

**ADVANCED SUBSIDIARY GCE**  
**MEI STATISTICS**  
Statistics 3 (Z3)

**G243**

Candidates answer on the Answer Booklet

**OCR Supplied Materials:**

- 8 page Answer Booklet
- Graph paper
- MEI Examination Formulae and Tables (MF2)

**Other Materials Required:**

- Scientific or graphical calculator

**Tuesday 15 June 2010**  
**Morning**

**Duration:** 1 hour 30 minutes



**INSTRUCTIONS TO CANDIDATES**

- Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
- Use black ink. Pencil may be used for graphs and diagrams only.
- Read each question carefully and make sure that 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 graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

**INFORMATION FOR CANDIDATES**

- The number of marks is given in brackets [ ] at the end of each question or part question.
- 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 document consists of **4** pages. Any blank pages are indicated.

## Section A (45 marks)

- 1 A student measures the pulse rates of a random sample of 10 runners both before and after they have run round an athletics track. The pulse rates are as follows.

Runner	A	B	C	D	E	F	G	H	I	J
Before	83	75	73	80	88	61	70	77	76	62
After	148	100	98	171	152	94	127	132	170	106

- (i) Draw a scatter diagram to illustrate these data. [3]
- (ii) Calculate the value of the product moment correlation coefficient. [2]
- (iii) Carry out a hypothesis test at the 1% significance level to determine whether there appears to be positive correlation between pulse rate before exercise and pulse rate after exercise, stating your hypotheses. [6]
- (iv) State the distributional assumption which is necessary for this test to be valid. Explain briefly in the light of the scatter diagram why it appears that the assumption may be valid. [2]
- (v) If this distributional assumption is not valid, what alternative test procedure could be performed? If you were to carry out this test, what null hypothesis would you be testing? [2]
- 2 Carrots are supplied to a supermarket chain by two different suppliers. The weights,  $x$  grams, of random samples of carrots from each supplier are summarised as follows.

Supplier A	Sample size $n_A = 100$	$\Sigma x_A = 10\,313$	$\Sigma x_A^2 = 1\,072\,660$
Supplier B	Sample size $n_B = 95$	$\Sigma x_B = 9\,672$	$\Sigma x_B^2 = 986\,540$

- (i) Carry out a test at the 1% significance level to determine whether carrots supplied by the two suppliers have the same mean weight, stating your hypotheses and your conclusion. [12]
- (ii) Explain why it is valid to carry out this test without knowing the population variances. [2]
- (iii) Assuming that the population mean weights are the same, explain why the supermarket chain might prefer carrots from supplier B. [1]

- 3 A researcher is investigating the speed of shipping in the English Channel. She records the speeds in knots of a random sample of 13 tankers and 15 general cargo vessels. The results, given in ascending order, are as follows.

Tanker	5.9	8.3	8.5	9.5	10.9	11.1	11.2	11.5	12.6	13.4	13.7	13.9	15.1		
General cargo vessels	9.6	10.6	11.3	12.0	12.5	12.7	13.1	13.2	13.5	13.8	14.5	16.1	16.7	17.0	24.4

The researcher wishes to investigate whether, on average, the speeds of tankers and the speeds of general cargo vessels are the same. The underlying populations are not considered to be Normally distributed.

- (i) State suitable null and alternative hypotheses for a test which can be performed to investigate whether, on average, the speeds are the same. Carry out the test at the 5% significance level. [10]
- (ii) Name a test which could be performed if the population distributions were Normal. What property of the variances of the two populations is required to perform this test? [3]
- (iii) If, instead, the numbers of tankers and general cargo vessels in the sample were the same, would it be appropriate to carry out a signed rank test? Briefly explain your answer. [2]

[Question 4 is printed overleaf.]

**Section B (27 marks)**

**4** As part of a large construction project, over the course of three weeks a large number of lorries deliver liquid concrete to the project site. In order to check the quality of the concrete, a number of samples of concrete are selected. For each sample, a small quantity of liquid concrete is taken from a lorry and left to harden. It is then tested for compressive strength.

