OCR
RECOGNISING ACHIEVEMENT

Thursday 6 June 2013 – Morning

AS GCE MEI STATISTICS

G242/01 Statistics 2 (Z2)

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:
- Printed Answer Book G242/01
- 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.

1 A health care inspector is monitoring the occurrence of the 'superbug' MRSA in a large hospital. Each day, he checks for the presence of MRSA in different areas of the hospital and records the number of areas where MRSA is detected. The following table summarises the results for a random sample of 300 days.

Number of areas where MRSA is detected	0	1	2	3	4
Observed frequency	86	108	74	24	8

(i) Show that the sample mean is 1.2. Given also that the sample standard deviation is 1.03, explain why the Poisson distribution may provide a suitable model for these data. **[3]**

The inspector wishes to test the goodness of fit of a Poisson model. The mean calculated from the data is used as an estimate for the mean of the underlying population to produce the following expected frequencies.

Number of areas where MRSA is detected	0	1	2	3	$\geqslant 4$
Expected frequency	90.36	108.43	65.06	26.02	10.13

(ii) Show how the expected frequency of 90.36 for 0 areas containing MRSA is calculated. **[3]**

(iii) Carry out the test of the goodness of fit of the Poisson model at the 5% level of significance. **[9]**

2 A gas boiler manufacturer is working to improve its domestic hot water boilers. It wishes to ensure that the boilers heat water as quickly as possible. New boilers are tested in controlled conditions by measuring the temperature of the water after 15 seconds of heating. The resulting temperatures, in degrees Celsius, for a random sample of 8 observations are as follows.

51.0 50.7 49.8 50.4 50.6 50.8 49.3 50.6

It is required to produce a confidence interval for the population mean temperature using these data.

(i) Stating a necessary assumption, explain why a confidence interval based on the *t* distribution is more appropriate, in this case, than a confidence interval based on the Normal distribution. **[4]**

(ii) Obtain a 95% confidence interval, based on the *t* distribution, for the population mean temperature. **[7]**

The manufacturer aims to produce a boiler that, in 15 seconds, heats water to 50 degrees Celsius on average.

(iii) With reference to the confidence interval found in part **(ii)**, comment on whether the manufacturer has been successful in its aim. **[2]**

3 A musical instrument manufacturer is developing a new range of violin strings made from a synthetic material. Part of the development process involves testing breaking strength. The breaking strengths of a random sample of 12 strings are measured. The results, in suitable units, are as follows.

347 345 349 348 350 349 349 344 354 347 351 350

(i) Explain, with reference to the sample, whether or not you think that these data could have an underlying Normal distribution. **[2]**

It is required that the mean breaking strength of synthetic strings should not be below 350.

(ii) Given that the sample standard deviation is 2.678, use a t test to examine, at the 5% significance level, whether this sample provides evidence that the mean breaking strength of these synthetic strings is below the minimum requirement. **[11]**

4 'Cool Milk Dairy' offers a delivery service for its milk. Milk is delivered each day in glass bottles. Empty bottles are collected so that they can be washed and used again. The dairy manager is monitoring the return of bottles to ensure there are enough bottles in stock to cover future demand. Over a long period of time she has established that, on average, 13 fewer bottles are collected each day than are delivered. For the past six months the dairy has been replacing lost stock with bottles from a different supplier and the dairy manager suspects that the average of 13 might have changed. A daily record of the difference between the number of bottles delivered, x, and the number of empty bottles collected, y, is kept. The differences, $x - y$, for ten consecutive days are as follows.

25 3 4 17 18 11 19 14 20 24

(i) Stating any necessary assumptions, use a Wilcoxon test to examine, at the 5% significance level, whether these data support the dairy manager's suspicions. **[13]**

(ii) Comment on the validity of any assumptions made in part **(i)**. **[2]**

[Question 5 is printed overleaf.]

5 A sea ferry operator is working to improve punctuality on one of its sailing routes. It decides to test for an association between the sea condition at the time of sailing and punctuality. The punctuality and sea condition of a random sample of 200 ferry crossings are recorded. The results are as follows.

		Punctuality	
		Not late	Late
Sea condition	Rough	36	12
	Moderate	91	12
	Slight	44	5

The following tables show some of the expected frequencies and contributions to the test statistic.

Expected frequencies		Punctuality	
		Not late	Late
Sea condition	Rough	41.040	6.960
	Moderate	88.065	
	Slight	41.895	

Contributions to the test statistic		Punctuality	
		Not late	Late
Sea condition	Rough	0.6189	3.6497
	Moderate	0.0978	
	Slight	0.1058	

(i) Calculate the remaining expected frequencies and contributions. Carry out the test using a 5% level of significance. **[11]**

The journey times for this particular route may be modelled using a Normal distribution with mean 160 minutes and standard deviation 5.5 minutes.

(ii) Given that the ferry departs at 08:30 and is scheduled to arrive at 11:15, find the probability that it is not late. **[5]**