



**GCE**  
**Mathematics (MEI)**

Unit **4769**: Statistics 4

Advanced GCE

**Mark Scheme for June 2015**

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 2015

## Annotations and abbreviations

Annotation in scores	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, with 3 significant figures often being the norm. Small variations in the degree of accuracy to which an answer is given (e.g. 2 or 4 significant figures where 3 is expected) should not normally be penalised, while answers which are grossly over- or under-specified should normally result in the loss of a mark. 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 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.

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

Question		Answer	Marks	Guidance
1	(i)	The integral is $-\exp(-x^2/a)$ , limits 0 and infinity. (Which evaluates to $0 - (-1) = 1$ .)	B1 [1]	Answer given
	(ii)	$\begin{aligned} E(X^2) &= \int_0^\infty \frac{2}{a} x^3 \exp\left(-\frac{x^2}{a}\right) dx \text{ or } \int_0^\infty x^2 \frac{2}{a} x \exp\left(-\frac{x^2}{a}\right) dx \\ &= \left[ x^2 \left( -\exp\left(-\frac{x^2}{a}\right) \right) \right]_0^\infty + \int_0^\infty 2x \exp\left(-\frac{x^2}{a}\right) dx \\ &= 0 + a = a \end{aligned}$ $\begin{aligned} E(X^4) &= \int_0^\infty \frac{2}{a} x^5 \exp\left(-\frac{x^2}{a}\right) dx \text{ or } \int_0^\infty x^4 \frac{2}{a} x \exp\left(-\frac{x^2}{a}\right) dx \\ &= \left[ x^4 \left( -\exp\left(-\frac{x^2}{a}\right) \right) \right]_0^\infty + \int_0^\infty 4x^3 \exp\left(-\frac{x^2}{a}\right) dx \\ &= 0 + 2a \times a = 2a^2 \end{aligned}$	M1 M1A1 A1 M1 A1 A1 [7]	M1 parts Cao Answer given so working must be convincing M1 for recognisable attempt to obtain the likelihood. A1 for $\prod X_i$ . A1 for the exp fn.
	(iii)	<p>Likelihood is <math>\left(\frac{2}{a}\right)^n \prod X_i \exp\left(-\frac{1}{a} \sum X_i^2\right)</math> (log likelihood as given) Differentiate log likelihood to obtain <math>-\frac{n}{a} + \frac{1}{a^2} \sum X_i^2</math></p> <p>Set this to zero to obtain <math>\text{MLE}(a) = \frac{1}{n} \sum X_i^2</math></p>	M1A1A1 M1 A1 M1A1 [7]	Justification of a maximum not required

Question		Answer	Marks	Guidance
(iv)	$E(\text{MLE}(a)) = \frac{1}{n} \sum E(X_i^2) = \frac{1}{n} n a = a$ <p>That is, the MLE is unbiased</p> $\begin{aligned} \text{Var}(\text{MLE}(a)) &= \frac{1}{n^2} \sum \text{Var}(X_i^2) = \frac{1}{n^2} \sum (E(X_i^4) - (E(X_i^2))^2) \\ &= \frac{1}{n^2} n (2a^2 - a^2) = \frac{a^2}{n} \end{aligned}$	B1	E1	Seen or very clearly implied
(v)	<p>Maximum likelihood estimate (also unbiased) of <math>a</math> is 1.471 with estimated standard error (<math>a / \sqrt{n}</math>) 0.1471.  <math>95\% \text{ CI is } 1.471 \pm 1.960 \times 0.1471 = 1.471 \pm 0.288</math></p>	M1M1	A1	Explanation not required if calculations are correct Accept 2 for 1.960
		[5]	B1	
		B1	M1A1	
		[4]	[24]	

