

OXFORD CAMBRIDGE AND RSA EXAMINATIONS
Advanced Subsidiary General Certificate of Education

MEI STATISTICS

G243

Statistics 3 (Z3)

Thursday **15 JUNE 2006** Afternoon 1 hour 30 minutes

Additional materials:
8 page answer booklet
Graph paper
MEI Examination Formulae and Tables (MF2)

TIME 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

- Write your name, centre number and candidate number in the spaces provided on the answer booklet.
- Answer **all** the questions.
- 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 question paper consists of 5 printed pages and 3 blank pages.

Section A (45 marks)

- 1 Agricultural scientists are studying the relationship between the yield of a crop and the amount of rainfall. Data from a random sample of 9 farms give the following values of y , the yield of the crop measured in a convenient standard unit, and x , the rainfall during the growing season measured in mm.

Farm	A	B	C	D	E	F	G	H	I
x	152	74	96	310	248	105	140	186	161
y	26.4	18.8	21.3	37.0	34.6	18.0	26.7	31.2	29.0

- (i) Draw a scatter diagram to illustrate these data, and comment on it briefly. [5]
- (ii) Calculate the value of Spearman's rank correlation coefficient. [4]
- (iii) Hence test at the 5% level of significance whether it is reasonable to assume that rainfall and yield are associated. [3]
- (iv) Give a reason, referring to your scatter diagram, why you might be unwilling to base a test on the value of the product moment correlation coefficient for these data. If you were to carry out such a test, what null hypothesis would you be testing? [3]
- 2 A study is being made of towns in two large areas, A and B, as part of a comparison of overall residential attractiveness. Each area has many towns and it is not practicable to include them all in the study. A sample of 8 towns is selected from area A and a sample of 10 from area B. Each is regarded as a random sample from a suitably defined underlying population.
- (a) (i) One of the criteria is accessibility to transport links. A transport analyst ranks the 18 towns in order of this accessibility (rank 1 denoting the most accessible). The ranks are as follows.
- | | | | | | | | | | | | | | | | | | | |
|-----------------|---|---|---|---|---|----|----|----|----|----|----|----|----|--|--|--|--|--|
| Towns in area A | 1 | 2 | 4 | 5 | 6 | 10 | 11 | 13 | | | | | | | | | | |
| Towns in area B | | | | 3 | 7 | 8 | 9 | 12 | 14 | 15 | 16 | 17 | 18 | | | | | |
- Test at the 5% level of significance whether, on the whole, the two areas may be assumed equivalent in this regard, stating your null and alternative hypotheses. [8]
- (ii) Comment very briefly on why this part of the study was conducted as a ranking exercise, and state one drawback of this. [2]
- (b) Another criterion is air pollution. The pollution level is measured for each town in the samples. Underlying Normality may be assumed. A statistician wishes to compare the means for the two areas by carrying out a two-sample test.
- (i) The statistician uses a test based on the Normal distribution. What further information does he need when calculating the value of the test statistic? [2]
- (ii) The value of the test statistic is 1.82. Carry out the test, using a 5% significance level and a two-sided interpretation, giving your conclusion. [3]

- 3 (i) Describe the circumstances when systematic sampling might sensibly be used as a method for selecting a sample for a survey. [2]

A Psychology Department at a university has 96 first-year students. There is a list of them in alphabetical order. A sample of size 12 is required for a survey.

- (ii) Describe how this sample could be selected by systematic sampling. [2]

The survey is concerned with overall academic ability and aptitude during the transition from school to university. The 12 selected students take two tests, one on entering the university and the other after three months. The 12 students are regarded as a random sample from the student population. The scoring systems for the tests have been designed so that, on the whole, equivalent performances in the two tests should result in equal scores and improved performance should result in a higher score. Underlying Normality may be assumed, as appropriate.

- (iii) The students' scores are as follows.

Student	A	B	C	D	E	F	G	H	I	J	K	L
First test	400	371	365	529	346	371	414	286	453	320	314	265
Second test	438	389	437	558	334	470	437	327	451	366	366	327

Use a t test to examine, at the 5% level of significance, whether it appears that the mean score for students in the second test is greater than that for the first. State carefully your null and alternative hypotheses and your conclusion. [11]

[Question 4 is printed overleaf.]

Section B (27 marks)

- 4 Engineers at a chemical company are trying to improve the performance of a process for making a particular type of plastic. The yield of the plastic is affected by the pressure and temperature in the reaction chamber in which it is made. The engineers hope that changing the pressure or the temperature, or both of them, will improve the average yield. However, they fear that such changes might increase the variability of the yield.

The process currently runs at temperature 550°C and at pressure 2.6 units.

- (i) The engineers carry out an experiment with four trial runs with settings as follows.

Run A: pressure 2.2, temperature 500°C .
 Run B: pressure 3.0, temperature 500°C .
 Run C: pressure 2.2, temperature 600°C .
 Run D: pressure 3.0, temperature 600°C .

Give a reason why the design of this experiment appears sensible for investigating whether the average yield can be improved, assuming there is no change in variability. [2]

Give a reason why the design of this experiment is not suitable for investigating whether the variability of yield has changed. [2]

The engineers establish that the best pressure is 2.6 and the best temperature is 625°C . All future experiments are carried out at this pressure and this temperature.