(i) Explain what population is being sampled in this case. [2]

The project engineer considers a number of methods of conducting the sampling.

- Choose a single day and sample all of the lorries which arrive on that day.
- Sample the first lorry which arrives on each day.
- On each day, select a lorry at random from the lorries due to arrive on that day.

(ii) Comment briefly on whether each of these methods is likely to result in a representative sample. [3]

(iii) If there were a list of all of the concrete deliveries expected at the site during the three weeks, explain briefly how a simple random sample of size 15 could be selected. [3]

The concrete is supplied by two different companies, A and B. The project engineer wishes to test whether there is any difference in the mean compressive strength of the concrete supplied by the two companies. The sampling process results in 9 samples from Company A and 6 samples from Company B. The compressive strengths, measured in suitable units, are summarised as follows.

Company A	$n = 9$	Sample mean 4.210	Sample standard deviation 0.2482
Company B	$n = 6$	Sample mean 4.295	Sample standard deviation 0.2459

(iv) State appropriate null and alternative hypotheses for a  $t$  test to determine whether there is any evidence that the mean compressive strengths of the concrete from the two companies are different. An assumption required for this test is equality of population variances. What other distributional assumption is required? Carry out the test at the 5% level. [13]

At a later stage in the project, a further large quantity of concrete is required. The project engineer again checks the quality of the concrete from each company. For a period of 10 days he takes two samples per day, one from a lorry belonging to Company A and one from a lorry belonging to Company B. He suspects that the weather conditions on each day may have an effect on the quality of the concrete. He wishes to check if the mean compressive strengths of the concrete supplied by the two companies are equal.

(v) Explain briefly why a paired  $t$  test should be carried out in this situation. [2]

(vi) The project engineer carries out a paired  $t$  test. He calculates the usual  $t$  statistic for this situation and finds this value to be 2.31. Complete the test, using a 5% significance level, stating your conclusion clearly. [4]

**Copyright Information**

OCR is committed to seeking permission to reproduce all third-party content that it uses in its assessment materials. OCR has attempted to identify and contact all copyright holders whose work is used in this paper. To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced in the OCR Copyright Acknowledgements Booklet. This is produced for each series of examinations, is given to all schools that receive assessment material and is freely available to download from our public website ([www.ocr.org.uk](http://www.ocr.org.uk)) after the live examination series.

If OCR has unwittingly failed to correctly acknowledge or clear any third-party content in this assessment material, OCR will be happy to correct its mistake at the earliest possible opportunity.

For queries or further information please contact the Copyright Team, First Floor, 9 Hills Road, Cambridge CB2 1GE.

OCR is part of the Cambridge Assessment Group; Cambridge Assessment is the brand name of University of Cambridge Local Examinations Syndicate (UCLES), which is itself a department of the University of Cambridge.

**Statistics (MEI)**

Advanced Subsidiary GCE **G243**

Statistics 3 (Z3)

**Mark Scheme for June 2010**

---

OCR (Oxford Cambridge and RSA) is a leading UK awarding body, providing a wide range of qualifications to meet the needs of pupils 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.

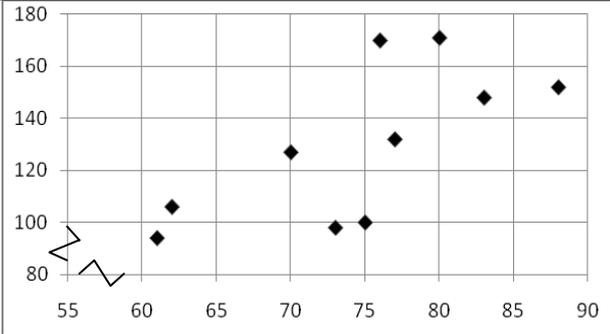
© OCR 2010

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](mailto:publications@ocr.org.uk)

