OCR
RECOGNISING ACHIEVEMENT

Monday 10 June 2013 – Morning

AS GCE MEI STATISTICS

G243/01 Statistics 3 (Z3)

QUESTION PAPER

G21601061 3

Candidates answer on the Printed Answer Book.

OCR supplied materials:
- Printed Answer Book G243/01
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- The Question Paper will be found in the centre of the Printed Answer Book.
- Write your name, centre number and candidate number in the spaces provided on the Printed Answer Book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the Printed Answer Book.** Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

INFORMATION FOR CANDIDATES

This information is the same on the Printed Answer Book and the Question Paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the Question Paper.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
- The total number of marks for this paper is **72**.
- This Printed Answer Book consists of **12** pages. The Question Paper consists of **4** pages. Any blank pages are indicated.

INSTRUCTION TO EXAMS OFFICER / INVIGILATOR

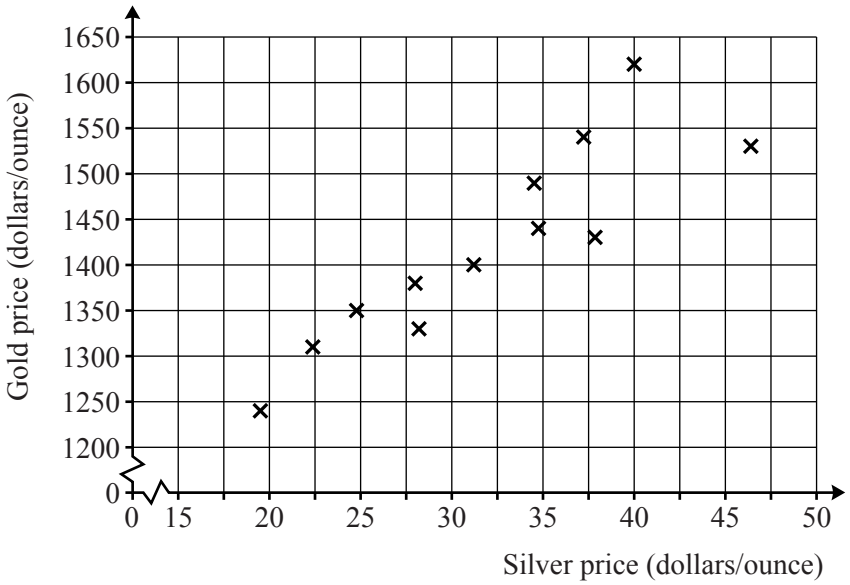
- Do not send this Question Paper for marking; it should be retained in the centre or recycled. Please contact OCR Copyright should you wish to re-use this document.

Section A (45 marks)

1 **(i)** A test is to be carried out to examine the relationship between two variables. Explain briefly why, provided that the conditions for its use are satisfied, it is preferable to carry out a test based on the product moment correlation coefficient, rather than a test based on Spearman's rank correlation coefficient. **[2]**

A student is investigating the prices of gold and silver over a period of one year. He randomly chooses 12 days over the course of the year and records the prices of both metals, in dollars per ounce, on those days. The prices are given below, together with a scatter diagram to illustrate the data.

Silver price	19.5	22.3	28.2	24.7	28.0	31.2	37.9	34.7	34.5	46.3	37.2	40.0
Gold price	1240	1310	1330	1350	1380	1400	1430	1440	1490	1530	1540	1620



(ii) State a condition which is required for the use of a test based on the product moment correlation coefficient. Explain why, in the light of the scatter diagram, it may not be appropriate to carry out such a test. **[2]**

(iii) Calculate the value of Spearman's rank correlation coefficient. **[5]**

(iv) Using your answer to part **(iii)** carry out a test, at the 5% level of significance, to determine whether it is reasonable to assume that prices of gold and silver are positively associated. **[6]**

2 A scientist from a company which manufactures baby milk is investigating whether babies fed with two different brands of baby milk from birth gain the same amount of weight, on average, in their first year of life. He randomly selects 100 babies who have been fed Brand A and another 100 babies who have been fed Brand B. The weight gains in kilograms have the following means and variances.

