OCR
RECOGNISING ACHIEVEMENT

GCE

Mathematics (MEI)

Advanced GCE

Unit **4767:** Statistics 2

Mark Scheme for January 2013

Annotations and abbreviations

Annotation in scoris	Meaning
✓and ✗	
BOD	Benefit of doubt
FT	Follow through
ISW	Ignore subsequent working
M0, M1	Method mark awarded 0, 1
A0, A1	Accuracy mark awarded 0, 1
B0, B1	Independent mark awarded 0, 1
SC	Special case
^	Omission sign
MR	Misread
Highlighting	
Other abbreviations in mark scheme	Meaning
E1	Mark for explaining
U1	Mark for correct units
G1	Mark for a correct feature on a graph
M1 dep*	Method mark dependent on a previous mark, indicated by *
cao	Correct answer only
oe	Or equivalent
rot	Rounded or truncated
soi	Seen or implied
www	Without wrong working

Subject-specific Marking Instructions for GCE Mathematics (MEI) Statistics strand

a   Annotations should be used whenever appropriate during your marking.

   The A, M and B annotations must be used on your standardisation scripts for responses that are not awarded either 0 or full marks. It is vital that you annotate standardisation scripts fully to show how the marks have been awarded.

   For subsequent marking you must make it clear how you have arrived at the mark you have awarded.

b   An element of professional judgement is required in the marking of any written paper. Remember that the mark scheme is designed to assist in marking incorrect solutions. Correct *solutions* leading to correct answers are awarded full marks but work must not be judged on the answer alone, and answers that are given in the question, especially, must be validly obtained; key steps in the working must always be looked at and anything unfamiliar must be investigated thoroughly.

   Correct but unfamiliar or unexpected methods are often signalled by a correct result following an *apparently* incorrect method. Such work must be carefully assessed. When a candidate adopts a method which does not correspond to the mark scheme, award marks according to the spirit of the basic scheme; if you are in any doubt whatsoever (especially if several marks or candidates are involved) you should contact your Team Leader.

c   The following types of marks are available.

   M
   A suitable method has been selected and *applied* in a manner which shows that the method is essentially understood. Method marks are not usually lost for numerical errors, algebraic slips or errors in units. However, it is not usually sufficient for a candidate just to indicate an intention of using some method or just to quote a formula; the formula or idea must be applied to the specific problem in hand, eg by substituting the relevant quantities into the formula. In some cases the nature of the errors allowed for the award of an M mark may be specified.

   A
   Accuracy mark, awarded for a correct answer or intermediate step correctly obtained. Accuracy marks cannot be given unless the associated Method mark is earned (or implied). Therefore M0 A1 cannot ever be awarded.

   B
   Mark for a correct result or statement independent of Method marks.

   E
   A given result is to be established or a result has to be explained. This usually requires more working or explanation than the establishment of an unknown result.

Unless otherwise indicated, marks once gained cannot subsequently be lost, eg wrong working following a correct form of answer is ignored. Sometimes this is reinforced in the mark scheme by the abbreviation isw. However, this would not apply to a case where a candidate passes through the correct answer as part of a wrong argument.

d   When a part of a question has two or more 'method' steps, the M marks are in principle independent unless the scheme specifically says otherwise; and similarly where there are several B marks allocated. (The notation 'dep *' is used to indicate that a particular mark is dependent on an earlier, asterisked, mark in the scheme.) Of course, in practice it may happen that when a candidate has once gone wrong in a part of a question, the work from there on is worthless so that no more marks can sensibly be given. On the other hand, when two or more steps are successfully run together by the candidate, the earlier marks are implied and full credit must be given.

e   The abbreviation ft implies that the A or B mark indicated is allowed for work correctly following on from previously incorrect results. Otherwise, A and B marks are given for correct work only – differences in notation are of course permitted. A (accuracy) marks are not given for answers obtained from incorrect working. When A or B marks are awarded for work at an intermediate stage of a solution, there may be various alternatives that are equally acceptable. In such cases, exactly what is acceptable will be detailed in the mark scheme rationale. If this is not the case please consult your Team Leader.