## Question 1

(i)	 <p style="text-align: center;">Before</p>	<p>G1 for axes including labels  G1 for correct zero or broken scales  G1 for all points (allow 1 error)</p>	<b>3</b>
(ii)	$r = 0.6919$	M1 A1	<b>2</b>
(iii)	$H_0: \rho = 0$ $H_1: \rho > 0$ (one-tailed test) where $\rho$ is the correlation coefficient for the underlying bivariate population For $n = 10$ , 1% critical value = 0.7155 Not significant It does not appear that there is correlation between pulse rate before exercise and pulse rate after exercise.	B1 B1  B1  A1  E1 E1	<b>6</b>
(iv)	Bivariate Normality Vaguely elliptical scatter diagram	E1 E1	<b>2</b>
(v)	A test based on Spearman's rank correlation coefficient. $H_0$ : No association between pulse rate before exercise and pulse rate after exercise.	E1  E1	<b>2</b>
			<b>15</b>

## Question 2

(i)	<p><math>H_0: \mu_A = \mu_B</math>  <math>H_1: \mu_A \neq \mu_B</math>  Where <math>\mu_A, \mu_B</math> denote the population mean weights for carrots supplied by the two suppliers</p> $\bar{x}_A = \frac{10313}{100} = 103.13 \quad \bar{x}_B = \frac{9672}{95} = 101.81$ $s_A^2 = \frac{1}{99} \left( 1072660 - \frac{10313^2}{100} \right) = \frac{9080.31}{99} = 91.72$ $s_B^2 = \frac{1}{94} \left( 986540 - \frac{9672^2}{95} \right) = \frac{1828.59}{94} = 19.45$ <p>2-sample test based on <math>N(0,1)</math></p> <p>Test statistic is</p> $\frac{103.13 - 101.81}{\sqrt{\frac{91.72}{100} + \frac{19.45}{95}}} = \frac{1.32}{\sqrt{1.1220}} = \frac{1.32}{1.0592} = 1.246$ <p>2-tailed 1% point of <math>N(0,1)</math> is 2.576  Not significant  No reason to suppose that mean weights are different.</p>	<p>B1  B1  B1</p> <p>B1</p> <p>M1  A1</p> <p>M1</p> <p>M1  A1</p> <p>B1  E1  E1</p>	<p><b>12</b></p>
(ii)	<p>Because the samples are large and so it is reasonable to use the sample variances as the population variances.</p>	<p>E1  E1</p>	<p><b>2</b></p>
(iii)	<p>EITHER: Because they have a much lower standard deviation.  OR: Because they are less variable.</p>	<p>E1</p>	<p><b>1</b></p>
			<p><b>15</b></p>

## Question 3

(i)	<p><math>H_0</math>: the medians of the two populations are the same  <math>H_1</math>: the medians of the two populations are different</p> <p>Wilcoxon rank sum test (or Mann-Whitney form thereof)  Ranks are  Tankers 1,2,3,4,7,8,9,11,14,18,20,22,24  General Cargo 5,6,10,12,13,15,16,17,19,21,23,25,26,27,28  Rank sum for smaller sample (tankers) is 143</p> <p>Refer to (13,15) table  2-tail 5% critical value is 145 [ or 54 for M-W]  Significant  There is sufficient evidence to suggest that the median speeds are different</p>	<p>B1 Allow 1 for medians  B1 need population for second mark</p> <p>M1</p> <p>M1 Combined ranking  A1 FT if M1 earned</p> <p>B1 (M-W stat =  0+0+0+0+2+2+2+3+5+8+9+10+11=52)</p> <p>M1  A1  E1  E1</p>	<b>10</b>
(ii)	<p>Unpaired <math>t</math> test –allow 2-sample <math>t</math> test</p> <p>The variances must be equal</p>	<p>E1 for unpaired  E1 for <math>t</math> test  E1</p>	<b>3</b>
(iii)	<p>No  Because the data are not paired.</p>	<p>E1 for no  E1 for explanation</p>	<b>2</b>
<b>TOTAL</b>			<b>15</b>

