General Certificate of Education

Statistics 6380

SS02   Statistics 2

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.

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.

SS02

Q	Solution	Marks	Total	Comments
1(a)	4	B1	1	CAO
(b)	short term variability about upward linear trend	E1 E1 E1	3	short term variability upward linear allow random variation about upward non-linear trend
(c)	`(-156 -216 -143)/3 = -172	M1 A1 A1	3	attempt to find mean deviation from line or by calculation -172 (-168 ~ -175) ignore sign negative sign
(d)	880 - 172 = 708 Estimated expenditure £708m s.c. B2 for answer within range with no or unclear method	B1 M1 A1	3	m.a. for Question 3 estimated from trend line seasonal effect subtracted from their trend 710 (705 ~ 715) allow 700 ignore units - disallow if more than 3sf given
	Total		**10**	
2(a)(i)	Although there had been a reduction in recorded crime nearly half the respondents thought there had been a large increase. Only a very small proportion (4%) correctly thought that there had been a decrease.	E1 E1	2	reduction in recorded crime most respondents believe there is at least as much
(ii)	The media highlight particularly horrific crimes, giving the impression there is more crime than is actually the case. Stories about crime are passed on from person to person and become exaggerated so that people think there is more crime than is actually the case. People whose friends or relatives experience a crime pay more attention to this than to national statistics. Amount of recorded crime may not accurately reflect the actual amount of crime.	E2(1)	2	E2 (1) Both marks for any valid well explained possible reason. May be earned in (b)(ii)
(iii)	The answers are divided into four non-numerical discrete categories.	E1	1	box and whisker requires numerical data
2(b)(i)	Although only a small proportion of crimes reported to the police (6%) are violent in nature the great majority of respondents (78%) believed that over 30% of crime involved violence or the threat of violence.	E1 E1	2	small proportion of reported crimes are violent majority of respondents overestimate the proportion of crimes which are violent
2(b)(ii)	(ii) as a(ii)	E1	1	any valid reason not used in (a)(ii) - may be earned in (a)(ii)
	Total		**8**	

SS02 (cont)

Q	Solution	Marks	Total	Comments
3(a)	$H_0: \mu = 90$ $H_1: \mu > 90$	B1		one hypothesis correct
	$\bar{x} = 109.56$	B1		both hypotheses correct
		M1		use of $55/\sqrt{9}$
	$z = (109.56 - 90)/(55/\sqrt{9})$	m1		correct method for z - ignore sign
	$= 1.07$	A1		1.07 (1.06 ~ 1.07)
	c.v. for 5% test is 1.6449	B1		1.6449 (1.64~1.65) - ignore sign
	Accept H_0. Conclude no significant evidence that the mean waiting time for calls made to Northgas exceeds 90 seconds.	A1√ A1√		conclusion - must compare correct tail of z conclusion in context
	(p-value 0.143)		8	
(b)	$H_0: \mu = 90$ $H_1: \mu > 90$	M1 m1		use of $12/\sqrt{85}$ method for z - ignore sign
	$z = (94 - 90)/(12/\sqrt{85})$			
	$= 3.07$	A1		3.07 (3.07~3.08)
	c.v. for 5% test is 1.6449	A1√		conclusion - must compare correct tail of z
	Reject H_0. Conclude there is significant evidence that mean waiting time for calls made to Southgas exceeds 90 seconds. (p-value 0.00107)	A1√	5	conclusion in context
	Apply mark scheme for (a) to (b) and vice versa if more favourable to candidate.			
(c)(i)	Sample mean in (a) greater than in (b) but population mean accepted as equal to 90 in (a) but concluded to be greater than 90 in (b).	E1 E1	2	sample mean greater in (a) than (b) comparison of conclusions
(ii)	Larger sample in (b) makes any difference from 90 more likely to be detected / More variable sample in (a) makes any difference from 90 less likely to be detected.	E1	1	sample size or variability
	Total		**16**	

SS02 (cont)