Brand A:	Sample mean 5.834	Sample variance 1.272
Brand B:	Sample mean 5.920	Sample variance 1.318

(i) Explain why, even though the distributions of the underlying populations are unknown, a test based on the Normal distribution is appropriate. **[2]**

(ii) Carry out a test, at the 5% level of significance, to examine whether there appears to be a difference in average weight gain between the two brands. **[11]**

(iii) Suppose instead that samples each of only 10 babies had been selected, and that the distributions of the underlying populations were unknown. Name another test which could have been used. State the null hypothesis which would have been tested. **[2]**

3 A university psychologist is investigating whether the time taken to complete a sudoku puzzle is affected by consumption of a small quantity of alcohol. She believes that consumption of alcohol will increase the median time taken. She randomly chooses 12 students from her classes who regularly do sudoku puzzles. Each student is given a particular puzzle to do; the time, in minutes, taken to complete it is recorded. Each student is then given a single unit of alcohol and required to complete a second puzzle of similar standard; the time taken to complete it is again recorded. The times are as follows.

Student	A	B	C	D	E	F	G	H	I	J	K	L
Time – no alcohol	6.2	13.7	15.5	11.0	11.6	9.6	14.6	10.2	19.8	12.9	5.9	14.7
Time – with alcohol	7.1	13.3	13.8	13.2	13.0	11.2	13.4	13.9	18.0	16.4	4.8	16.0

(i) Carry out a test, at the 5% level of significance, to examine whether the psychologist's belief concerning the median time taken appears to be correct. **[11]**

(ii) Explain why the psychologist has used a paired design. **[2]**

(iii) Explain briefly one improvement which could be made to the design of this experiment. **[2]**

Section B (27 marks)

4 **(i)** State two requirements for systematic sampling to be a sensible method for selecting a sample for a survey. **[2]**

(ii) Describe how to choose a systematic sample of 8 potato plants from a single row of 80 potato plants. **[2]**

A researcher is running a trial of two new varieties of potato, P and Q. A field contains 20 rows of each variety, with 80 plants in each row; the varieties are planted in alternate rows. The researcher intends to dig up 8 plants of each variety and measure the weight of potatoes produced by each plant. He considers the following methods of choosing a sample for each variety.

A Choose plants near the field entrance to avoid disturbing the rest of the crop.
B Systematically select 8 plants from one row of each variety in the field.
C Number all of the plants and then randomly select 8 of each variety.

(iii) Discuss advantages and disadvantages of these methods. **[6]**

(iv) Name the types of sampling described in methods A and C. **[2]**

The researcher wishes to compare the weights in kilograms of potatoes in the two samples to test whether the population mean for variety P is equal to the population mean for variety Q. He selects a random sample of 8 potatoes of each variety.

(v) State the assumptions which are required for a *t* test to examine whether the means appear to be equal. Given that these assumptions are valid, use the data below to carry out the test at the 5% significance level.

Variety P:	Sample mean 1.824	Sample variance 0.124
Variety Q:	Sample mean 2.191	Sample variance 0.230

[15]

OCR
RECOGNISING ACHIEVEMENT

Monday 10 June 2013 – Morning

AS GCE MEI STATISTICS

G243/01 Statistics 3 (Z3)

PRINTED ANSWER BOOK

G233580613

Candidates answer on this Printed Answer Book.

OCR supplied materials:
- Question Paper G243/01 (inserted)
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

G 2 4 3 0 1

Candidate forename		Candidate surname	

Centre number						Candidate number				

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- The Question Paper will be found in the centre of the Printed Answer Book.
- Write your name, centre number and candidate number in the spaces provided on the Printed Answer Book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the Printed Answer Book**. Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

INFORMATION FOR CANDIDATES

This information is the same on the Printed Answer Book and the Question Paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the Question Paper.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
- The total number of marks for this paper is **72**.
- The Printed Answer Book consists of **12** pages. The Question Paper consists of **4** pages. Any blank pages are indicated.

