GCE

Further Mathematics B (MEI)

Y422/01: Statistics major

Advanced GCE

Mark Scheme for June 2019

OCR (Oxford Cambridge and RSA) is a leading UK awarding body, providing a wide range of qualifications to meet the needs of candidates of all ages and abilities. OCR qualifications include AS/A Levels, Diplomas, GCSEs, Cambridge Nationals, Cambridge Technicals, Functional Skills, Key Skills, Entry Level qualifications, NVQs and vocational qualifications in areas such as IT, business, languages, teaching/training, administration and secretarial skills.

It is also responsible for developing new specifications to meet national requirements and the needs of students and teachers. OCR is a not-for-profit organisation; any surplus made is invested back into the establishment to help towards the development of qualifications and support, which keep pace with the changing needs of today's society.

This mark scheme is published as an aid to teachers and students, to indicate the requirements of the examination. It shows the basis on which marks were awarded by examiners. It does not indicate the details of the discussions which took place at an examiners' meeting before marking commenced.

All examiners are instructed that alternative correct answers and unexpected approaches in candidates' scripts must be given marks that fairly reflect the relevant knowledge and skills demonstrated.

Mark schemes should be read in conjunction with the published question papers and the report on the examination.

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 a result or establishing a given result
dep*	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
AG	Answer given
awrt	Anything which rounds to
BC	By Calculator
DR	This indicates that the instruction **In this question you must show detailed reasoning** appears in the question.

Subject-specific Marking Instructions for A Level Further Mathematics B (MEI)

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, escalate the question to your Team Leader who will decide on a course of action with the Principal Examiner.
If you are in any doubt whatsoever 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, e.g. 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, e.g. 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, what is acceptable will be detailed in the mark scheme. If this is not the case, please escalate the question to your Team Leader who will decide on a course of action with the Principal Examiner.
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.

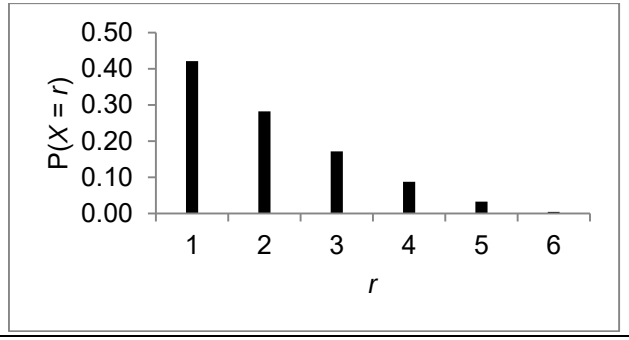
f Unless units are specifically requested, there is no penalty for wrong or missing units as long as the answer is numerically correct and expressed either in SI or in the units of the question. (e.g. lengths will be assumed to be in metres unless in a particular question all the lengths are in km, when this would be assumed to be the unspecified unit.) We are usually quite flexible about the accuracy to which the final answer is expressed; over-specification is usually only penalised where the scheme explicitly says so. When a value is given in the paper only accept an answer correct to at least as many significant figures as the given value. This rule should be applied to each case. When a value is not given in the paper accept any answer that agrees with the correct value to 2 s.f. Accept any value within a probability model that agrees with the correct value to 4 d.p.
Follow through should be used so that only one mark is lost for each distinct accuracy error, except for errors due to premature approximation which should be penalised only once in the examination. There is no penalty for using a wrong value for *g*. E marks will be lost except when results agree to the accuracy required in the question.

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 For a genuine misreading (of numbers or symbols) which is such that the object and the difficulty of the question remain unaltered, mark according to the scheme but following through from the candidate's data. A penalty is then applied; 1 mark is generally appropriate, though this may differ for some units. This is achieved by withholding one A mark in the question. Marks designated as cao may be awarded as long as there are no other errors. E marks are lost unless, by chance, the given results are established by equivalent working. 'Fresh starts' will not affect an earlier decision about a misread. Note that a miscopy of the candidate's own working is not a misread but an accuracy error.

i If a calculator is used, some answers may be obtained with little or no working visible. Allow full marks for correct answers (provided, of course, that there is nothing in the wording of the question specifying that analytical methods are required). Where an answer is wrong but there is some evidence of method, allow appropriate method marks. Wrong answers with no supporting method score zero. If in doubt, consult your Team Leader.

j If in any case the scheme operates with considerable unfairness consult your Team Leader.