Sometimes the answer to one part of a question is used in a later part of the same question. In this case, A marks will often be 'follow through'. In such cases you must ensure that you refer back to the answer of the previous part question even if this is not shown within the image zone. You may find it easier to mark follow through questions candidate-by-candidate rather than question-by-question.

f   Wrong or missing units in an answer should not lead to the loss of a mark unless the scheme specifically indicates otherwise.

Candidates are expected to give numerical answers to an appropriate degree of accuracy. 3 significant figures may often be the norm for this, but this always needs to be considered in the context of the problem in hand. For example, in quoting probabilities from Normal tables, we generally expect *some* evidence of interpolation and so quotation to 4 decimal places will often be appropriate. But even this does not always apply – quotations of the standard critical points for significance tests such as 1.96, 1.645, 2.576 (maybe even 2.58 – but not 2.57) will commonly suffice, especially if the calculated value of a test statistic is nowhere near any of these values. Sensible discretion *must* be exercised in such cases.

Discretion must also be exercised in the case of small variations in the degree of accuracy to which an answer is given. For example, if 3 significant figures are expected (either because of an explicit instruction or because the general context of a problem demands it) but only 2 are given, loss of an accuracy ("A") mark is likely to be appropriate; but if 4 significant figures are given, this should not normally be penalised. Likewise, answers which are slightly deviant from what is expected in a very minor manner (for example a Normal probability given, after an attempt at interpolation, as 0.6418 whereas 0.6417 was expected) should not be penalised. However, answers which are *grossly* over- or under-specified should normally result in the loss of a mark. This includes cases such as, for example, insistence that the value of a test statistic is (say) 2.128888446667 merely because that is the value that happened to come off the candidate's calculator. Note that this applies to answers that are given as final stages of calculations; intermediate working should usually be carried out, and quoted, to a greater degree of accuracy to avoid the danger of premature approximation.

The situation regarding any particular cases where the accuracy of the answer may be a marking issue should be detailed in the mark scheme rationale. If in doubt, contact your Team Leader.

g   Rules for replaced work

If a candidate attempts a question more than once, and indicates which attempt he/she wishes to be marked, then examiners should do as the candidate requests.

If there are two or more attempts at a question which have not been crossed out, examiners should mark what appears to be the last (complete) attempt and ignore the others.

NB Follow these maths-specific instructions rather than those in the assessor handbook.

h   Genuine misreading (of numbers or symbols, occasionally even of text) occurs. If this results in the object and/or difficulty of the question being considerably changed, it is likely that all the marks for that question, or section of the question, will be lost. However, misreads are often such that the object and/or difficulty remain substantially unaltered; these cases are considered below.

The simple rule is that *all* method ("M") marks [and of course all independent ("B") marks] remain accessible but at least some accuracy ("A") marks do not. It is difficult to legislate in an overall sense beyond this global statement because misreads, even when the object and/or difficulty remains unchanged, can vary greatly in their effects. For example, a misread of 1.02 as 10.2 (perhaps as a quoted value of a sample mean) may well be catastrophic; whereas a misread of 1.6748 as 1.6746 may have so slight an effect as to be almost unnoticeable in the candidate's work.

A misread should normally attract *some* penalty, though this would often be only 1 mark and should rarely if ever be more than 2. Commonly in sections of questions where there is a numerical answer either at the end of the section or to be obtained and commented on (eg the value of a test statistic), this answer will have an "A" mark that may actually be designated as "cao" [correct answer only]. This should be interpreted *strictly* – if the misread has led to failure to obtain this value, then this "A" mark must be withheld even if all method marks have been earned. It will also often be the case that such a mark is implicitly "cao" even if not explicitly designated as such.

On the other hand, we commonly allow "fresh starts" within a question or part of question. For example, a follow-through of the candidate's value of a test statistic is generally allowed (and often explicitly stated as such within the marking scheme), so that the candidate may exhibit knowledge of how to compare it with a critical value and draw conclusions. Such "fresh starts" are not affected by any earlier misreads.