Question		Answer	Marks	Guidance
2	(i)	$M_Y(\theta) = E(e^{\theta Y}) = \int_0^\infty e^{\theta y} \frac{1}{\sqrt{2\pi y}} e^{-\frac{1}{2}y} dy$ $= \int_0^\infty \frac{1}{\sqrt{2\pi y}} e^{-\frac{1}{2}y(1-2\theta)} dy$ <p>Substitute <math>u = y(1-2\theta)</math>, <math>du = dy(1-2\theta)</math> to obtain</p> $\int_0^\infty \frac{\sqrt{1-2\theta}}{\sqrt{2\pi u}} e^{-\frac{1}{2}u} \frac{1}{1-2\theta} du$ $= (1-2\theta)^{-\frac{1}{2}} \int_0^\infty \frac{1}{\sqrt{2\pi u}} e^{-\frac{1}{2}u} du$ $= (1-2\theta)^{-\frac{1}{2}}$	M1, A1 A1 M1 A1 B1 [6]	Answer given
	(ii)	<p>Either expand the mgf as a power series:</p> $(1-2\theta)^{-\frac{1}{2}} = 1 + \theta + 3 \frac{\theta^2}{2!} + \dots$ <p><math>E(Y) = 1</math> (coefficient of <math>\theta</math>)  <math>E(Y^2) = 3</math> (coefficient of <math>\theta^2/2!</math>)  Hence <math>\text{Var}(Y) = 2</math></p> <p>-----</p> <p>Or differentiate the mgf:</p> <p>1<sup>st</sup> derivative simplifies to <math>(1-2\theta)^{-\frac{3}{2}}</math>  Putting <math>\theta = 0</math> gives <math>E(Y) = 1</math>  2<sup>nd</sup> derivative simplifies to <math>3(1-2\theta)^{-\frac{5}{2}}</math>  Putting <math>\theta = 0</math> gives <math>E(Y^2) = 3</math>, hence <math>\text{Var}(Y) = 2</math></p>	M1A1 B1 B1 B1 [5]	
	(iii)	<p>For independent rvs <math>X</math> and <math>Y</math>,</p> $M_{X+Y}(\theta) = M_X(\theta) M_Y(\theta)$	B1	

Question		Answer	Marks	Guidance
(iv)	<p>Hence <math>M_U(\theta) = (1 - 2\theta)^{-\frac{n}{2}}</math></p> <p><math>E(U) = n</math></p> <p><math>\text{Var}(U) = 2n</math></p> $\begin{aligned} M_W(\theta) &= E\left(\exp\left(\theta\left(\frac{U-n}{\sqrt{2n}}\right)\right)\right) \\ &= \exp\left(-\frac{\theta n}{\sqrt{2n}}\right) E\left(\exp\left(\frac{\theta}{\sqrt{2n}}U\right)\right) \\ &= \exp\left(-\frac{\theta n}{\sqrt{2n}}\right) M_U\left(\frac{\theta}{\sqrt{2n}}\right) \\ &= \exp\left(-\frac{\theta n}{\sqrt{2n}}\right) \left(1 - \frac{2\theta}{\sqrt{2n}}\right)^{-\frac{n}{2}} \end{aligned}$ <p>Hence <math>\ln(M_W(\theta))</math> as given</p> <p>Expanding <math>\ln(M_W(\theta))</math> gives</p> $-\sqrt{\frac{n}{2}}\theta + \frac{n}{2}\left(\sqrt{\frac{2}{n}}\theta + \frac{1}{2}\left(\sqrt{\frac{2}{n}}\theta\right)^2 + \text{terms of order } n^{-\frac{3}{2}}\right)$ <p>Convincing simplification to <math>\frac{1}{2}\theta^2 + \text{terms of order } n^{-\frac{1}{2}}</math></p> <p>Hence tends to <math>\frac{1}{2}\theta^2</math>.</p> <p>The mgf therefore tends to <math>\exp(\frac{1}{2}\theta^2)</math>.</p> <p>That is, the distribution of <math>W</math> tends to the standard Normal.</p>	<p>B1</p> <p>B1</p> <p>B1</p> <p>[4]</p> <p>B1</p> <p>B1</p> <p>B1</p> <p>B1</p> <p>M1A1</p> <p>B1</p> <p>B1</p> <p>[9]</p> <p>[24]</p>	<p>Or by use of general linear transformation result</p> <p>Award a maximum of 3 marks from here on if no account is taken of terms beyond <math>\theta^2</math>. (answer given)</p> <p>Uniqueness of mgfs may be implied</p>	