Question			Answer	Marks	Guidance	
1	(a)		r | 1 | 2 | 3 | 4 | 5 | 6 P($X = r$) | $91k$ | $61k$ | $37k$ | $19k$ | $7k$ | k	**B1** **[1]**		
1	(b)		$91k + 61k + 37k + 19k + 7k + k = 1$ $216k = 1$ So $k = \frac{1}{216}$	**M1** **A1** **[2]**	For equation **AG**	Zero if only write $216k = 1$ so $k = \frac{1}{216}$
1	(c)			**B1** **B1** **[2]**	For heights by eye For axes (including scales) and labels Should be a probability scale rather 91, 61 etc so $91k$ is ok	
1	(d)		The distribution has (strong) positive skew	**B1** **[1]**	Allow 'The distribution is J-shaped' Do not allow 'Decreasing distribution'	
1	(e)		**DR** $E(X) = 1\times\frac{91}{216} + 2\times\frac{61}{216} + 3\times\frac{37}{216} + 4\times\frac{19}{216} + 5\times\frac{7}{216} + 6\times\frac{1}{216}$ $= \frac{49}{24} = 2.0417$ $E(X^2) = 1^2\times\frac{91}{216} + 2^2\times\frac{61}{216} + 3^2\times\frac{37}{216} + 4^2\times\frac{19}{216}$ $+ 5^2\times\frac{7}{216} + 6^2\times\frac{1}{216} = \frac{1183}{216} = 5.4769$ $\text{Var}(X) = 5.4769 - (2.0417)^2$ $= 1.308$ or $\frac{2261}{1728}$	**M1** **A1** **M1*** ***DM1** **A1** **[5]**	FT their probabilities for all M marks (provided they sum to 1) Allow fraction or decimal form Dep on valid attempt at $E(X^2)$ – at least four correct terms but could be ito k	

Question			Answer	Marks	Guidance	
2	(a)		Faults occur randomly, independently and at a uniform average rate	**E1** **E1** **[2]**	Allow constant average rate	Minus 1 mark for no context
2	(b)		P(≥ 2 faults) = 1 – 0.5249 = 0.4751	**M1** **A1** **[2]**	Or P(≥ 2 faults) = 1 – P(≤ 1 fault) BC	
2	(c)		Mean = 5 × 1.6 = 8 P(≤ 10 faults) = 0.8159	**B1** **B1** **[2]**	 BC	
2	(d)		Exactly 1 fault in 10 km So can use Poisson(3.2) P(1 fault) = 0.1304	**M1** **M1** **A1**	seen or implied BC	
			Alternative solution In 1 km length, P(0) = 0.7261, P(1) = 0.2324 For 10 1 km lengths, P(1) = $10 \times 0.7261^9 \times 0.2324$ = 0.1304	 **M1** **M1** **A1**	 BC (for both)	
				[3]		

3	(a)		Total weight : N(5×205, 5×11^2) P(Total ≥ 1000 g) = 0.8453	**M1** **A1** **[2]**	For distribution BC	
3	(b)		Peeled weight : N(0.65×205, $0.65^2 \times 11^2$) P(weight ≤ 150 g) = 0.9904	**B1** **M1** **A1** **[3]**	For N and mean For variance BC	(mean = 133.25) (variance = 51.1225)

Question			Answer	Marks	Guidance	
3	(c)		Weight of smoothie : N($2\times133.25 + 20\times22.5$, $2\times51.1225 + 20\times2.7^2$) N(716.5, 248.045) P(weight < 700 g) = 0.1474	**M1** **M1** **A1** **A1** **[4]**	Method for mean FT their part (b) Method for variance FT their part (b) For both correct BC	
4	(a)		Underlying distribution of caesium levels needs to be Normal	**E1** **[1]**	Context not required	
4	(b)		$2.369 < \mu < 2.836$	**B1** **[1]**	.	
4	(c)		$\text{SE} = \frac{0.2793}{\sqrt{8}}$ $= 0.09875$	**M1** **A1** **[2]**	 Allow 0.0987 or 0.0988	
4	(d)		t value = 2.365 t value × SE = 2.365 × 0.09875 = 0.2335	**B1** **B1** **[2]**	 **AG**	
4	(e)		By using a higher confidence level	**E1** **[1]**	Do NOT allow 'a stricter interval' Do not allow higher significance level. Do not allow higher confidence level if then contradicted by writing eg 90%	