A misread may be of a symbol rather than a number – for example, an algebraic symbol in a mathematical expression. Such misreads are more likely to bring about a considerable change in the object and/or difficulty of the question; but, if they do not, they should be treated as far as possible in the same way as numerical misreads, *mutatis mutandis*. This also applied to misreads of text, which are fairly rare but can cause major problems in fair marking.

The situation regarding any particular cases that arise while you are marking for which you feel you need detailed guidance should be discussed with your Team Leader.

Note that a miscopy of the candidate's own working is not a misread but an accuracy error.

Question	Answer	Marks	Guidance	
1 (i)		G1 G2,1,0	G1 For axes suitably labelled with some indication of **linear** scale provided. G2 for points plotted correctly. G1 if 4 points plotted correctly. G0 if two or more incorrectly plotted/omitted points. Special Case SC1 for points visibly correct on axes where no indication of scale has been provided.	Allow *x* & *y* Allow axes reversed.
		[3]		
1 (ii)	Thickness is the independent variable since the values of 'Thickness' are not subject to random variation, but are determined by the manufacturer.	E1 **[1]**	Allow explanations referring to thickness being "controlled" by the manufacturer. Allow equivalent interpretations.	
1 (iii)	$\bar{t} = 60$, $\bar{h} = 1.548$ $b = \dfrac{S_{th}}{S_{tt}} = \dfrac{546.8 - (300 \times 7.74 / 5)}{22000 - 300^2 / 5} = \dfrac{82.4}{4000} =$ 0.0206 OR $\quad b = \dfrac{546.8 / 5 - (60 \times 1.548)}{22000 / 5 - 60^2} = \dfrac{16.48}{800} =$ 0.0206 hence least squares regression line is: $\quad h - \bar{h} = b(t - \bar{t})$	B1 M1* A1	For $\bar{t}$ and $\bar{h}$ used. SOI (e.g. can be implied by $b = 0.0206$) For attempt at **calculating** gradient (b) for h on t. For 0.0206 cao	

Question			Answer	Marks	Guidance	
			$\Rightarrow \quad h - 1.548 = 0.0206\,(t - 60)$	M1 dep*	For equation of line, using their b, $b > 0$, and passing through their $(\bar{t}, \bar{h})$	
			$\Rightarrow \quad h = 0.0206\,t + 0.312$	A1	Final equation must have h as the subject. **CAO** Allow $h = 0.021\,t + 0.31$, Allow $h = 0.021\,t + 0.288$ NOTE If equation given in terms of y and x then A0 unless x & y defined appropriately	
				[5]		
1	(iv)	(*A*)	$(0.0206 \times 70) + 0.312 = 1.754$	B1	Allow 1.75 FT their equation provided $b > 0$	
			Likely to be reliable as interpolation	E1		
				[2]		
1	(iv)	(*B*)	$(0.0206 \times 120) + 0.312 = 2.784$	B1	Allow 2.78 FT their equation provided $b > 0$	
			Could be unreliable as extrapolation	E1	Condone "reliable as 120 is not too far away from the data used to produce the equation"	
				[2]		
1	(v)		Thickness = 40 $\Rightarrow$ predicted max height $\qquad = (0.0206 \times 40) + 0.312 = 1.136$ Residual = 1.09 − 1.136 $= -0.046$	M1 M1 A1	For prediction. FT their equation provided $b > 0$ For difference between 1.09 and prediction. Allow −0.05	
				[3]		
1	(vi)		Regression line gives a prediction of $(0.0206 \times 200) + 0.312 = 4.432$ This is **well above** the observed value. It could be that the relationship breaks down for larger thickness, or that the relationship is not linear	B1* E1 dep* E1	B1 for obtaining a prediction from regression equation or from graph E1 for noting the large difference between prediction and actual value E1 for suitable interpretation regarding the relationship between maximum height and thickness	
				[3]		