Question		Answer	Marks	Guidance
3	(a) (i)	<p>Sample means: standard 107.2766, GM 113.6192</p> <p>Sample variances: standard 138.2243, GM 121.9372</p> <p><math>H_0: \mu_1 = \mu_2</math> (the means of the underlying distributions)</p> <p><math>H_0: \mu_1 \neq \mu_2</math></p> <p>Test statistic: <math display="block">\frac{113.6192 - 107.2766}{\sqrt{\frac{138.2243}{30} + \frac{121.9372}{26}}} = 2.08(0103)</math></p> <p>Compare with <math>z</math> distribution</p> <p>Critical value, 2 tails, 5%, 1.960</p> <p>So 2.08(0103) is just in the critical region</p> <p>Hence reject <math>H_0</math> and conclude that there is a difference in mean weight between standard and GM tomatoes.</p> <p>Assumption: the tomatoes may be regarded, in some sense, as random samples of their respective varieties.</p>	B1 B1 B1 M1A1 B1 B1 B1B1 B1 [10]	Must be clear that the hypotheses refer to underlying means
	(ii)	<p>Additional assumptions: underlying Normality of the tomatoes' weights common variance in the underlying distributions</p> <p>Given the sample sizes it seems safe to use the Normal distribution.</p> <p>The Normal test is better in that it makes fewer assumptions.</p>	B1 B1 E1 E1 [4]	
	(b) (i)	<p>The tomatoes should be chosen at random.</p> <p>The panel should not know which tomatoes are GM / standard.</p> <p>Making fine judgements on the appearances of tomatoes is unlikely to be reliable.</p>	B1 B1 B1 [3]	Accept other sensible comments

Question		Answer	Marks	Guidance
	(ii)	<p>Rank sums: Standard 134, GM 76</p> <p><math>H_0</math>: GM and std tomatoes have, on the whole, the same appearance</p> <p><math>H_1</math>: GM tomatoes have, on the whole, a better appearance than std tomatoes</p> <p>(lower rank sum for GM indicates that the evidence is in the correct tail)</p> <p>Wilcoxon rank sum test</p> <p>Critical value for <math>m = n = 10</math>, 1 tail, 1% level is 74</p> <p>The observed value of 76 is not in the critical region</p> <p>so accept <math>H_0</math>. That is, conclude that there is insufficient evidence to suppose that, on the whole, GM tomatoes have a better appearance than standard tomatoes.</p>	M1A1 B1 B1 B1 B1 B1 [7] [24]	<p>Hope for, but don't expect, a formulation in terms of a shift in location parameter for underlying distributions of appearances</p> <p>May be implied</p> <p>Critical value 19</p>

Question		Answer	Marks	Guidance
4	(a)	(i) $Y_{ij}$ denotes the $j$ th observation (or measurement) in the $i$ th group $\mu$ is the underlying mean for whole population $\alpha_i$ denotes the population mean difference for the $i$ th group $\varepsilon_{ij}$ denotes the random error in the $ij$ th observation (or measurement)	B1 B1 B1 B1 [4]	Accept just $Y_{ij}$ denotes the observations (measurements) Withhold these marks if there is no reference to populations In marking section (a), evidence of understanding is more important than precise wording.
		(ii) $\varepsilon_{ij} \sim \text{ind N}(0, \sigma^2)$ , where $\sigma^2$ is constant across groups. The variance ratio test (F test) requires underlying Normality (and common variance).	B1B1 B1 B1 [4]	B1 ind N, B1 zero mean May be implied by later comment Common variance may be implied by earlier statement
		(iii) $H_0$ : all the $\alpha_i$ are zero $H_1$ : not all the $\alpha_i$ are zero If $H_0$ is accepted then we proceed on the basis that the group means are all equal. I.e. there are no 'treatment effects'. If $H_0$ is rejected the we proceed on the basis that the group means are not all equal. I.e. there are some 'treatment effects'.	B1 B1 E1 E1 [4]	Accept clear wording in terms of all group effects being zero, or all groups having same underlying distribution H <sub>1</sub> : all the $\alpha_i$ are non-zero scores zero
	(b)	(i) $k = 4$ (number of routes is 4) $N = 19$ (number of journeys is 19)	B1 B1 [2]	

Question		Answer					Marks	Guidance
	(ii)	Source of var'n	SS	df	MS	F		
		Between groups	333.77	<b>3</b>	<b>111.2567</b>	<b>3.98(113)</b>	M1	Evidence of understanding how to complete the table
		Within groups	<b>419.19</b>	15	<b>27.9460</b>		A4	for the figures shown in bold -1 each error
		Total	752.96	18				
		$H_0$ : all routes have the same underlying mean time $H_1$ : routes do not all have the same underlying mean time The observed variance ratio is compared with $F_{3, 15}$ The 5% critical value is 3.29 (tables) or 3.287382 (calculator) Observed value is in critical region, so reject $H_0$ . That is, proceed on the assumption that the routes do differ in underlying mean time.					B1	
							M1	Hypotheses may be implied by later comments
							A1	May be implied by use of a correct value
							B1	Or equivalent wording
							[10]	
							[24]	

**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](mailto:general.qualifications@ocr.org.uk)

[www.ocr.org.uk](http://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 2015