Section A (45 marks)

1 (i)	
1 (ii)	
1 (iii)	
	(answer space continued on next page)

1 (iii)	(continued)
1 (iv)	

2 (i)

2 (ii)

(answer space continued on next page)

2 (ii)	**(continued)**
2 (iii)	

3 (i)

(answer space continued on next page)

3 (i)	**(continued)**
3 (ii)	
3 (iii)	

Section B (27 marks)

4 (i)	
4 (ii)	

4 (iii)	Method A advantage:
	Method A disadvantage:
	Method B advantage:
	Method B disadvantage:
	Method C advantage:
	Method C disadvantage:

4 (iv)

4 (v)

(answer space continued on next page)

4 (v)	**(continued)**

PLEASE DO NOT WRITE ON THIS PAGE

OCR
RECOGNISING ACHIEVEMENT

GCE

Statistics (MEI)

Advanced Subsidiary GCE

Unit **G243:** Statistics 3 (Z3)

Mark Scheme for June 2013

Oxford Cambridge and RSA Examinations

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.

Annotations

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.

<table>
<tr><th colspan="3">Question</th><th>Answer</th><th>Marks</th><th>Guidance</th><th></th></tr>
<tr><td>1</td><td>(i)</td><td></td><td>Spearman's uses ranked data, rather than the raw data.
Thus information is lost.</td><td>E1

E1
[2]</td><td>For ranked. (May be implied by second E1)

For 'information is lost' oe</td><td>Spearman is less powerful gets E2

Do not allow 'not as accurate'</td></tr>
<tr><td>1</td><td>(ii)</td><td></td><td>The population should have a bivariate Normal distribution.
In this case the points do not appear to lie in an elliptical pattern which suggests that the population may not have a bivariate Normal distribution.</td><td>E1

E1

[2]</td><td></td><td>Do not allow 'Normality of both populations'</td></tr>
<tr><td>1</td><td>(iii)</td><td></td><td>
<table>
<tr><th>Silver</th><th>Gold</th><th>R silver</th><th>R gold</th><th>d</th><th>d^2</th></tr>
<tr><td>19.5</td><td>1240</td><td>1</td><td>1</td><td>0</td><td>0</td></tr>
<tr><td>22.3</td><td>1310</td><td>2</td><td>2</td><td>0</td><td>0</td></tr>
<tr><td>28.2</td><td>1330</td><td>5</td><td>3</td><td>2</td><td>4</td></tr>
<tr><td>24.7</td><td>1350</td><td>3</td><td>4</td><td>-1</td><td>1</td></tr>
<tr><td>28.0</td><td>1380</td><td>4</td><td>5</td><td>-1</td><td>1</td></tr>
<tr><td>31.2</td><td>1400</td><td>6</td><td>6</td><td>0</td><td>0</td></tr>
<tr><td>37.9</td><td>1430</td><td>10</td><td>7</td><td>3</td><td>9</td></tr>
<tr><td>34.7</td><td>1440</td><td>8</td><td>8</td><td>0</td><td>0</td></tr>
<tr><td>34.5</td><td>1490</td><td>7</td><td>9</td><td>-2</td><td>4</td></tr>
<tr><td>46.3</td><td>1530</td><td>12</td><td>10</td><td>2</td><td>4</td></tr>
<tr><td>37.2</td><td>1540</td><td>9</td><td>11</td><td>-2</td><td>4</td></tr>
<tr><td>40.0</td><td>1620</td><td>11</td><td>12</td><td>-1</td><td>1</td></tr>
</table>
$\Sigma d^2 = 28$
$r_s = 1 - \frac{6 \times 28}{12 \times 143} = 1 - 0.098$
= 0.902</td><td>M1

M1

A1

M1

A1

[5]</td><td>For good attempt at ranking (allow all ranks reversed)

For d^2

For method for r_s

FT their ranks provided $|r_s| < 1$
NB No ranking scores zero</td><td></td></tr>
</table>

Question			Answer	Marks	Guidance	
1	(iv)		H_0: no association between prices of gold and silver H_1: positive association between prices of gold and silver	B1 B1	B0B0 for hypotheses not in context	B0 B0 for H_0 H_1 ito ρ
				B1	For mention of population, either in hypotheses or in concluding statement	Allow -0.5035
			One tail test critical value at 5% level is 0.5035	B1		
			0.902 > 0.5035	M1	For comparison with c.v., leading to correct conclusion, provided $\|r_s\| < 1$	Dep on sensible CV Allow $r_s > 0.5035$
			There is sufficient evidence to reject H_0. Thus there is enough evidence to suggest positive association between prices of gold and silver in the population.	A1	For conclusion in words in context – FT their r_s and sensible cv Allow 'Significant' Condone 'Accept H_1'	Do not allow if also say 'not significant'
				[6]		