Question			Answer	Marks	Guidance	
5	(a)		$B11 = \frac{144 \times 23}{500} = 6.6240$ $C10 = \frac{118 \times 139}{500} = 32.8040$ $C14 = \frac{(52 - 66.164)^2}{66.164} = 3.0321$	**B1** **B1** **M1** **A1** **[4]**	Allow these to be found by subtraction from row or column totals For $\frac{(O-E)^2}{E}$ used	
5	(b)		H_0: no association between smoking status and weight H_1: some association between smoking status and weight Degrees of freedom = 4 Critical value = 13.28 Test statistic = 14.716 14.716 > 13.28 so reject H_0 There is sufficient evidence to suggest that there is some association between smoking status and weight	**B1** **B1** **B1** **M1** **A1** **[6]**	For both Do NOT allow 'relationship' in place of association FT their test statistic provided that critical value is correct. Do NOT allow 'relationship' here If hypotheses the wrong way around MAX B0B1B1B1M0A0	Allow independent, not independent
5	(c)		For non-smokers the contribution of 3.0321 shows that rather fewer than expected are normal weight For light smokers the contribution of 3.8510 shows that more than expected are underweight For heavy smokers the contribution of 1.2129 shows that rather more than expected are of normal weight	**B1** **B1** **B1** **[3]**	Do NOT allow 'are more underweight than expected'	

Question			Answer	Marks	Guidance	
6	(a)	(i)	$b = \frac{S_{xy}}{S_{yy}} = \frac{116\,724 - (1131 \times 1227/12)}{126\,725 - (1227^2/12)} = 0.8537$ Correct regression line is x on y so equation is $x - \bar{x} = b(y - \bar{y})$ $\Rightarrow x - 94.25 = 0.8537(y - 102.25)$ $\Rightarrow x = 0.8537y + 6.962$	**M1** **A1** **B1** **DM1** **A1** **[5]**	For attempt at gradient (b) For 0.8537 cao May be implied by correct form of equation FT provided first M1 earned CAO Accept either form Allow M1M1 for y on x regression line with $b = 0.9098$	$Sxy = 1079.25$ $Syy = 1264.25$ $b = 0.853668\ldots$ Allow $b = \frac{4317}{5057}$ Allow constant between 6.9 and 7.0
6	(a)	(ii)	Prediction for 95 is 88 Prediction for 60 is 58	**B1FT** **B1FT** **[2]**	If answers given to more than 1 decimal place then MAX B1B0 as these are estimates	FT only if any reasonable x on y line
6	(a)	(iii)	Because the points do not lie very close to the line, the first prediction is only moderately reliable. The second prediction is rather less reliable because in addition it is extrapolation	**E1** **E1** **[2]**	Allow 1 mark for either not very close to line and so not very reliable or for second value is extrapolation so unreliable.	
6	(b)	(i)	The shape of the scatter diagram is very approximately elliptical, so bivariate Normality is possible	**B1** **B1** **[2]**	For identifying 'elliptical' shape For conclusion about 'bivariate Normal'	Do NOT allow 'data is bivariate Normal' but can get first mark
6	(b)	(ii)	PMCC = –0.5638	**B1** **B1** **[2]**	NB B1 for +0.5638 or for answer given to less than 3 dp	

Question			Answer	Marks	Guidance	
6	(b)	(iii)	H_0: $\rho = 0$, H_1: $\rho \neq 0$ (two-tailed test) where ρ is the population correlation coefficient between temperature and rainfall For $n = 10$, 5% critical value = 0.6319 Since $\lvert -0.5640 \rvert < 0.6319$ the result is not significant There is insufficient evidence to reject H_0 There is insufficient evidence at the 5% level to suggest that there is correlation between temperature and rainfall	**B1** **B1** **B1** **M1** **A1** **[5]**	For both hypotheses Allor r if defined For defining ρ For critical value For comparison leading to a conclusion NB M0 for $-0.5640 < 0.6319$ Do NOT allow M1 for incorrect conclusion FT for conclusion in context Provided critical value is correct Do NOT allow 'there is evidence to suggest that there is no correlation between temperature and rainfall'	Only penalise lack of context once. For hypotheses in words allow both marks if population and context mentioned, but zero if no mention of population
7	(a)		0.1442 ± 1.96 $\times \frac{0.2580}{\sqrt{40}}$ $= 0.1442 \pm 0.0800$ or (0.0642, 0.2242)	**B1** **M1** **M1** **A1** **[4]**	Allow 0.064 to 0.224	