Q	Solution	Marks	Total	Comments
4(a)	Plot points: (1, 2166), (2, 2383), (3, 2638), (4, 2825), (5, 3017), (6, 3248), (7, 3402), (8, 3433), (9, 3475)	M1 A1	 2	method of plotting reasonably accurate plot – by eye
(b)	$y = 2093 + 172.3t$ $t = 0$ $y = 2093$ $t = 10$ $y = 3816$	B2 B1 M1 A1	 5	172.3 (172~173) one mark for (170~175) 2093 (2090~2100) method for line - their equation accurate line - by eye
(c)	Forecast of number of screens in UK in 2006 is $2093 + 172.3 \times 11$ $= 3988$	M1 A1	 2	attempt to substitute 11 in their equation 3988 (3980~4000) - disallow if not whole number
(d)	Graph suggests actual figure will be below the regression line - say 3500	M1 A1	 2	forecast reduced by a non-trivial amount 3500 (3400 ~ 3800)
	Total		**11**	
5(a)	0.8335	B1	1	0.834 (0.833 ~ 0.834)
(b)(i)	0.0273	B1	1	0.0273 (0.027~0.0274)
(ii)	0.3027 - 0.1257 = 0.177	M1 A1	 2	P (2 or fewer) - P(1 or fewer) 0.177 (0.1765 ~ 0.1775)
(iii)	$P_o(5)$ P (>6) = 1 - 0.7622 = 0.2378	B1 M1 A1	 3	$P_o(5)$ method - their mean 0.238 (0.237 ~ 0.238)
(c)	$E(Y) = 0 \times 0.24 + 1 \times 0.27 + 2 \times 0.29$ $+ 3 \times 0.12 + 4 \times 0.08$ $= 1.53$ $E(Y^2) = 0^2 \times 0.24 + 1^2 \times 0.27 + 2^2 \times 0.29$ $+ 3^2 \times 0.12 + 4^2 \times 0.08$ $= 3.79$ $V(Y) = 3.79 - 1.53^2$ $= 1.4491$ s.d. $= \sqrt{1.4491} = 1.20$	M1 A1 M1 m1 A1	 5	method for E(Y) 1.53 CAO method for E(Y^2) method for s.d. 1.20 (1.20 ~ 1.21) SC allow max 4 for variance = 1.45 (1.44 ~1.46)
(d)	s.d. of $Z = \sqrt{2.89} = 1.7$	E1	1	$\sqrt{2.89}$ AG
(e)(i)	X	B1		CAO
(ii)	s.d. $W = \sqrt{1.4} = 1.18$, s.d of X, Y and Z are all larger W is least variable	M1 M1 A1	 4	attempt to compare s.d. or variances method for s.d. of W or X W least variable
	Total		**17**	

SS02 (cont)

Q	Solution	Marks	Total	Comments
6(a)	Number examiners 000 to 399	B1 E1		400 examiners - may be implied valid numbering
	Select 3 digit random numbers	E1		select 3 digit random numbers
	Ignore repeats and greater than 399	E1		ignore repeats and out of range
	Continue until 40 selected choose corresponding examiners	E1	5	select 40 and choose corresponding examiners
(b)(i)	Cluster sampling	B1	1	cluster sampling
(ii)	More geographically localised - less travelling	E1	1	less travelling
(iii)	Views will differ between regions e.g. examiners from South East likely to prefer London and examiners from North West likely to prefer Manchester	E1 E1	2	views likely to be more homogeneous in context
(iv)	No - examiners from regions with small number of examiners e.g. North West more likely to be selected than those from regions with a large number of examiners.	B1 E1	2	no explanation allow B1 for no, examiners in regions not chosen have no chance
(c)	(i) (b)(ii) no longer valid since no travelling required (ii) (b)(iii) still valid - using telephone email does not affect it.	E1 E1	2	no travelling needed views will still differ between regions allow E1 for (i) no (ii) yes without explanation
	Total		**13**	
	TOTAL		**75**	