



GCE

Mathematics (MEI)

Advanced GCE

Unit 4767: Statistics 2

Mark Scheme for June 2013

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.

OCR will not enter into any discussion or correspondence in connection with this mark scheme.

© OCR 2013

1. Annotations and abbreviations

Annotation in scoris	Meaning
 and 	
	Benefit of doubt
	Follow through
	Ignore subsequent working
 ,	Method mark awarded 0, 1
 ,	Accuracy mark awarded 0, 1
 ,	Independent mark awarded 0, 1
	Special case
	Omission sign
	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

12. 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)	<p>EITHER:</p> $S_{xy} = \sum xy - \frac{1}{n} \sum x \sum y = 40.66 - \frac{1}{60} \times 43.62 \times 55.15$ $= 0.56595$ $S_{xx} = \sum x^2 - \frac{1}{n} (\sum x)^2 = 32.68 - \frac{1}{60} \times 43.62^2$ $= 0.96826$ $S_{yy} = \sum y^2 - \frac{1}{n} (\sum y)^2 = 51.44 - \frac{1}{60} \times 55.15^2$ $= 0.74796$ $r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{0.56595}{\sqrt{0.96826 \times 0.74796}} = 0.665$ <p>OR:</p> $\text{cov}(x,y) = \frac{\sum xy}{n} - \bar{xy} = 40.66/60 - (43.62/60 \times 55.15/60)$ $= 0.0094325$ $\text{rmsd}(x) = \sqrt{\frac{S_{xx}}{n}} = \sqrt{(0.96826/60)} = \sqrt{0.016137\dots} = 0.1270$ $\text{rmsd}(y) = \sqrt{\frac{S_{yy}}{n}} = \sqrt{(0.74796/60)} = \sqrt{0.012466} = 0.1117$ $r = \frac{\text{cov}(x,y)}{\text{rmsd}(x)\text{rmsd}(y)} = \frac{0.0094325}{0.1270 \times 0.1117} = 0.665$	M1* M1* A1 M1 dep* A1 [5] M1* M1* A1 M1 dep* A1 [5]	For method for S_{xy} For method for at least one of S_{xx} or S_{yy} For at least one of S_{xy} , S_{xx} or S_{yy} (to 2 sf) Note Allow 0.57322 for S_{xy} and 0.76634 for S_{yy} from rounding mean of y to 0.919. For structure of r For answer rounding to 0.66 or 0.67 For method for cov (x,y) For method for at least one msd or rmsd For at least one of cov (x,y) , msd or rmsd correct (to 2 sf) For structure of r For answer rounding to 0.66 or 0.67 Methods mixed – max M0M1A1M0A0	

Question		Answer	Marks	Guidance
1	(ii)	<p>$H_0: \rho = 0$ $H_1: \rho > 0$ (one-tailed test)</p> <p>where ρ is the population correlation coefficient</p> <p>For $n = 60$, 5% critical value = 0.2144</p> <p>Since $0.665 > 0.2144$, the result is significant.</p> <p>Thus we have sufficient evidence to reject H_0</p> <p>There is sufficient evidence at the 5% level to suggest that there is positive correlation between FEV1 before and after the two-week course.</p>	B1 B1 B1 M1 A1 E1 [6]	<p>For H_0, H_1 in symbols. Hypotheses in words must refer to population. Do not allow alternative symbols unless clearly defined as the population correlation coefficient.</p> <p>For defining ρ. Condone omission of “population” if correct notation ρ is used, but if ρ is defined as the sample correlation coefficient then award B0. Allow “ρ is the pmcc”.</p> <p>For critical value</p> <p>For sensible comparison leading to a conclusion provided that $r < 1$. The comparison can be in the form of a diagram as long as it is clear and unambiguous. Sensible comparison: e.g. $0.665 > 0.2144$ is ‘sensible’ whereas $0.665 > -0.2144$ is ‘not sensible’. Reversed inequality sign e.g. $0.665 < 0.2144$ etc. gets max M1 A0.</p> <p>For reject H_0 o.e. FT their r and critical value from 5% 1-tail column.</p> <p>For correct, non-assertive conclusion in context (allow ‘x and y’ for context). E0 if H_0 and H_1 not stated, reversed or mention a value other than zero for ρ in H_0.</p>