Question			Answer	Marks	Guidance	
7	(b)		It seems that the coach's belief may be correct, as the confidence interval contains 0.2	**E1** **E1** **[2]**	FT their interval Allow E1E0 for whole interval is above zero so evidence to suggest coach may be correct	Condone 'The coaches' belief is correct'
7	(c)		By the CLT, for large samples the distribution of the sample mean is approximately Normal	**B1** **B1** **[2]**	For mention of central limit theorem For full statement (including CLT) and including sample mean	
7	(d)		For halving 0.12 $1.96 \times \frac{0.2580}{\sqrt{n}} \leq 0.06$ $n = 71.03$ so minimum sample size is 72	**M1** **M1** **A1** **[3]**	If 0.12 not halved allow M0M1A0	

Question			Answer	Marks	Guidance	
8	(a)		A Wilcoxon test should be carried out since a *t* test requires the population to be Normally distributed, but the Normal probability plot is not roughly straight and the *p*-value is low which both suggest that the data does not come from a Normal distribution	**B1** **E1** **E1** **[3]**	For conclusion of Wilcoxon For either For both and correct conclusion	
8	(b)		There is no indication that her sample is random The students from whom the sample is taken may not be representative of students in general EG Test scores may not be independent due to having the same teaching. Students at this school may be more or less able than at other schools The sample size is too small The test requires a symmetrical distribution which may not be the case (or median not equal to the mean) Accept possible improvements such as a paired sample test might be more suitable, if scores both with and without tea drinking could be obtained	**E1** **E1** **E1** **[3]**	1 mark for each correct statement Allow any valid features Do NOT allow 'not independent' with no context Needs to be more than 'these students may not have previously averaged 35' Do not allow 'It could be that many students who previously sat the test had drunk tea beforehand'	

<table>
<tr><th colspan="3">Question</th><th>Answer</th><th>Marks</th><th>Guidance</th><th></th></tr>
<tr><td>8</td><td>(c)</td><td></td><td>H_0: population median is 33.5
H_1: population median is greater than 33.5

<table>
<tr><th>Result</th><th>Res – 33.5</th><th>Abs value</th><th>Rank</th></tr>
<tr><td>26</td><td>–7.5</td><td>7.5</td><td>7</td></tr>
<tr><td>28</td><td>–5.5</td><td>5.5</td><td>6</td></tr>
<tr><td>29</td><td>–4.5</td><td>4.5</td><td>5</td></tr>
<tr><td>30</td><td>–3.5</td><td>3.5</td><td>4</td></tr>
<tr><td>31</td><td>–2.5</td><td>2.5</td><td>3</td></tr>
<tr><td>32</td><td>–1.5</td><td>1.5</td><td>2</td></tr>
<tr><td>34</td><td>0.5</td><td>0.5</td><td>1</td></tr>
<tr><td>42</td><td>8.5</td><td>8.5</td><td>8</td></tr>
<tr><td>49</td><td>15.5</td><td>15.5</td><td>9</td></tr>
<tr><td>54</td><td>20.5</td><td>20.5</td><td>10</td></tr>
<tr><td>55</td><td>21.5</td><td>21.5</td><td>11</td></tr>
<tr><td>56</td><td>22.5</td><td>22.5</td><td>12</td></tr>
<tr><td>61</td><td>27.5</td><td>27.5</td><td>13</td></tr>
</table>

$W_- = 7 + 6 + 5 + 4 + 3 + 2 = 27$
($W_+ = 1 + 8 + 9 + 10 + 11 + 12 + 13 = 64$)
Test statistic = $W_- = 27$
Critical value = 21
So do not reject H_0; there is insufficient evidence to suggest that the tea drinking improves arithmetic ability</td><td>B1
B1

M1

M1

A1
B1

A1
[7]</td><td>Population median used
Both correct
Can get B1 for H_0: median is 33.5 (no mention of population) but Max B1B0 if no mention of population
Zero for use of ‘average’

For attempt at ranking

Allow if they omit second column of table but otherwise all correct

Attempt to calculate either W_+ or W_-

For correct test statistic 27
For correct critical value 21

For conclusion in context; no FT from incorrect test statistic or critical value
FT for A1 (Not B1) from two tailed hypothesis and ‘correct’ CV of 17.