The company's factory has two reaction chambers in which the process can be run, and the next task is to establish whether or not these two chambers perform similarly in terms of the average yield of the plastic.

- (ii) The engineers carry out a further experiment. They undertake five trial runs in one chamber, one run on each day of a working week. The next week, they similarly undertake five trial runs in the other chamber. The yields, in kg, of the plastic are as follows.

First chamber	26.8	28.4	27.1	21.9	30.6
	Sample mean $\bar{x} = 26.96$				
	Sample variance $s_1^2 = 10.243$				
Second chamber	23.2	22.5	20.9	28.6	29.4
	Sample mean $\bar{y} = 24.92$				
	Sample variance $s_2^2 = 14.647$				

Assuming that the underlying variances for the two chambers are the same, test the null hypothesis $\mu_1 = \mu_2$, where μ_1 and μ_2 are the population mean yields for the two chambers, against the alternative $\mu_1 \neq \mu_2$, at the 5% level of significance, stating your conclusion carefully. [10]

The company statistician suggests that yields may be affected by day-to-day variability in the quality of the raw materials supplied to the factory.

- (iii) Explain why the experiment in part (ii) is unsatisfactory. Explain how a paired comparison experiment could be carried out in this context and why it would be an improvement. [4]

[Question 4 continues on the next page.]

The statistician also suggests that the assumption of underlying Normality that was needed for the test in part (ii) may not be valid. The engineers therefore carry out a new experiment according to a paired comparison design, and ask the statistician to analyse the data.

(iv) The data are as follows. There are 10 paired trials, labelled I, II, ..., X.

Trial	I	II	III	IV	V	VI	VII	VIII	IX	X
Yield (kg) in first chamber	29.4	28.0	27.8	24.2	26.9	30.0	23.4	21.8	21.7	31.0
Yield (kg) in second chamber	29.0	28.5	26.6	24.5	26.2	28.0	23.5	21.2	20.2	30.0

Use the Wilcoxon signed rank test, at the 5% level of significance, to examine whether the underlying medians may be assumed equal. [9]

BLANK PAGE

BLANK PAGE

BLANK PAGE

G243 Statistics 3 (Z3)

June 2006

Mark Scheme

MAXIMUM MARK	72
---------------------	-----------

Q1																																					
(i)	<p>Looks reasonably linear, maybe flattens a bit at the top.</p>	G1 G1 G1 E1 E1	Axes, including labels. “x” and “y” suffice as they are defined in the question. Correct zero, or clear “breaks”. All points correct. Allow one error.	5																																	
(ii)	<table border="0"> <tr> <td>x ranks</td> <td>5</td> <td>1</td> <td>2</td> <td>9</td> <td>8</td> <td>3</td> <td>4</td> <td>7</td> <td>6</td> <td></td> </tr> <tr> <td>y ranks</td> <td>4</td> <td>2</td> <td>3</td> <td>9</td> <td>8</td> <td>1</td> <td>5</td> <td>7</td> <td>6</td> <td></td> </tr> <tr> <td>d</td> <td>1</td> <td>1</td> <td>1</td> <td>0</td> <td>0</td> <td>2</td> <td>1</td> <td>0</td> <td>0</td> <td>$\Sigma d^2 = 8$</td> </tr> </table> $r_s = 1 - \frac{6 \times 8}{9 \times 80} = 0.9333$	x ranks	5	1	2	9	8	3	4	7	6		y ranks	4	2	3	9	8	1	5	7	6		$ d $	1	1	1	0	0	2	1	0	0	$\Sigma d^2 = 8$	B1 B1 M1 A1	c.a.o.	4
x ranks	5	1	2	9	8	3	4	7	6																												
y ranks	4	2	3	9	8	1	5	7	6																												
$ d $	1	1	1	0	0	2	1	0	0	$\Sigma d^2 = 8$																											
(iii)	Critical value for $n = 9$ at two-sided 5% level is 0.7000. Significant. Seems there is an association between rainfall and yield.	B1 E1 E1	No fit if wrong. S.C. Use of the 1-tail point (0.6000) can get either, but not both, of these E marks.	3																																	
(iv)	No real suggestion of bivariate Normality. H_0 is $\rho = 0$, where ρ is the correlation coefficient for the underlying bivariate population.	E1 B1 B1	Or equivalent statement.	3																																	
				15																																	

Q2				
(a) (i)	H_0 : the medians of the two populations are the same. H_1 : the medians of the two populations are different. Wilcoxon rank sum test (or Mann-Whitney form thereof). Ranks are A 1 2 4 5 6 10 11 13 B 3 7 8 9 12 14 15 16 17 18 Rank sum for smaller sample is 52. Refer to (8,10) table. 2-tail 5% critical value is 53. Significant. Seems areas are not equivalent in this regard.	B1 B1 M1 B1 M1 A1 E1 E1	Or more formal statements. (Or M-W statistic = 16) No fit from here if wrong. (Or 17 for M-W). No fit if wrong. ft only c's test statistic. ft only c's test statistic.	8
(ii)	There is nothing obvious that can be measured. Depends on subjective judgement of analyst.	E1 E1	Or other sensible comments. Or other sensible comments.	2
(b) (i)	The population standard deviation.	B1 B1		2
(ii)	The critical point is 1.96. Not significant. Appears that the average air pollution in towns in the two regions may be assumed to be the same.	M1 E1 E1	No fit if wrong.	3
				15