Question		Answer	Marks	Guidance
1	(iii)	<p>The underlying population must have a bivariate Normal distribution.</p> <p>Yes, since the scatter diagram appears to have a roughly elliptical shape.</p>	<p>B1</p> <p>E1 [2]</p>	<p>Condone “bivariate Normal distribution”, “underlying bivariate Normal distribution”, but do not allow “the data have a bivariate Normal distribution”</p> <p>Condone ‘oval’ or suitable diagram</p>
1	(iv)	The significance level is the probability of rejecting the null hypothesis when in fact it is true.	<p>E1* E1dep* [2]</p>	<p>For “probability of rejecting H_0” or “probability of a significant result”.</p> <p>For “when H_0 is true”</p>
1	(v)	$\sum x = 43.62 + 0.45 = 44.07$ $\sum y = 55.15 - 0.45 = 54.70$ $\sum xy = 40.66$ $\sum x^2 = 32.68 + 1 - 0.55^2 = 33.3775$ $\sum y^2 = 51.44 - 1 + 0.55^2 = 50.7425$	<p>B1</p> <p>B1</p> <p>B1 [3]</p>	<p>For $\sum x$ or $\sum y$ or $\sum xy$</p> <p>For $\sum x^2$ or $\sum y^2$ (to 2 dp)</p> <p>For all correct (ignore n)</p>
2	(i)	$P(\text{At least one has red hair}) = 1 - 0.97^{10}$ $= 0.263$	<p>M1 A1 [2]</p>	<p>M1 for $1 - 0.97^{10}$</p> <p>Allow 0.26</p>
2	(ii)	<p>(Because X is binomially distributed), n is large and p is small.</p> <p>Mean = 1.8</p>	<p>E1 E1 B1 [3]</p>	<p>Allow “sample is large” for n is large</p> <p>Allow “$np < 10$” or “mean \approx variance” for “p is small”</p> <p>Do not allow “the probability is small”</p>

Question		Answer	Marks	Guidance
2	(iii)	(A) $P(X=2) = e^{-1.8} \frac{1.8^2}{2!} = 0.2678$ OR $0.7306 - 0.4628 = 0.2678$	M1 A1 [2]	For calculation for $P(X=2)$ FT their mean. Allow answer to 3sf.
2	(iii)	(B) $P(X > 2) = 1 - P(X \leq 2) = 1 - 0.7306 = 0.2694$	M1 A1 [2]	$1 - P(X \leq 2)$ used. e.g. $1 - P(X \leq 2) = 1 - 0.4628$ gets M0 CAO
2	(iv)	The mean ($np = 1.8$) is too small It is not appropriate to use a Normal approximation	E1* E1dep* [2]	For “mean is too small” or “mean < 10” For “not appropriate”. Do not allow “ p is too small”.
2	(v)	Binomial(5000, 0.03)	B1* B1dep* [2]	For binomial, or B(,) For parameters
2	(vi)	Mean $5000 \times 0.03 = 150$ Variance $= 5000 \times 0.03 \times 0.97 = 145.5$ Using Normal approx. to the binomial, $X \sim N(150, 145.5)$ $P(X \geq 160) = P\left(Z \geq \frac{159.5 - 150}{\sqrt{145.5}}\right)$ $= P(Z > 0.7876) = 1 - \Phi(0.7876) = 1 - 0.7846$ $= 0.215$ (to 3 sig.fig.)	B1 B1 B1 M1 A1 [5]	For mean (soi) For variance (soi) For continuity corr. For probability using correct tail and structure (condone omission of/incorrect c.c.) CAO, (Do not FT wrong or omitted CC) Allow 0.2155. Do not allow 0.216

Question		Answer	Marks	Guidance
3	(i)	$\begin{aligned} P(Y = 76) &= P\left(\frac{75.5 - 76}{12} \leq Z \leq \frac{76.5 - 76}{12}\right) \\ &= P(-0.04166... < Z < 0.04166...) \\ &= \Phi(0.04166...) - (1 - \Phi(0.04166...)) \\ &= 2 \times \Phi(0.04166...) - 1 \\ &= 2 \times 0.5167 - 1 \\ &= 0.0334 \end{aligned}$	B1 M1 M1 A1 [4]	For one correct continuity correction used For standardizing For correctly structured probability calculation. CAO inc use of diff tables. Allow 0.0330 – 0.0340 www.
3	(ii)	$\begin{aligned} P(Y \geq 80) &= P\left(Z \geq \frac{79.5 - 76}{12}\right) \\ &= P(Z > 0.2917) = 1 - \Phi(0.2917) \\ &= 1 - 0.6148 = 0.3852 = 0.385 \text{ to 3 sig fig} \end{aligned}$	B1 M1 A1 [3]	For correct cc used For correct structure CAO do not allow 0.386
3	(iii)	$3 \times 0.3852 \times 0.6148^2 = 0.4368$	M1 A1 [2]	$3 \times \text{their } p \times (1 - \text{their } p)^2$ FT their p . Allow 2sf if working seen.