Question			Answer	Marks	Guidance	
2	(i)		The samples are large, so by the central limit theorem the underlying distribution of the sample means will be approximately Normal.	E2,1,0 **[2]**	Do not allow 'The <u>data</u> will be approximately normally distributed'	
2	(ii)		H_0: $\mu_A = \mu_B$ H_1: $\mu_A \neq \mu_B$ Where μ_A, μ_B denote the population mean weight gains with brands A and B respectively	B1 B1 B1	Condone absence of "population" if correct notation "μ" has been used, but do NOT accept $\overline{X}$ and $\overline{Y}$ or similar unless explicitly stated to be population means. Accept hypothesis explained in words, provided 'population' appears.	
			2–sample test based on N(0,1)	E1	soi	
			Test statistic is $\frac{5.920-5.834}{\sqrt{\frac{1.318}{100}+\frac{1.272}{100}}}=\frac{0.086}{0.161}=0.534$	M1 M1 A1	Numerator Denominator CAO (0.53 to 0.54) or (-0.54 to -0.53)	M0M0 for *t* statistic, even if numerator correct FT from here if all M marks earned
			2–tailed 5% point of N(0,1) is 1.96 0.534 < 1.96 Not significant There is insufficient evidence to suggest that the mean weight gain differs in the population.	B1 M1 A1 E1 **[11]**	No further marks if CV is wrong or -0.534 > -1.96 Allow 'Accept H_0'	
2	(iii)		Wilcoxon rank sum test H_0: The medians of the two populations are the same	B1 B1 **[2]**		Condone unpaired Wilcoxon test

Question			Answer	Marks	Guidance	
3	(i)		Wilcoxon signed rank test: H_0: Median population difference is zero H_1: Median population difference is positive Differences are 0.9 -0.4 -1.7 2.2 1.4 1.6 -1.2 3.7 -1.8 3.5 -1.1 1.3 Ranks of |d| are 2 1 8 10 6 7 4 12 9 11 3 5 Test statistic is 1+3+4+8+9 = 25 Refer to paired Wilcoxon table with n=12 Lower 5% 1-tailed value is 17 Not significant Insufficient evidence to suggest that the median time to do a puzzle after consuming one unit of alcohol is greater than the median time to do a puzzle without alcohol	B1 B1 M1 M1 A1 M1 A1 M1 A1 E1 E1 **[11]**	B1B0 if population not mentioned No marks if differences not used CAO FT if ranks wrong CAO FT if previous M1 earned No further marks if CV is wrong Allow 'Accept H_0'	Do not allow 'difference in medians is zero'
3	(ii)		The pairing will eliminate any differences in individual subjects' aptitude in the sudoku puzzle … …and so will compare the before and after times.	E2,1,0 **[2]**		
3	(iii)		EG: The two puzzles could be done on different days, in case having done one puzzle already, the students might perform worse on the second due to fatigue, or better due to having practised one already.	B1 E1 **[2]**	Sensible improvement to design Valid explanation	Eg 'more repetition' or 'sample from a larger population' Eg 'Have a control group, who have a 'placebo' rather than alcohol'

