OCR
RECOGNISING ACHIEVEMENT

GCE

Statistics (MEI)

Advanced Subsidiary GCE

Unit **G242:** Statistics 2 (Z2)

Mark Scheme for June 2012

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, OCR Nationals, 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.

Any enquiries about publications should be addressed to:

OCR Publications
PO Box 5050
Annesley
NOTTINGHAM
NG15 0DL

Telephone: 0870 770 6622
Facsimile: 01223 552610
E-mail: publications@ocr.org.uk

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 awarde$ 0(1
B0$ B1	Independent mark awarded 0, 1
SC	Cpecial 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 Mr 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.

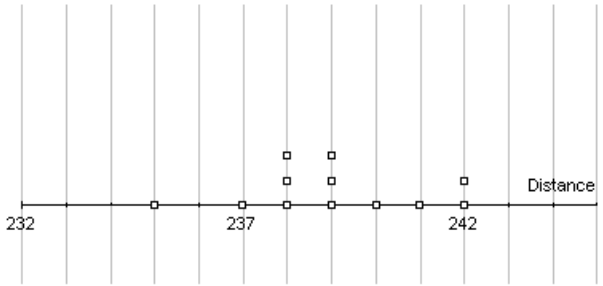
NB Follow these maths-specific instructions rather than those in the assessor handbook.

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)		Independence and constant probability	B1 **[1]**	
1	(ii)		Poisson approximation is appropriate as p is small and n is large The parameter of the Poisson distribution is 0.4	B1 B1 **[2]**	Allow '$np < 10$' or 'mean ≈ variance' for 'p small' This mark may be recovered in part (iii)
1	(iii)		Using $X \sim$ Poisson(0.4) $P(X \geq 3) = 1 - P(X \leq 2)$ $= 1 - 0.9921$ $= 0.0079$	 M1 A1 **[2]**	 Correct structure with any λ. FT their λ from part (ii). 0.0076 from B(50, 0.008) gets 0/2
1	(iv)		Using $X \sim$ Poisson(1.5) $P(X = 1) = \frac{e^{-1.5} \times 1.5}{1!} = 0.33469\ldots$ $\ldots \times 200 = 66.9390\ldots = 66.94$ (2d.p.)	 M1 M1 **[2]**	 M1 for calculation as seen Or for 0.5578 – 0.2231 from tables
1	(v)		H_0: The Poisson model is suitable H_1: The Poisson model is not suitable Number of degrees of freedom = 3 (5 – 1 – 1) At 5% significance level, critical value is 7.815 9.279 > 7.815 Result is significant. Evidence suggests that this Poisson model is not a good fit to these data.	 B1 B1 B1 B1 B1 E1 **[6]**	 Condone 'data is a good fit to the Poisson model' 3 degrees of freedom **seen** cao Sensible comparison **seen**. FT only c.v. = 9.488 [note that c.v. = 9.488 leads to 'not significant'] Conclusion in words must be non-assertive and consistent with their result and hypotheses.
2	(i)		By symmetry, μ lies halfway between 72 and 98	B1 **[1]**	Allow (72 + 98) ÷ 2 = 85. Allow μ is the mid-point of 72 and 98. Allow 98 = 85 + 13 & 72 = 85 – 13. Allow demonstration that the values standardise to give ± 2.907
2	(ii)		$P(M < 80) = P(Z < \frac{80-85}{\sqrt{20}})$ $= 1 - \Phi(1.118\ldots)$ $= 1 - 0.8682 = 0.132$ (3sf)	M1 M1 A1 **[3]**	M1 for standardising. Allow numerator reversed. M0 if continuity correction used. Correct structure FT their σ cao

Question			Answer	Marks	Guidance
2	(iii)		A population mean is the mean of all possible values and is a constant A sample mean is the mean of a sample and is variable.	E2,1,0 **[2]**	For E2 need recognition that population mean is constant and sample mean is variable. E1 for explaining that the sample mean represents the mean of only part of the population.
2	(iv)		Standard error = $\sqrt{20} \div \sqrt{4}$ $= \sqrt{5} = 2.24$ (3sf) cao	M1 A1 **[2]**	M1 for $\sqrt{20} \div \sqrt{4}$ even if seen as part of probability calculation. Allow M1 for $20 \div \sqrt{4}$ if use of $\sigma = 20$ has already been penalised in part (ii)
2	(v)		$\frac{90-85}{\sqrt{5}} = 2.236\ldots$ $1 - \Phi(2.236\ldots) = 0.0127$ (3sf)	M1 M1 A1 **[3]**	M1 for standardising. FT their standard error M1 for correct structure cao
2	(vi)		As the sample was taken from a Normal distribution, it is not necessary to use the CLT in part (v)	B1 E1 dep **[2]**	
3	(i)		Assume that the underlying population is distributed symmetrically and the sample is random. H_0: population median = 258 H_1: population median < 258 Actual differences –5 –3 –7 –12 –8 13 6 –2 –10 9 Associated ranks 3 2 5 9 6 10 4 1 8 7 $T^- = 3 + 2 + 5 + 9 + 6 + 1 + 8 = 34$ $T^+ = 10 + 4 + 7 = 21$ $\therefore T = 21$ From $n = 10$ tables – at the 5% level of significance in a one-tailed Wilcoxon single sample test, the critical value of T is 10 $21 > 10$ $\therefore$ the result is not significant The evidence does not suggest that there has been a reduction in the average lap time.	B1 B1 B1 B1 B1 M1*A1 B1 B1 M1* A1 M1dep* A1 **[13]**	Condone 'the values are random' B1 for using 258 in hypotheses B1 for both correct and including 'population' Condone opposite signs M1 for ranking absolute values of their differences. A1 FT provided differences are used and the sum of the ranks is 55. B1 for $T^+ = 21$or $T^- = 34$. FT only if the sum of the ranks is 55 B1 FT for test statistic cao FT their test statistic only if both previous M1 marks earned A1 for non-assertive conclusion in context