Question			Answer	Marks	Guidance	
2	(i)	(A)	$P(X = 0) = \dfrac{e^{-2.1}2.1^0}{0!}$	M1	For calculation	
			$= 0.1225$	A1	**CAO** Allow 0.122	
			Or from tables $P(X = 0) = 0.1225$			
				[2]		
2	(i)	(B)	$P(X \geq 2) = 1 - P(X \leq 1) = 1 - 0.3796$	M1	M1 for **use of** correct structure. i.e. M0 for use of $1 - P(X \leq 2)$ or $1 - 0.6796$ Using $\lambda = 2.0$ leading to $1 - 0.4060$ gets M1	
			$= 0.6204$	A1	CAO Allow 0.6203, 0.620	
				[2]		
2	(i)	(C)	New $\lambda = 5 \times 2.1 = 10.5$ P(Between 5 and 10 in 5 mins)	B1	For mean (SOI)	
			$= 0.5207 - 0.0211$	M1	For $P(X \leq 10) - P(X \leq 4)$ used.	e.g. $1 - 0.9379$ leads to B0M1A0
			$= 0.4996$	A1	CAO Allow 0.500, 0.50. Condone 0.5 www.	
				[3]		
2	(ii)		Mean number in 60 minutes $= 60 \times 2.1 = 126$ Using Normal approx. to the Poisson, $X \sim N(126, 126)$	B1 B1	For Normal approx. For correct parameters (SOI)	
			$P(X \geq 130) = P\left(Z \geq \dfrac{129.5 - 126}{\sqrt{126}}\right)$	B1	For correct continuity correction	
			$= P(Z > 0.3118) = 1 - \Phi(0.3118)$ $= 1 - 0.6224$	M1	For correct probability structure	
			$= 0.3776$	A1	CAO, (Do not FT wrong or omitted CC). Allow 0.378www & 0.3775	
				[5]		

Question			Answer	Marks	Guidance	
2	(iii)		(Because if butterflies are blown in pairs,) the events will no longer be occurring singly.	E1		Accept 'not independent'
				[1]		
2	(iv)		P(3 or fewer) = P(3 or fewer individuals and no pairs) + P(0 or 1 individual and 1 pair)	M1	For P(0 pairs) (= 0.8187)	
			= $(0.9068 \times 0.8187) + (0.4932 \times (0.9825 - 0.8187))$	M1	For P(1 pair) (= 0.1638 or 0.1637)	First two M1s can be awarded for 0.9825
			= $(0.9068 \times 0.8187) + (0.4932 \times 0.1638)$	M2	For structure M2 for correct 6 combinations identified and their probabilities added, M1 for 5 combinations identified and their probabilities added.	
			= 0.7424 + 0.0808			
			= 0.8232	A1	CAO Allow awrt 0.823	
			Or			
			using D for the number of pairs and S for the number of singles			
			$P(D = 0) \times P(S = 0) = e^{-0.2} \times e^{-1.7}$ $= 0.1495\ldots$			
			$P(D = 0) \times P(S = 1) = e^{-0.2} \times 1.7e^{-1.7}$ $= 0.2542\ldots$			
			$P(D = 0) \times P(S = 2) = e^{-0.2} \times 1.7^2 e^{-1.7} \div 2$ $= 0.2161\ldots$			
			$P(D = 0) \times P(S = 3) = e^{-0.2} \times 1.7^3 e^{-1.7} \div 3!$ $= 0.1224\ldots$			
			$P(D = 1) \times P(S = 0) = 0.2e^{-0.2} \times e^{-1.7}$ $= 0.0299\ldots$			
			$P(D = 1) \times P(S = 1) = 0.2e^{-0.2} \times 1.7e^{-1.7}$ $= 0.0508\ldots$			
			Or			
			$P(D = 0) \times P(S = 3) + P(D = 1) \times P(S = 1)$		P(3 butterflies) (= 0.1733…)	
			$P(D = 0) \times P(S = 2) + P(D = 1) \times P(S = 0)$		P(2 butterflies) (= 0.2460…)	
			$P(D = 0) \times P(S = 1)$		P(1 butterfly) (= 0.2542…)	
			$P(D = 0) \times P(S = 0)$		P(0 butterflies) (= 0.1495…)	