## Question 4

(i)	The population consists of all of the concrete in all of the lorries arriving at the project site.	E1 all of the concrete E1 in all of the lorries	<b>2</b>
(ii)	The day chosen may not be representative of all days There may be some variation in compressive strength during the day, so selecting the first lorry may not give a representative sample This is much better than the other two methods and should produce a fairly representative sample	E1 E1  E1	<b>3</b>
(iii)	Allocate numbers 1 to $n$ to the deliveries. Use random number tables or a random number generator to choose 15 random numbers. If any repeats appear, choose further random numbers to replace them.	E1 E1 E1	<b>3</b>
(iv)	$H_0: \mu_A = \mu_B$ $H_1: \mu_A \neq \mu_B$ Where $\mu_A, \mu_B$ denote the population mean compressive strengths for Companies A and B  Assumption required is Normality of <u>both populations</u>  $\text{Pooled } s^2 = \frac{8 \times 0.2482^2 + 5 \times 0.2459^2}{13} = \frac{0.79516}{13} = 0.06116$  $\text{Test statistic} = \frac{4.21 - 4.295}{\sqrt{0.06116} \sqrt{\frac{1}{9} + \frac{1}{6}}} = \frac{-0.085}{0.1303} = -0.652$  Refer to $t_{13}$ 2-tail 5% point is 2.160  Not significant There is insufficient evidence to show that the population means are different	B1  B1  B1  M1 for attempt at pooling A1  M1 for numerator M1 for $\sqrt{0.06116}$ M1 for $\sqrt{\frac{1}{9} + \frac{1}{6}}$ A1 (allow +0.652 if doing B – A) FT from here if all M marks earned  M1 A1  E1 E1	<b>13</b>
(v)	Because each day one lorry from each company is sampled and there is expected to be day to day variation in the compressive strength which will be eliminated by pairing.	E1 E1	<b>2</b>
(vi)	Refer to $t_9$ 2-tailed 5% point is 2.262 Significant It appears that the population means are different.	M1 A1 E1 E1	<b>4</b>
		<b>TOTAL</b>	<b>27</b>

**OCR (Oxford Cambridge and RSA Examinations)**  
**1 Hills Road**  
**Cambridge**  
**CB1 2EU**

**OCR Customer Contact Centre**

**14 – 19 Qualifications (General)**

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**

## Chief Examiner's Report

The Principal Examiners' reports that follow discuss the candidates' performances on the individual modules. There is one matter that should be discussed in a general way as it applies to all the statistics modules. This is in respect of arithmetical accuracy in intermediate working and in quotation of final answers.

Most candidates are sensible in their arithmetical work, but there is some unease as to exactly what level of accuracy the examiners are expecting. There is no general answer to this! The standard rubric for all the papers sums the situation up by including "final answers should be given to a degree of accuracy appropriate to the context". Three 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 from Normal tables, *some* evidence of interpolation is generally expected and so quotation to four 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.

Talking now in general terms, the examiners always exercise sensible discretion in cases 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, a candidate is likely to lose an Accuracy mark; but if 4 significant figures are given, there would normally be no penalty. Likewise, answers which are slightly deviant from what is expected in a very minor manner are not penalised (for example, a Normal probability given, after an attempt at interpolation, as 0.6418 whereas 0.6417 was expected). However, there are increasing numbers of cases where candidates give answers which are *grossly* over- or under-specified, such as insistence that the value of a test statistic is (say) 2.128888446667 merely because that is the value that happens to come off the candidate's calculator. **Such gross over-specification indicates a lack of appreciation of the nature of statistical work and, with effect from the January 2011 examinations, will be penalised by withholding of associated Accuracy marks.**

Candidates must however always be aware of the dangers of premature rounding if there are several steps in a calculation. If, say, a final answer is desired that is correct to 3 decimal places, this can in no way be guaranteed if only 3 decimal places are used in intermediate steps; indeed, it may not be safe to carry out the intermediate work even to 4 decimal places. The issue of over-specification may arise for the final answer but not for intermediate stages of the working.