Q3				
(i)	There is a list of the population. (As far as is known) there are no cycles or patterns in the list that might be related to the aim of the sampling.	E1 E1		2
(ii)	Pick one of first 8 in alphabetical list at random then every 8 th .	M1 M1		2
(iii)	<p>$H_0: \mu_D = 0$ (or $\mu_1 = \mu_1$ etc)</p> <p>$H_1: \mu_D > 0$ (or $\mu_2 > \mu_1$ etc) where μ_D is the population mean for the differences.</p> <p><u>MUST</u> be PAIRED COMPARISON t test. Use of differences. Differences are: 38 18 72 29 -12 99 23 41 -2 46 52 62 $\bar{d} = 38.8\dot{3}$ $s_{n-1} = 30.924(64)$</p> <p>Test statistic is $\frac{38.8\dot{3} - 0}{\frac{30.924(64)}{\sqrt{12}}}$ $= 4.35.$</p> <p>Refer to t_{11}. Upper 5% point is 1.796. Significant. Seems mean score for second test is greater than for first.</p>	<p>B1</p> <p>B1 B1</p> <p>M1</p> <p>A1</p> <p>M1</p> <p>A1</p> <p>M1 A1 E1 E1</p>	<p>Do NOT allow \bar{D} or similar unless it is clearly and explicitly stated to be a <u>population</u> mean. Hypotheses in words only must include “population”. Or “<” for “first” – “second”. For adequate verbal definition. Allow absence of “population” here if correct notation μ has been used.</p> <p>For both. [$s_n = 29.6081$ <u>NOT</u> allowed.]</p> <p>Allow c’s \bar{d} and/or s_{n-1}. Allow alternative: $0 + (c's 1.796) \times \frac{30.924(64)}{\sqrt{12}}$ (= 16.03) for subsequent comparison with \bar{d}. (Or $\bar{d} - (c's 1.796) \times \frac{30.924(64)}{\sqrt{12}}$ (= 22.80) for comparison with 0.) c.a.o. but ft from here if M1 awarded, but no marks from here on if not paired t test. Use of $0 - \bar{d}$ scores M1A0, but ft.</p> <p>No ft from here if wrong. No ft from here if wrong. ft only c’s test statistic. ft only c’s test statistic.</p>	11
				15

Q4				
(i)	Both the pressure and the temperature are tried at settings above and below those currently used. There is no replication so variability cannot be assessed.	E2	(E0, E1, E2)	2
		E2	(E0, E1, E2)	2
(ii)	<p>Pooled $s^2 = \frac{(4 \times 10 \cdot 243) + (4 \times 14 \cdot 647)}{8} = 12 \cdot 445$</p> <p>Test statistic is</p> $\frac{26 \cdot 96 - 24 \cdot 92}{\sqrt{12 \cdot 445 \sqrt{\frac{1}{5} + \frac{1}{5}}}} = \frac{2 \cdot 04}{2 \cdot 2311} = 0 \cdot 9143$ <p>Refer to t_8. Double-tailed 5% point is 2.306. Not significant. No evidence that population mean yields are different.</p>	M1 A1	For any reasonable attempt at pooling. If correct.	10
		M1	For numerator.	
		M1	$\sqrt{12 \cdot 445}$. Use of c's pooled variance.	
		M1	$\sqrt{\frac{1}{5} + \frac{1}{5}}$.	
		A1	ft from here if all M marks earned.	
		M1	No ft from here if wrong.	
		A1	No ft from here if wrong.	
		E1	ft only c's test statistic.	
		E1	ft only c's test statistic.	
(iii)	All runs in first chamber are on different days from all runs in second, so any variability in the raw materials may affect the chambers in different ways. Do a run in each chamber on Monday, then a run in each chamber on Tuesday, and so on. Any variability would affect each chamber equally.	E1		4
		E1		
		E1	Or equivalent.	
		E1		
(iv)	<p>Differences are 0.4 -0.5 1.2 -0.3 0.7 2.0 -0.1 0.6 1.5 1.0</p> <p>Ranks of d 3 4 8 2 6 10 1 5 9 7</p> <p>Test statistic is $4 + 2 + 1 = 7$ (or $3 + 8 + 6 + 10 + 5 + 9 + 7 = 48$)</p> <p>Refer to paired Wilcoxon table with $n = 10$ Need lower 2½ % point which is 8 (or, if 48 used, upper 2½ % point which is 47). Result is significant. Seems underlying medians are not the same.</p>	B1 M1 A1		9
		M1		
		A1	No ft from here if wrong.	
		M1	No ft from here if wrong.	
		A1		
		E1	ft only c's test statistic.	
		E1	ft only c's test statistic.	
				27