MEI STRUCTURED MATHEMATICS
STATISTICS 2, S2   **4767**

Specimen Paper

Additional materials: Answer booklet
                     Graph paper
                     MEI Examination Formulae and Tables (MF 2)

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 **may** use a graphical or scientific calculator in this paper.

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.
- Final answers should be given to a degree of accuracy appropriate to the context.
- The total number of marks for this paper is **72**.

1 A medical statistician wishes to carry out a hypothesis test to see if there is any correlation between the head circumference and body length of newly-born babies.

(i) State appropriate null and alternative hypotheses for the test. **[2]**

A random sample of 20 newly-born babies have had their head circumference, x cm, and body length, y cm, measured. This bivariate sample is illustrated in Fig. 1.
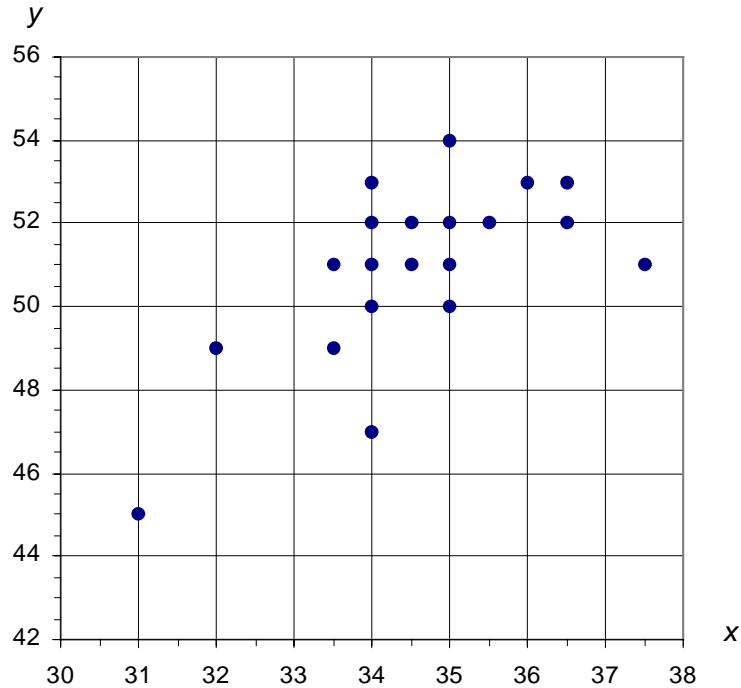

Fig. 1

Summary statistics for this data set are as follows.

$$n = 20 \qquad \sum x = 691 \qquad \sum y = 1018 \qquad \sum x^2 = 23\ 917 \qquad \sum y^2 = 51\ 904 \qquad \sum xy = 35\ 212.5$$

(ii) Calculate the product-moment correlation coefficient for the data.
Carry out the hypothesis test at the 1% significance level, stating the conclusion carefully.
What assumption is necessary for the test to be valid? **[10]**

Originally, the point $x = 34$, $y = 51$ had been recorded incorrectly as $x = 51$, $y = 34$.

(iii) Calculate the values of the summary statistics if this error had gone undetected.
Use the uncorrected summary statistics to show that the value of the product-moment correlation coefficient would be negative. **[4]**

(iv) How is it that this one error produces such a large change in the value of the correlation coefficient and also changes its sign? **[2]**

2 *Extralite* are testing a new long-life bulb. The life-times, in hours, are assumed to be Normally distributed with mean μ and standard deviation σ.

After extensive tests, they find that 19% of bulbs have a life-time exceeding 5000 hours, while 5% have a life-time under 4000 hours.

(i) Illustrate this information on a sketch. **[2]**

(ii) Show that $\sigma = 396$ and find the value of μ. **[5]**

In the remainder of this question take μ to be 4650 and σ to be 400.

(iii) Find the probability that a bulb chosen at random has a life-time between 4250 and 4750 hours. **[3]**

(iv) Find the probability that a bulb has a life-time of over 5450 hours. **[1]**

(v) *Extralite* wish to quote a life-time which will be exceeded by 99% of bulbs. What time, correct to the nearest 100 hours, should they quote? **[3]**

A new school classroom has 6 light-fittings, each fitted with an *Extralite* long-life bulb.

(vi) Find the probability that no more than one bulb needs to be replaced within the first 4250 hours of use. **[4]**

3 The numbers of goals per game scored by teams playing at home and away in the Premier League are modelled by independent Poisson distributions with means 1.63 and 1.17 respectively.

(i) Find the probability that, in a game chosen at random,

    (*A*) the home team scores at least 2 goals, **[4]**

    (*B*) the result is a 1-1 draw, **[3]**

    (*C*) the teams score 5 goals between them. **[4]**

(ii) Give **two** reasons why the proposed model might not be suitable. **[2]**

The number of goals scored per game at home by *Rovers* is modelled by the Poisson distribution with mean 1.63. In a season they play 19 home games.

(iii) Use a suitable approximating distribution to find the probability that *Rovers* will score more than 35 goals in their home games. **[5]**

4 **(a)** The length of metal rods used in an engineering structure is specified as being 40 cm.
It does not matter if they are slightly longer, but they should not be any shorter.
These rods are made by a machine in such a way that their lengths are Normally distributed with standard deviation 0.2 cm.
The mean, μ cm, of the lengths is set to a value slightly above 40 cm to give a margin for error.

To examine whether the specification is being met, a random sample of 12 rods is taken.
Their lengths, in cm, are found to be:

| 40.43 | 40.49 | 40.19 | 40.36 | 40.81 | 40.47 |
| 40.46 | 40.63 | 40.41 | 40.27 | 40.34 | 40.54 |

It is desired to test whether $\mu = 40.5$.

(i) State a suitable alternative hypothesis for the test. **[1]**

(ii) Carry out the test at the 5% level of significance, stating your conclusion carefully. **[8]**

(b) Data are extracted from the medical records of a random sample of patients of a large general practice, showing for part of a particular year the frequencies of contracting or not contracting influenza for patients who had or had not had influenza inoculations.

		Influenza	
		Yes	No
Inoculated	Yes	8	18
	No	35	17

State null and alternative hypotheses for a suitable test for independence of inoculation and occurrences of influenza.
Carry out the test at the 5% level of significance. **[9]**