It is worth repeating that most candidates act sensibly in all these respects, but it is hoped that this note may help those who are perhaps a little less confident in how to proceed.

## G243 Statistics 3

### Comments on individual questions

- 1) (i) Most candidates gained full marks here, but some had axes starting at for example 90 without using a broken scale. Others did not know how to draw a scatter diagram, instead plotting the runners' 'names' on the  $x$ -axis and the times before and after on the  $y$ -axis.
- 1) (ii) Very few candidates used their calculators' built in functions to calculate the correlation coefficient and there were a number of errors in using the formula to calculate the coefficient.
- 1) (iii) Most candidates gave their hypotheses correctly in terms of  $\rho$ , but hardly any defined  $\rho$ . Most found the critical value although a few gave the two-tailed value instead. Most candidates then completed the test correctly. A very few tried to carry out other types of test.
- 1) (iv) Most candidates correctly said that a bivariate Normal distribution was required and most knew that an elliptical scatter diagram would indicate that the sample came from this distribution.
- 1) (v) Many correct answers were seen but a number of candidates thought that the null hypothesis was 'no correlation' rather than the correct 'no association'.
- 2) (i) There were a number of fully correct solutions. Almost all candidates found the mean and most also found the variance or standard deviation correctly. The majority of candidates made a good attempt at the test statistic, but occasionally there were errors such as using  $s$  instead of  $s^2$  in calculating the pooled variance. The hypotheses were often correctly given either in symbols or, more often, in words, but most candidates failed to mention 'population' when giving hypotheses in words. The critical value was usually correct, and most candidates who had got this far went on to complete the test correctly. A very small number of candidates tried to carry out a  $t$ -test.
- 2) (ii) Most candidates said that the samples were large, but rather less went on to explain that the sample variances could be used as if they were the population variances without appreciable error as the samples were large.
- 2) (iii) The majority of candidates correctly realised that the lower variance of the carrots from Supplier B might make them preferable.
- 3) (i) Most candidates ranked the data and found the Wilcoxon test statistic, although a few attempted to carry out a paired test despite the values of  $n$  being different. Most then went on to complete the test correctly, although a few compared the Wilcoxon test statistic to the Mann-Whitney critical value (it would of course be acceptable to carry out the whole procedure according to the Mann-Whitney formulation). Most gave their hypotheses in terms of the median, although a few incorrectly wrote average or even mean. Many failed to mention population.
- 3) (ii) Most candidates answered this well although several did not mention that the test should be unpaired.

*Report on the Unit taken in June 2010*

- 3) (iii) This was not well answered. Very few gave a sensible reason for a signed rank test being inappropriate.
- 4) (i) Many candidates correctly identified the population as consisting of the concrete, but few gave a fully correct description – ‘all of the concrete in all of the lorries’.
- 4) (ii) Most candidates realised that the third method was the best and correctly identified the failings in the first two methods.
- 4) (iii) There were many correct solutions to this part, although several failed to mention repeats. A few tried to describe a systematic sample.
- 4) (iv) Some candidates gave the hypotheses correctly in symbols and defined their symbols. Rather more gave them in words, usually correctly. Few stated the necessary assumption of Normality of both populations. Most candidates found the pooled variance although several candidates gave their result as 0.2473, the standard deviation. The method for finding the test statistic was well known and candidates could follow through to gain the final marks (provided that they scored the method marks here). Most realised that they needed to use 13 degrees of freedom and found the correct critical value, usually going on to complete the test correctly.
- 4) (v) There was a wide variety of wrong answers here, such as ‘the samples are small’, or ‘there are equal numbers in the two samples’. Very few gave an answer that was reasonable.
- 4) (vi) Most candidates gained full marks here, although a few had the wrong critical value and scored zero.

**OCR (Oxford Cambridge and RSA Examinations)**  
**1 Hills Road**  
**Cambridge**  
**CB1 2EU**

**OCR Customer Contact Centre**

**14 – 19 Qualifications (General)**

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 2010