Question			Answer	Marks	Guidance	
4	(i)		There is a list of the population. There are no cycles or patterns in the list.	E1 E1 **[2]**		
4	(ii)		Randomly select one of the first 10 plants … … then select every tenth plant	E1 E1 **[2]**		Do not allow 'randomly select a potato plant' unless it is made clear that when you get to the end of the list you go back to the beginning
4	(iii)		A: Advantage – easy to do A: Disadvantage – conditions near the entrance may be different from elsewhere in the field B: Advantage – relative easy to carry out and only disturbs one row for each variety B: Disadvantage – conditions in these two rows may be different from elsewhere in the field C: Advantage – probability based so can carry out statistical inference C: Disadvantage – difficult to carry out	E6 **[6]**	E 3,2,1,0 for advantages. Allow any sensible advantages E 3,2,1,0 for disadvantages. Allow any sensible disadvantages	Other suitable answers include: Advantage for A: 'will save time' Advantage for C 'Every plant has equal chance of being selected' or 'Sample more likely to be representative than with the other methods' Do not allow: Advantage for A; 'Rest of crop not disturbed' Do not allow: Advantage for B; 'gives a representative sample from one row' Do not allow: Advantage for C 'sample <u>will</u> be more representative' or 'non-biased' or 'fair test' Do not allow: Disadvantage for A or B 'not a random sample' or 'may be biased'
4	(iv)		A: Opportunity sampling C: Simple random sampling	B1 B1 **[2]**	Accept quota sampling for A	Allow 'convenience sampling' Allow just 'random sampling'

Question			Answer	Marks	Guidance	
4	(v)		Assumptions required are: Normality of both populations, equal population variances	 B1 B1	Do not allow 'The <u>data</u> is normally distributed'	
			H_0: $\mu_Q = \mu_P$ H_1: $\mu_Q \neq \mu_P$ Where μ_P, μ_Q denote the population mean yields for varieties P and Q	B1 B1 B1	Condone absence of "population" if correct notation "μ" has been used, but do NOT accept $\overline{X}$ and $\overline{Y}$ or similar unless explicitly stated to be population means. Accept hypothesis explained in words, provided 'population' appears.	
			Pooled $s^2 = \frac{(7\times 0.230)+(7\times 0.124)}{14}$	M1	For attempt at pooling	
			$=0.177$	A1		
			Test statistic $= \frac{2.191-1.824}{\sqrt{0.177}\times\sqrt{\frac{1}{8}+\frac{1}{8}}} = \frac{0.367}{0.2104}$	M1 M1 M1	For numerator for $\sqrt{0.177}$ or √their pooled variance for $\sqrt{\frac{1}{8}+\frac{1}{8}}$	
			$=1.74$	A1	CAO Allow -1.74	FT from here if all M marks earned
			Refer to t_{14} 2–tail 5% point is 2.145	M1 A1	No further marks if CV is wrong	
			Not significant There is insufficient evidence to suggest that the population mean weights for the two varieties are different.	E1 E1		
				[15]		

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

PART OF THE CAMBRIDGE ASSESSMENT GROUP

LLOYD'S REGISTER QUALITY ASSURANCE · ISO9001 · UKAS MANAGEMENT SYSTEMS 001

OCR
RECOGNISING ACHIEVEMENT

GCE

Statistics (MEI)

Advanced Subsidiary GCE **AS H132**

OCR Report to Centres

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 report on the examination provides information on the performance of candidates which it is hoped will be useful to teachers in their preparation of candidates for future examinations. It is intended to be constructive and informative and to promote better understanding of the specification content, of the operation of the scheme of assessment and of the application of assessment criteria.

Reports should be read in conjunction with the published question papers and mark schemes for the examination.

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

G243 Statistics 3

General Comments

There was no evidence of candidates being unable to complete the paper in the allocated time. However, although some candidates were well prepared for the paper, a disappointingly large number gained relatively few marks. In general, candidates supported their numerical answers with appropriate explanations and working.

It is pleasing to report that fewer candidates than last year lost marks because they gave their hypotheses in words and failed to mention 'population' (other than in Question 1). Indeed, most candidates in tests for the mean did use the parameter μ as mentioned in last year's report and, of these, most then went on to define this as the population mean. As last year, relatively few candidates lost marks because their answers were too assertive. Perhaps surprisingly, Question 1 − on correlation − was found to be more difficult than the remaining questions, with many candidates scoring fairly well on the hypothesis tests in Questions 2, 3 and 4.

Comments on Individual Questions

1)(i) This question was not well answered, with many candidates simply stating that a test based on the product moment correlation coefficient is more accurate than one based on Spearman's rank correlation coefficient. This type of answer did not gain any credit.

