OCR
Oxford Cambridge and RSA

Friday 17 June 2016 – Afternoon

AS GCE MEI STATISTICS

G243/01 Statistics 3 (Z3)

QUESTION PAPER

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 inside 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.** If additional space is required, you should use the lined page(s) at the end of this booklet. The question number(s) must be clearly shown.
- 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.

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.

DC (LK/SW) 124825/2

Section A (47 marks)

1 A researcher wishes to test whether two mathematics examinations based on the same syllabus are of equal difficulty. In order to do this the researcher randomly selects 120 students to take the exams, randomly allocating 60 of them to the first and the other 60 to the second.

(i) Explain why, even though the distributions of the parent populations and the population variances are unknown, a test based on the Normal distribution is appropriate. **[4]**

The means and variances of the marks for the 60 students who took each exam are shown below.

Exam A:	mean 55.7	variance 124.3
Exam B:	mean 57.2	variance 141.5

(ii) Carry out a test at the 10% significance level to investigate whether the examinations appear to be of equal difficulty. **[11]**

2 A researcher is investigating the effect of listening to different types of music on the times taken by athletes to run a distance of 400 metres. The researcher selects a random sample of 12 athletes and on one evening gets each of them to run this distance listening to classical music. The next evening the researcher gets each of the athletes to run this distance listening to rock music. The times in seconds are shown below.

Athlete	A	B	C	D	E	F	G	H	I	J	K	L
Time with classical music	57.2	59.1	68.3	63.5	57.8	63.5	64.8	60.0	63.0	65.0	57.1	56.1
Time with rock music	62.3	62.2	61.4	70.1	56.4	71.7	59.8	59.2	68.5	63.5	68.6	62.1

A test is to be carried out to investigate whether it appears that the times taken to run 400 metres listening to the two types of music are different.

(i) Explain why the researcher used the same set of athletes for both runs. **[2]**

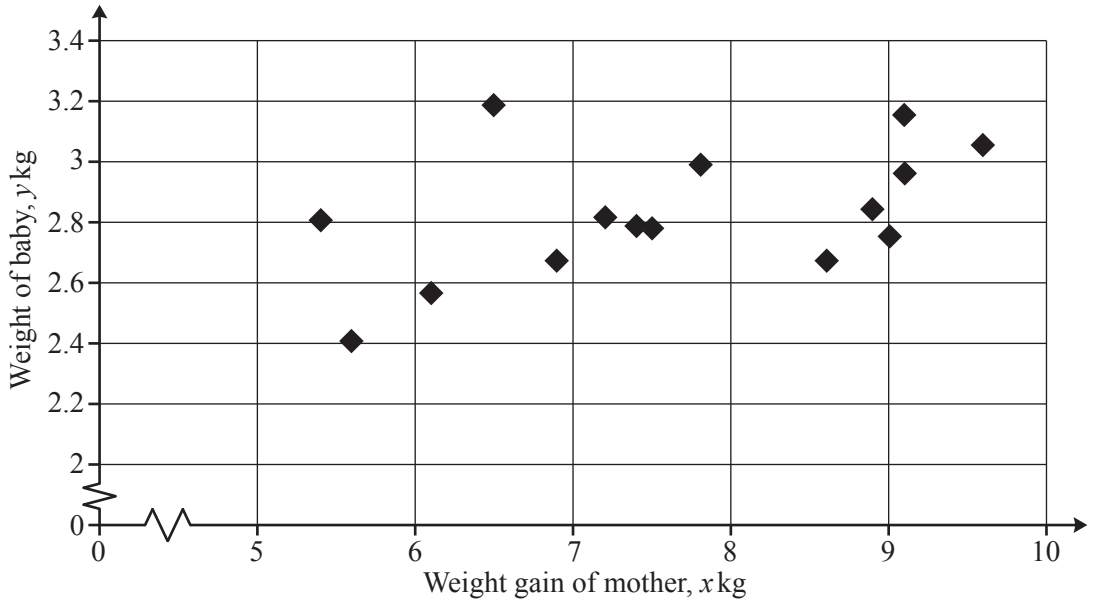
(ii) Given that the distribution of the population of differences is not known, use a suitable test to examine, at the 5% significance level, whether it appears that times taken are different. **[11]**

(iii) Explain why the researcher did not get the athletes to do both runs on the same evening. **[1]**

(iv) Briefly describe one improvement that the researcher could have made to the design of this experiment and explain why this change is an improvement. **[2]**

3 A scientist is investigating whether the weight, y kg, of a new-born baby is related to the gain in weight, x kg, of the mother during pregnancy. Summary statistics and a scatter diagram for the data collected by the scientist are shown below.

$\Sigma x = 114.7$ $\Sigma y = 42.45$ $\Sigma x^2 = 903.2$ $\Sigma y^2 = 120.8$ $\Sigma xy = 326.5$ $n = 15$



(i) Calculate the product moment correlation coefficient. **[5]**

(ii) What distributional assumption is required for a test based on the product moment correlation coefficient? Explain why, in view of the scatter diagram, it may be appropriate to carry out a test based on the product moment correlation coefficient. **[3]**

(iii) Carry out a test at the 5% significance level to investigate whether there is positive correlation between x and y. **[6]**

(iv) Because the correlation coefficient is positive it is suggested that, in order to have heavier babies, mothers should gain more weight during pregnancy. Give two possible reasons why this may not be the case. **[2]**

Section B (25 marks)

4 A motoring magazine correspondent is investigating the fuel consumption figures for a particular model of car on urban roads. (Urban roads are roads within a town.) The correspondent has previously tested the urban fuel consumption of a sample of 7 of these cars and has found

- the sample mean to be 5.07 litres per 100 km,
- the sample standard deviation to be 0.23 litres per 100 km.

The car manufacturer states that a new 'lower emission' engine will reduce the fuel consumption by 0.5 litres per 100 km. The correspondent thinks that the reduction will be less than 0.5 litres per 100 km. In order to investigate this, the magazine's editor suggests checking the fuel consumption of one car with the new engine over a distance of 10 km.

(i) Give two reasons why the editor's approach is not suitable. **[2]**

The correspondent decides to select a sample of cars and measure the fuel consumption of these cars on urban roads. In order to do this, the magazine asks its readers to contact the magazine if they drive cars of this model with the new engine. A total of 124 people contact the magazine. The correspondent selects a random sample of 10 of these people.

(ii) Explain how a random sample of 10 people could be selected. **[3]**

The fuel consumption figures (in litres per 100 km) for the cars driven by the random sample of 10 people are as follows.

4.89 4.67 4.55 4.36 5.13 4.52 5.05 4.89 4.29 4.57

(iii) Calculate the sample mean and sample standard deviation for these data. **[2]**

(iv) State two distributional assumptions which are required for a two-sample *t* test to investigate whether the manufacturer's claim is true. **[2]**

The correspondent wishes to test whether the reduction will be less than 0.5 litres per 100 km. The null hypothesis for a *t* test to determine whether this appears to be true is $H_0: \mu_{new} = \mu_{old} - 0.5$.

(v) State the meaning of μ_{new}. State the alternative hypothesis and explain why it takes the form that it does. Given that the assumptions required in part **(iv)** are valid, carry out the test at the 5% significance level. **[13]**

(vi) Sometimes the assumptions required for an unpaired *t* test are not valid. Under such circumstances, what alternative test can be carried out? What distributional assumption does this alternative test require? **[3]**

END OF QUESTION PAPER