Question			Answer	Marks	Guidance	
			Or $P(D \leq 1) \times P(S \leq 1) = 0.98247 \times 0.49324$ $D(D = 0) \times P(S = 2) = 0.21613$ $D(D = 0) \times P(S = 3) = 0.12247$ **[5]**			
3	(i)	(A)	$P(X < 450) = P\left(Z < \dfrac{450 - 435}{30}\right)$ $= P(Z < 0.5) = \Phi(0.5)$ $= 0.6915$ **[3]**	M1 M1 A1	For standardising. M0 if 'continuity correction' applied For correct structure CAO Allow 0.692	
3	(i)	(B)	$P(400 < X < 450)$ $= P\left(\dfrac{400 - 435}{30} < Z < \dfrac{450 - 435}{30}\right)$ $= P(-1.1667 < X < 0.5)$ $= \Phi(0.5) - \Phi(-1.1667)$ $= 0.6915 - 0.1216$ $= 0.5699$ **[3]**	 M1 B1 A1	 For correct structure For use of difference column to obtain 0.8784, 0.8783, 0.1216 or 0.1217. Condone 0.8782 or 0.1218 FT "their 0.6915" – 0.1216 (or 0.1217)	
3	(ii)		P(all 5 between 400 and 450) $= 0.5699^5$ $= 0.0601$ **[2]**	 M1 A1	 FT Allow 0.060	

Question			Answer	Marks	Guidance	
3	**(iii)**		P($Y < 350$) = 0.2, P($Y > 390$) = 0.1 $$P(Z < \frac{350 - \mu}{\sigma}) = 0.2$$ $\Phi^{-1}(0.2) = -0.8416$ $$\frac{350 - \mu}{\sigma} = -0.8416$$ $$P(Z > \frac{390 - \mu}{\sigma}) = 0.1$$	M1	For equation as seen or equivalent with their –ive z value	If 'continuity corrections' applied allow M marks but do not award final A marks
			$\Phi^{-1}(0.9) = 1.282$	B1	For 1.282 or -0.8416	
			$$\frac{390 - \mu}{\sigma} = 1.282$$ $350 = \mu - 0.8416\,\sigma$ $390 = \mu + 1.282\,\sigma$ $2.1236\,\sigma = 40$	M1	For equation as seen or equivalent with their +ive z value	
			$\sigma = 18.84$	A1	Allow 18.8	Answers to max 2 d.p.
			$\mu = 350 + (0.8416 \times 18.84) = 365.85$	A1	Allow 365.86, 366, 365.9	
				[5]		
3	**(iv)**		$\Phi^{-1}(0.975) = 1.96$	B1	For using a suitable pair of z values e.g. ± 1.96	
			$a = 365.85 - (1.96 \times 18.84)$	M1	For either equation provided that a suitable pair of z-values is used. e.g. +2.326 and –1.751	
			$= 328.9$	A1	FT their μ and σ to 2 d.p. (A0 if 'continuity correction' used)	Accept any correct values of a and b.
			$b = 365.85 + (1.96 \times 18.84)$ $= 402.8$	A1	FT their μ and σ to 2 d.p. (A0 if 'continuity correction' used)	
				[4]		

Question	Answer	Marks	Guidance
4 **(a)**	H_0: no association between grade and hours worked H_1: some association between grade and hours worked;	B1	Hypotheses in context

	Less than 5hrs	At least 5hrs
A or B	17.05	13.95
C or lower	15.95	13.05

M1 — Any row/column correct
A1 — For expected values (to 2 dp)

	Less than 5hrs	At least 5hrs
A or B	0.5104	0.6238
C or lower	0.5456	0.6669

M1 — For valid attempt at $(O-E)^2/E$. Any row column correct.
A1 — For all correct

$X^2 = 2.347$ — B1

Refer to χ_1^2 — M1 — For 1 deg of freedom. No FT from here if wrong.