Question		Answer	Marks	Guidance	
3	(iv)	<p>EITHER: $P(\text{Score} \geq k) = 0.1$</p> $\Phi^{-1}(0.9) = 1.282$ $\frac{k - 76}{12} = 1.282$ $k = 76 + (1.282 \times 12) = 91.38 \text{ or}$ $k = 76 + 0.5 + (1.282 \times 12) = 91.88$ $91.38 > 90.5 \text{ or } 91.88 > 91$ <p>so lowest reported mark = 92</p> <p>OR Trial and improvement method</p> $P(\text{Mark} \geq 91) = P(\text{Score} \geq 90.5) = 0.1135$ $P(\text{Mark} \geq 92) = P(\text{Score} \geq 91.5) = 0.0982$ $P(\text{Mark} \geq 91) > 10\% \text{ and } P(\text{Mark} \geq 92) < 10\%$ <p>so lowest reported mark = 92</p>	B1 M1 A1 M1 A1 M1 A1 A1 [5]	For 1.282 Allow $k - 0.5$ used for k . Positive z used. For 91.38 or 91.88 Relevant comparison (e.g. diagram) M1 for attempt to find $P(\text{Mark} \geq \text{integer})$ A1 for 0.1135 A1 for 0.0982 M1 for comparisons	www
3	(v)	$P(Y \leq 50) = 0.2$ $P(Z \leq \frac{50.5 - \mu}{12}) = 0.2$ $\frac{50.5 - \mu}{12} = \Phi^{-1}(0.2) = -0.8416$ $\mu = 50.5 + (12 \times 0.8416) = 60.6$	B1 B1 M1 A1 [4]	For 50.5 used For -0.8416 . Condone -0.842 Condone 0.8416 if numerator reversed. For structure. CAO	

Question		Answer					Marks	Guidance																														
4	(i)	H_0 : no association between sex and artist preferred H_1 : some association between sex and artist preferred					B1	For both hypotheses in context																														
		<table border="1"> <thead> <tr> <th>EXPECTED</th><th>Monet</th><th>Renoir</th><th>Degas</th><th>Cézanne</th></tr> </thead> <tbody> <tr> <td>Male</td><td>12.13</td><td>28</td><td>13.07</td><td>16.8</td></tr> <tr> <td>Female</td><td>13.87</td><td>32</td><td>14.93</td><td>19.2</td></tr> </tbody> </table> <table border="1"> <thead> <tr> <th>CONTRIB'N</th><th>Monet</th><th>Renoir</th><th>Degas</th><th>Cézanne</th></tr> </thead> <tbody> <tr> <td>Male</td><td>1.4081</td><td>0.3214</td><td>1.8626</td><td>0.2881</td></tr> <tr> <td>Female</td><td>1.2321</td><td>0.2813</td><td>1.6298</td><td>0.2521</td></tr> </tbody> </table>		EXPECTED	Monet	Renoir	Degas	Cézanne	Male	12.13	28	13.07	16.8	Female	13.87	32	14.93	19.2	CONTRIB'N	Monet	Renoir	Degas	Cézanne	Male	1.4081	0.3214	1.8626	0.2881	Female	1.2321	0.2813	1.6298	0.2521	M1 A2	For expected values (to 2 dp where appropriate) (allow A1 for at least one row or column correct)			
EXPECTED	Monet	Renoir	Degas	Cézanne																																		
Male	12.13	28	13.07	16.8																																		
Female	13.87	32	14.93	19.2																																		
CONTRIB'N	Monet	Renoir	Degas	Cézanne																																		
Male	1.4081	0.3214	1.8626	0.2881																																		
Female	1.2321	0.2813	1.6298	0.2521																																		
		$X^2 = 7.28$ Refer to χ^2 Critical value at 10% level = 6.251 Result is significant There is evidence to suggest that there is some association between sex and artist preferred		M1 A2	For valid attempt at $(O-E)^2/E$ For all correct (to 2 dp) and presented in a table or clear list. (Allow A1 for at least one row or column correct)		NB: These three marks cannot be implied by a correct final value of X^2 www																															
		NB if H_0 H_1 reversed, or 'correlation' mentioned, do not award first B1 or final E1		B1 B1	Allow 7.27 for 3 deg of f																																	
				B1	CAO for cv No FT from here if wrong or omitted, unless <i>p</i> -value used instead		B1 for <i>p</i> -value = 0.0636																															
				B1	FT their X^2																																	
				E1	For correct (FT their X^2), non-assertive conclusion, in context.																																	
				[12]																																		

Question		Answer	Marks	Guidance	
4	(ii)	<p>Monet: More females and fewer males than expected prefer Monet, as indicated by large contribution(s) (of 1.4081 and 1.2321).</p> <p>Renoir: Preferences are much as expected, as indicated by small contributions.</p> <p>Degas: Fewer females and more males than expected prefer Degas, as indicated by large contribution(s) (of 1.8626 and 1.6298).</p> <p>Cézanne: Preferences are much as expected, as indicated by small contributions.</p>	E1* E1dep* E1 E1* depE1* E1 [6]	FT their table of contributions	NB MAX 3/6 for answers not referring to contributions (explicitly or implicitly). SC1 Renoir and Cézanne have correct comments for both but without referring to contributions

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