1)(ii) Most candidates gained a mark for stating that the points did not appear to lie in an elliptical pattern, but rather fewer knew the requirement for bivariate Normality.

1)(iii) Under half of the candidature gained any credit here. A very large number of candidates failed to rank the data, or ranked from 1 to 24 rather than ranking each variable from 1 to 12. Most of those who did rank correctly gained full marks.

1)(iv) Most candidates stated the hypotheses correctly in terms of association rather than incorrectly mentioning correlation and only a few tried to use symbols. Many then went on to complete the test correctly although a few got the critical value wrong. However, very few candidates gained the mark for 'population'.

2)(i) Most candidates gained a mark for saying that the samples are large, but rather fewer mentioned the central limit theorem for the second mark. Many candidates discussed the variances – such comments were ignored, whether correct or incorrect.

2)(ii) This was generally fairly well answered, with candidates often gaining all or most of the marks available. A few failed to define the parameter μ, thus losing a mark. However a number of candidates squared the variances, presumably thinking that they were standard deviations. Some instead tried to find a pooled estimate for sample variance and thus failed to gain any credit other than possibly for the hypotheses.

2)(iii) About half of the candidates answered this correctly, but many candidates suggested a *t* test and a few thought that a signed rank test was appropriate.

3)(i) Although a good number of essentially correct responses were seen, very few candidates stated the hypotheses in terms of the population **difference**. Many candidates also made errors in the signs of the differences or in the ranking. However, provided that their essential method was correct, many such candidates gained method marks and marks for the critical value and the conclusion.

3)(ii) Candidates found this question difficult, with many simply saying something such as 'the psychologist wanted to compare with and without alcohol'. The best responses mentioned subjects' ability or aptitude.

3)(iii) Many candidates suggested using a larger sample and some gave a good explanation of why this would be an improvement, thus gaining one or two marks. A wide variety of incorrect responses was also seen.

4)(i) There were very few correct responses.

4)(ii) Most candidates gained a mark for stating that every tenth potato plant should be selected. It is disappointing to report that rather fewer gained the second mark for mentioning randomly selecting one of the first 10 plants as the point at which to start. Some suggested randomly choosing one of the 80 plants as the starting point, but this only gained a mark if it was clear that when the end of the list was reached one went back and started again from the beginning.

4)(iii) Most candidates gained some credit, but only a few scored all 6 marks. Those who thought about the ease or otherwise of carrying out the methods tended to score well, with 3 of the 6 marks available for this. A reasonable number of candidates mentioned the soil conditions or shading in the field being a possible disadvantage for methods A and B.

4)(iv) Many correct responses were seen, and most candidates gained at least one of the two marks. A few candidates did not distinguish their two answers and some gave three answers, so scored zero.

4)(v) A small number of fully correct responses were seen. Most candidates correctly gave their hypotheses in terms of μ. However, the majority of candidates did not know the necessary assumptions. Many candidates also made errors in the variance, often squaring the given variances before calculating a pooled estimate. Others forgot to square root their pooled variance in calculating the test statistic. A disappointing number of candidates gave a wrong critical value, some thinking there were 7 degrees of freedom and others using the Normal tables.

For a description of how UMS marks are calculated see:
www.ocr.org.uk/learners/ums_results.html

Unit level raw mark and UMS grade boundaries June 2013 series
AS GCE / Advanced GCE / AS GCE Double Award / Advanced GCE Double Award

GCE Mathematics (MEI)

