comparison of value from B(20,0.2) with 0.1
	Accept H_0, no significant evidence that the proportion of customers who do not buy fuel has increased.	A1√	5	A1√ in context - requires previous A mark
(d)	There is significant evidence that the amount of fuel bought on each visit has reduced.	E1√		E1√ Any point consistent with their results
	There is some evidence of reduction in the number of customers - at least on Friday afternoon but this evidence is not significant at the 1% level.	E1√		E1√ second point consistent their results E1 three points based on correct results and methods
	No significant evidence that the proportion of customers who do not buy fuel has increased.	E1		E1 earned either for overall comment or for comment that there is evidence of a reduction in number of customers but it is not significant.
	Overall he is right to be concerned.	E1	4	
		Total	**25**	
4(a)	$z = (120 - 110)/25 = 0.4$	M1		M1 attempt to find probability of <120 minutes, from normal.
	P(<120) = 0.65542	M1		M1 method for probability their minutes- allow wrong tail
		A1	3	A1 0.655 (0.655~0.656)
(b)	$2T$ is normally distributed with mean 220	B1		B1 mean 220
	minutes and standard deviation 50 minutes.	B1		B1 s.d. 50 or variance = 2500
	$z = (180 - 220)/50 = -0.8$	M1		M1 method - including method for distribution of $2T$ - allow wrong tail and slip in number of minutes to 12 noon
	probability taxi before noon is $1 - 0.78814 = 0.212$	A1	4	A1 0.212 (0.211 ~ 0.213)

SS04 (cont)

Q	Solution	Marks	Total	Comments
4(c)	Time for two appeals plus break is $T_1 + B + T_2$ normally distributed with mean = 110 +12 + 110 = 232 Variance $25^2 + 4^2 + 25^2 = 1266$ standard deviation = 35.58 $z = (180 - 232)/\sqrt{1266} = -1.461$ Probability second appeal not completed by noon is 0.928	M1 M1 A1 m1 A1	 5	M1 method for mean M1 method for s.d. or variance Allow for adding variances of independent variables even if model incorrect A1 232 and 35.58 or 1266 m1 correct method allow wrong tail and slip in number of minutes to 12 noon A1 0.928 (0.9275 ~ 0.9285)
(d)(i)	$B + T_2 - T_1$ has mean 12 and standard deviation 35.58. $z = 12/\sqrt{1266} = 0.3373$ $P(<0) = 1 - 0.632 = 0.368$	M1 m1 A1	 3	M1 method for mean and s.d./variance m1 method - allow wrong tail A1 0.368 (0.366 ~0.371)
(ii)	Taxi is due T_1 minutes after first appeal is completed. Second appeal is completed $B + T_2$ minutes after first appeal is completed. $\therefore$ Second appeal completed before taxi due if $B + T_2 < T_1$ i.e. $B + T_2 - T_1 < 0$	E1 E1	 2	E1 reasonable attempt E1 complete explanation
		Total	**17**	
5(a)(i)	$\bar{x} = 134.51 \quad s = 1.0181$ 95% confidence interval for mean $134.51 \pm 2.262 \times 1.0181/\sqrt{10}$ 134.51 ± 0.728 $133.78 \sim 135.24$	B1 M1 m1 B1 B1$\surd$ A1	 6	B1 134.51 (134.5 ~ 134.52) and 1.0181 (1.018 ~ 1.02) M1 use of their s.d/$\sqrt{10}$ m1 correct method for t B1 9df B1$\surd$ 2.262 their df A1 133.78 and 135.24 from correct working AG
(ii)	As all lengths start with 13, 3sf is in effect 1sf which is too few. **or** Width of confidence interval is 1.456 - if the limits had been rounded to 3sf the width would have apparently been 1 – a large % error.	E1	1	E1 reason
(b)	**Statement 1** C; Interval is based on the sample mean. There is no reason why 95% of individual lengths should lie in the interval. There is a very small possibility that it could occur by chance. **Statement 2** C; this would be true for an interval based on a known population mean and s.d. using z. It is extremely unlikely to be true of an interval based on estimates and t. **Statement 3** D; the interval is centred on x and so is certain to contain x.	B1 E1 B1 E1 B1 E1	 6	B1 C E1 explanation - allow both marks for a good explanation even if option D chosen. B1 C E1 explanation-allow both marks for a good explanation even if option D chosen. B1 D E1 explanation
	Total		**13**	
	TOTAL		**75**	