Question			Answer	Marks	Guidance
3	(ii)		The Wilcoxon test does not require the underlying population to be Normally distributed whereas the *t* test does.	E1 **[1]**	Allow 'lap times do not need to be Normally distributed' Condone 'data do not need to be Normally distributed'
3	(iii)		Sensible comment with justification. e.g. Lap times are unlikely to be distributed Normally so a *t* test is probably not appropriate. e.g. consideration of dot plot	E2,1,0 **[2]**	E2 for 'lap times are unlikely to be distributed Normally so a *t* test is (probably) not appropriate'. i.e. not overly assertive. E1 for 'lap times are not Normally distributed so a *t* test is not appropriate'. Allow sensible alternatives
4	(i)		$\Sigma x \div n = 2868 \div 12$ (= 239 A.G.) $\sqrt{\dfrac{685498 - \dfrac{2868^2}{12}}{11}}$ = 2.045 (4s.f.)	B1 M1 A1 **[3]**	Full method shown or 2868 ÷ 12 seen Allow M1A0 for σ^2
4	(ii)		Sensible comment, with justification.  eg These (continuous) data appear to be symmetrically distributed and so it is possible that they have an underlying Normal distribution.	E1 E1 **[2]**	E1 for suitable comment regarding similarity between mean and median or for suitable comment relating to central tendency or comment about continuous variable. E1 for suitable comment regarding symmetry
4	(iii)		$239 \pm 2.201 \times \dfrac{2.0449..}{\sqrt{12}}$ (237.70, 240.30)	B1 B1 M1 A1A1 **[5]**	centred on 239 for 2.201 structure Allow A1A0 for (237.700…, 240.299…)

<table>
<tr><th colspan="3">Question</th><th>Answer</th><th>Marks</th><th>Guidance</th></tr>
<tr><td>4</td><td>(iv)</td><td></td><td>The interval in (iii) contains 240.
This supports the view that this type of battery is suitable.
The mean could be at least 240.</td><td>E1*
E1dep*
E1
[3]</td><td>Can be implied. Allow ‘manufacturer’s requirement’ for ‘240’.
FT their CI
Allow other sensible comments.</td></tr>
<tr><td>4</td><td>(v)</td><td></td><td>Variance is unknown.
Sample is small.</td><td>E1
E1
[2]</td><td>Allow ‘n is small’ or ‘n < 30’</td></tr>
<tr><td>5</td><td>(i)</td><td></td><td>H_0: No association between pesticide use and change in bee population.
H_1: There is an association between pesticide use and change in bee population.
Expected frequencies
<table>
<tr><th colspan="2" rowspan="2">Expected frequencies</th><th colspan="3">Pesticide use</th></tr>
<tr><th>High</th><th>Medium</th><th>Low</th></tr>
<tr><td rowspan="3">Change in bee population</td><td>Minimal change</td><td>9.425</td><td>14.625</td><td>14.95</td></tr>
<tr><td>Decrease of 10% to 20%</td><td>10.15</td><td>15.75</td><td>16.1</td></tr>
<tr><td>Over 20% decrease</td><td>9.425</td><td>14.625</td><td>14.95</td></tr>
</table>
Contributions to X^2
<table>
<tr><th colspan="2" rowspan="2">Contributions to the test statistic</th><th colspan="3">Pesticide use</th></tr>
<tr><th>High</th><th>Medium</th><th>Low</th></tr>
<tr><td rowspan="3">Change in bee population</td><td>Minimal change</td><td>4.380</td><td>0.129</td><td>1.706</td></tr>
<tr><td>Decrease of 10% to 20%</td><td>0.071</td><td>0.004</td><td>0.075</td></tr>
<tr><td>Over 20% decrease</td><td>3.298</td><td>0.181</td><td>1.044</td></tr>
</table>
$X^2 = 10.887$
4 degrees of freedom
Critical value for 5% significance level is 9.488
As 10.887 > 9.488 the result is significant
There is evidence to suggest an association between pesticide use and change in bee population</td><td>B1

B2,1,0

M1
A1

A1
B1
B1
M1, A1
B1

[11]</td><td>Hypotheses in context

B1 if at least two expected frequencies correct
B2 if all correct

Attempt at (obs – exp)^2 ÷ exp
All correct

cao AWRT 10.89

M1 for sensible comparison, A0 if wrong cv used or if first M mark not awarded
B1 for non-assertive conclusion in context. If hypotheses reversed, do not award first B1 or final A1 B1</td></tr>
</table>

Question			Answer	Marks	Guidance
5	(ii)		This cell provides the strongest evidence of association between the classification factors. There were significantly fewer observed values than would be expected if there were no association.	E1 E1 **[2]**	Do not allow 'positive association'. Allow 'this supports the alternative hypothesis' Allow other sensible comments related to the hypothesis test
5	(iii)		Medium pesticide use shows the least association. The column for medium use has the lowest total contribution to the test statistic.	E1 E1 **[2]**	

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