			Max Mark	a	b	c	d	e	u
4751/01	(C1) MEI Introduction to Advanced Mathematics	Raw	72	62	56	51	46	41	0
		UMS	100	80	70	60	50	40	0
4752/01	(C2) MEI Concepts for Advanced Mathematics	Raw	72	54	48	43	38	33	0
		UMS	100	80	70	60	50	40	0
4753/01	(C3) MEI Methods for Advanced Mathematics with Coursework: Written Paper	Raw	72	58	52	46	40	33	0
4753/02	(C3) MEI Methods for Advanced Mathematics with Coursework: Coursework	Raw	18	15	13	11	9	8	0
4753/82	(C3) MEI Methods for Advanced Mathematics with Coursework: Carried Forward Coursework Mark	Raw	18	15	13	11	9	8	0
4753	(C3) MEI Methods for Advanced Mathematics with Coursework	UMS	100	80	70	60	50	40	0
4754/01	(C4) MEI Applications of Advanced Mathematics	Raw	90	66	59	53	47	41	0
		UMS	100	80	70	60	50	40	0
4755/01	(FP1) MEI Further Concepts for Advanced Mathematics	Raw	72	63	57	51	45	40	0
		UMS	100	80	70	60	50	40	0
4756/01	(FP2) MEI Further Methods for Advanced Mathematics	Raw	72	61	54	48	42	36	0
		UMS	100	80	70	60	50	40	0
4757/01	(FP3) MEI Further Applications of Advanced Mathematics	Raw	72	60	52	44	36	28	0
		UMS	100	80	70	60	50	40	0
4758/01	(DE) MEI Differential Equations with Coursework: Written Paper	Raw	72	62	56	51	46	40	0
4758/02	(DE) MEI Differential Equations with Coursework: Coursework	Raw	18	15	13	11	9	8	0
4758/82	(DE) MEI Differential Equations with Coursework: Carried Forward Coursework Mark	Raw	18	15	13	11	9	8	0
4758	(DE) MEI Differential Equations with Coursework	UMS	100	80	70	60	50	40	0
4761/01	(M1) MEI Mechanics 1	Raw	72	57	49	41	33	25	0
		UMS	100	80	70	60	50	40	0
4762/01	(M2) MEI Mechanics 2	Raw	72	50	43	36	29	22	0
		UMS	100	80	70	60	50	40	0
4763/01	(M3) MEI Mechanics 3	Raw	72	64	56	48	41	34	0
		UMS	100	80	70	60	50	40	0
4764/01	(M4) MEI Mechanics 4	Raw	72	56	49	42	35	29	0
		UMS	100	80	70	60	50	40	0
4766/01	(S1) MEI Statistics 1	Raw	72	55	48	41	35	29	0
		UMS	100	80	70	60	50	40	0
4767/01	(S2) MEI Statistics 2	Raw	72	58	52	46	41	36	0
		UMS	100	80	70	60	50	40	0
4768/01	(S3) MEI Statistics 3	Raw	72	61	55	49	44	39	0
		UMS	100	80	70	60	50	40	0
4769/01	(S4) MEI Statistics 4	Raw	72	56	49	42	35	28	0
		UMS	100	80	70	60	50	40	0
4771/01	(D1) MEI Decision Mathematics 1	Raw	72	58	52	46	40	35	0
		UMS	100	80	70	60	50	40	0
4772/01	(D2) MEI Decision Mathematics 2	Raw	72	58	52	46	41	36	0
		UMS	100	80	70	60	50	40	0
4773/01	(DC) MEI Decision Mathematics Computation	Raw	72	46	40	34	29	24	0
		UMS	100	80	70	60	50	40	0
4776/01	(NM) MEI Numerical Methods with Coursework: Written Paper	Raw	72	56	50	44	38	31	0
4776/02	(NM) MEI Numerical Methods with Coursework: Coursework	Raw	18	14	12	10	8	7	0
4776/82	(NM) MEI Numerical Methods with Coursework: Carried Forward Coursework Mark	Raw	18	14	12	10	8	7	0
4776	(NM) MEI Numerical Methods with Coursework	UMS	100	80	70	60	50	40	0
4777/01	(NC) MEI Numerical Computation	Raw	72	55	47	39	32	25	0
		UMS	100	80	70	60	50	40	0
4798/01	(FPT) Further Pure Mathematics with Technology	Raw	72	57	49	41	33	26	0
		UMS	100	80	70	60	50	40	0

GCE Statistics (MEI)

			Max Mark	a	b	c	d	e	u
G241/01	(Z1) Statistics 1	Raw	72	55	48	41	35	29	0
		UMS	100	80	70	60	50	40	0
G242/01	(Z2) Statistics 2	Raw	72	55	48	41	34	27	0
		UMS	100	80	70	60	50	40	0
G243/01	(Z3) Statistics 3	Raw	72	56	48	41	34	27	0
		UMS	100	80	70	60	50	40	0