Zero marks for Normal distribution test</td><td></td></tr>
<tr><td>9</td><td>(a)</td><td></td><td>P(at least 3 minutes) = 0.4</td><td>B1
[1]</td><td></td><td></td></tr>
</table>

Question			Answer	Marks	Guidance	
9	(b)		P(at most 6 minutes) = 0.5 $+\int_5^6 \frac{1}{25}(10-x)\,dx$ = 0.5 + 0.18 = 0.68	 **M1** **M1** **A1** **[3]**	 For addition of 0.5 oe use of area of trapezium or triangle	 oe, using 1 – P(> 6)
9	(c)		Estimate of $P(T>18)=\frac{2}{20}=0.1$	**B1** **[1]**		
9	(d)		Use more rows in the simulation	**E1** **[1]**	Allow 'increase the sample size', 'use more weeks', etc	
9	(e)		E(T) = 12.5 $\text{Var}(T)=5\times\frac{25}{12}$ $=\frac{125}{12}$ (= 10.417)	**B1** **M1** **A1** **[3]**		
9	(f)		Total ~ $N\left(12.5, \frac{125}{12}\right)$ Estimate of P(T > 18) = 0.0442	**M1** **A1** **[2]**	oe use of Mean for 5 days ~ $N\left(2.5, \frac{5}{12}\right)$ BC	FT their 9(e) for M1 only
9	(g)		The sample size of 5 is small, so the CLT may not give an accurate estimate	**B1** **[1]**	Allow n is small, etc (but not 20 is small)	
9	(h)		Total ~ $N\left(200\times2.5, 200\times\frac{25}{12}\right)$ so $N\left(500, \frac{5000}{12}\right)$ gives P(Total > 510) = 0.3121	**B1** **M1** **A1** **[3]**	For Normal and correct mean (500) For variance (416.67) BC	oe, using the distribution of the mean time for a year

Question			Answer	Marks	Guidance	
10	(a)		$\int_0^a kx^m \, \mathrm{d}x = 1$	**M1**	For integral equated to 1	
			$\left[\frac{kx^{m+1}}{m+1}\right]_0^a = 1$	**M1**		
			$\frac{ka^{m+1}}{m+1} = 1 \Rightarrow k = \frac{m+1}{a^{m+1}}$	**A1** **[3]**	**AG**	
10	(b)		$\int_0^x \frac{m+1}{a^{m+1}} u^m \, \mathrm{d}u$	**B1**	Limits needed. Can get first two marks if in terms of k	
			$= \left[\frac{m+1}{a^{m+1}} \frac{u^{m+1}}{m+1}\right]_0^x$	**M1**	oe use of constant of integration	
			$= \frac{x^{m+1}}{a^{m+1}}$	**A1**	oe but not with k	
			$\mathrm{F}(x) = \begin{cases} 0 & x < 0 \\ \frac{x^{m+1}}{a^{m+1}} & 0 \le x \le a \\ 1 & x > a \end{cases}$	**A1** **[4]**	Fully simplified For fully correct answer with no incorrect working allow all 4 marks.	

Question			Answer	Marks	Guidance	
10	(c)	(i)	$\mathrm{P}\left(\frac{1}{4}a < X < \frac{1}{2}a\right) = \frac{\left(\frac{1}{2}a\right)^{m+1}}{a^{m+1}} - \frac{\left(\frac{1}{4}a\right)^{m+1}}{a^{m+1}}$	**M1**	For use of $\mathrm{F}\left(\frac{1}{2}a\right) - \mathrm{F}\left(\frac{1}{4}a\right)$, oe	Allow in terms of k
			$= \frac{1}{2^{m+1}} - \frac{1}{4^{m+1}}$	**M1**		
			$\Rightarrow \frac{1}{2p} - \frac{1}{4p^2} = \frac{1}{10}$	**M1**	Forming equation in p	
			$\Rightarrow 10p - 5 = 2p^2 \Rightarrow 2p^2 - 10p + 5 = 0$	**A1** **[4]**	**AG**	
10	(c)	(ii)	$p = 0.5635$ or $p = 4.4365$	**B1**	BC	Or $p = \frac{5 \pm \sqrt{15}}{2}$
			so $m = \frac{\log 0.5635}{\log 2}$ or $\frac{\log 4.4365}{\log 2}$	**M1**	For attempt to find m	
			giving $m = 2.149$ (and reject negative value -0.8275)	**A1** **[3]**		

OCR (Oxford Cambridge and RSA Examinations)
The Triangle Building
Shaftesbury Road
Cambridge
CB2 8EA

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