Critical value at 5% level = 3.841 — A1 — CAO for cv or p-value = 0.1255. SC1 for cv or p-value if 1 dof not seen.

Result is not significant.
There is insufficient evidence to suggest that there is any association between hours worked and grade. — E1 — For conclusion in context. NB if H_0 H_1 reversed, or 'correlation' mentioned, do not award first B1 or final E1

[9]

NB These M1A1 marks cannot be implied by a correct final value of X^2

Question	Answer	Marks	Guidance
4 **(b)**	$\bar{x} = 417.79$	B1	For $\bar{x}$
	$H_0: \mu = 420;$	B1	For use of 420 in hypotheses. Hypotheses in words must refer to population. Do not allow alternative symbols unless clearly defined as the population mean.
	$H_1: \mu \neq 420$	B1	For both correct
	Where μ denotes the mean volume of the cans of tomato purée (in the population)	B1	For definition of μ. Condone omission of "population" if correct notation μ is used, but if μ is defined as the **sample** mean then award **B0**.
	Test statistic $= \dfrac{417.79 - 420}{3.5 / \sqrt{10}} = \dfrac{-2.21}{1.107} = -1.997$	M1* A1	must include $\sqrt{10}$ FT their $\bar{x}$
	Lower 1% level 2 tailed critical value of $z = -2.576$	B1*	For -2.576 Must be -2.576 unless it is clear that absolute values are being used.
	$-1.997 > -2.576$	M1 dep*	For sensible comparison leading to a conclusion.
	So not significant. There is insufficient evidence to reject H0		
	There is insufficient evidence to suggest that the average volumes of the cans of tomato purée is not 420ml	A1	For conclusion in words in context provided that correct cv used. FT only candidate's test statistic.
		[9]	

ADDITIONAL NOTES REGARDING QUESTION 4 (b)

Critical Value Method
$420 - 2.576 \times 3.5 \div \sqrt{10}$ gets M1*B1*
$= 417.148\ldots$ gets A1
$417.79 > 417.148..$ gets M1dep* for sensible comparison
A1 still available for correct conclusion in words & context

Confidence Interval Method
CI centred on 417.79 + or – $2.5756 \times 3.5 \div \sqrt{10}$ gets M1* B1*
$= (414.93\ldots, 420.64..)$ gets A1
NOTE that the final M1dep* A1 available only if 2.576 used.
"Contains 420" gets M1dep*
A1 still available for correct conclusion in words & context

Probability Method
Finding P(sample mean < 417.79) = 0.0229 gets M1* A1 B1*
$0.0229 > $ **0.005*** gets M1dep* for a sensible comparison if a conclusion is made.
A1 available for a correct conclusion in words & context.
Condone P(sample mean > 417.79) = 0.9771 for M1* but only allow A1 B1* if sensible comparison made, at which point the final M1dep* and A1 are still available

ADDITIONAL NOTE REGARDING OVER-SPECIFICATION OF ANSWERS
Over-specification by providing final answers correct to 5 or more significant figures will be penalised. When this applies, candidates may lose no more than 2 marks per question and no more than 4 marks in total. The only exception to this rule is in Question 3 parts (iii) & (iv) – see guidance notes.

OCR (Oxford Cambridge and RSA Examinations)
1 Hills Road
Cambridge
CB1 2EU

OCR Customer Contact Centre

Education and Learning
Telephone: 01223 553998
Facsimile: 01223 552627
Email: general.qualifications@ocr.org.uk

www.ocr.org.uk

For staff training purposes and as part of our quality assurance
programme your call may be recorded or monitored

Oxford Cambridge and RSA Examinations
is a Company Limited by Guarantee
Registered in England
Registered Office; 1 Hills Road, Cambridge, CB1 2EU
Registered Company Number: 3484466
OCR is an exempt Charity

OCR (Oxford Cambridge and RSA Examinations)
Head office
Telephone: 01223 552552
Facsimile: 01223 552553

© OCR 2013