OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE

MATHEMATICS (MEI) 4766

Statistics 1

QUESTION PAPER

Candidates answer on the printed answer book.

OCR supplied materials:
- Printed answer book 4766
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Thursday 26 May 2011
Morning

Duration: 1 hour 30 minutes

MODIFIED LANGUAGE

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. 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.

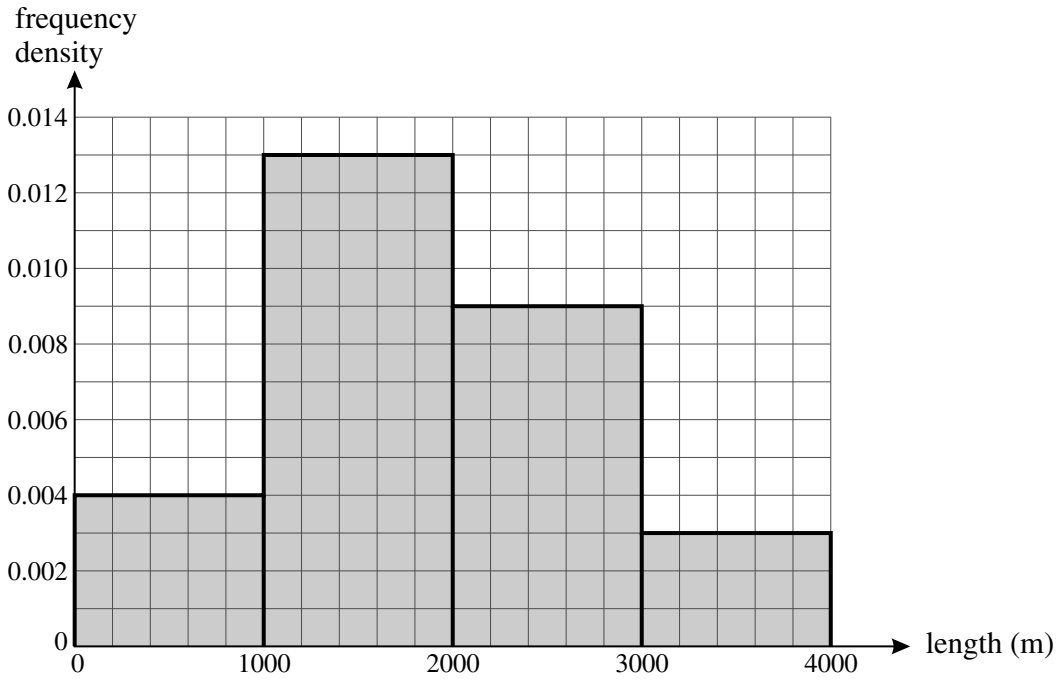
INSTRUCTION TO EXAMS OFFICER / INVIGILATOR

- Do not send this question paper for marking; it should be retained in the centre or destroyed.

Section A (36 marks)

1 In the Paris-Roubaix cycling race, there are a number of sections of cobbled road.

The lengths of these sections, measured in metres, are illustrated in the histogram.



(i) Find the number of sections which are between 1000 and 2000 metres in length. **[2]**

(ii) Name the type of skewness suggested by the histogram. **[1]**

(iii) State the minimum and maximum possible values of the midrange. **[2]**

2 I have 5 books, each by a different author. The authors are Austen, Brontë, Clarke, Dickens and Eliot.

(i) If I arrange the books in a random order on my bookshelf, find the probability that the authors are in alphabetical order with Austen on the left. **[2]**

(ii) If I choose two of the books at random, find the probability that I choose the books written by Austen and Brontë. **[3]**

3 25% of the plants of a particular species have red flowers. A random sample of 6 plants is selected.

(i) Find the probability that there are no plants with red flowers in the sample. **[2]**

(ii) If 50 random samples of 6 plants are selected, find the expected number of samples in which there are no plants with red flowers. **[2]**

4 Two fair six-sided dice are thrown. The random variable X denotes the difference between the scores on the two dice. The table shows the probability distribution of X.

r	0	1	2	3	4	5
$\mathrm{P}(X = r)$	$\frac{1}{6}$	$\frac{5}{18}$	$\frac{2}{9}$	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$

(i) Draw a vertical line chart to illustrate the probability distribution. **[2]**

(ii) Use a probability argument to show that

(*A*) $\mathrm{P}(X = 1) = \frac{5}{18}$, **[2]**

(*B*) $\mathrm{P}(X = 0) = \frac{1}{6}$. **[1]**

(iii) Find the mean value of X. **[2]**

5 In a recent survey, a large number of working people were asked whether they worked full-time or part-time. Part-time was defined as less than 25 hours per week. One of the respondents is selected at random.

- W is the event that this person works part-time.
- F is the event that this person is female.

You are given that $\mathrm{P}(W) = 0.14$, $\mathrm{P}(F) = 0.41$ and $\mathrm{P}(W \cap F) = 0.11$.

(i) Draw a Venn diagram showing the events W and F, and fill in the probability corresponding to each of the four regions of your diagram. **[3]**

(ii) Determine whether the events W and F are independent. **[2]**

(iii) Find $\mathrm{P}(W \mid F)$ and explain what this probability represents. **[3]**

6 The numbers of eggs laid by a sample of 70 female herring gulls are shown in the table.

Number of eggs	1	2	3	4
Frequency	10	40	15	5

(i) Find the mean and standard deviation of the number of eggs laid per gull. **[5]**

(ii) The sample did not include female herring gulls that laid no eggs.

How would the mean and standard deviation change if these gulls were included? **[2]**

Section B (36 marks)

7 Any patient who fails to turn up for an outpatient appointment at a hospital is described as a 'no-show'. At a particular hospital, on average 15% of patients are no-shows. A random sample of 20 patients who have outpatient appointments is selected.

(i) Find the probability that

(*A*) there is exactly 1 no-show in the sample, **[3]**

(*B*) there are at least 2 no-shows in the sample. **[2]**

The hospital management introduces a policy of telephoning patients before appointments. It is hoped that this will reduce the proportion of no-shows. In order to check this, a random sample of n patients is selected. The number of no-shows in the sample is recorded and a hypothesis test is carried out at the 5% level.

(ii) Write down suitable null and alternative hypotheses for the test. Give a reason for your choice of alternative hypothesis. **[4]**

(iii) In the case that $n = 20$ and the number of no-shows in the sample is 1, carry out the test. **[4]**

(iv) In another case, where n is large, the number of no-shows in the sample is 6 and the critical value for the test is 8. Complete the test. **[3]**

(v) In the case that $n \leqslant 18$, explain why there is no point in carrying out the test at the 5% level. **[2]**

8 The heating quality of the coal in a sample of 50 sacks is measured in suitable units. The data are summarised below.

Heating quality (x)	$9.1 \leqslant x \leqslant 9.3$	$9.3 < x \leqslant 9.5$	$9.5 < x \leqslant 9.7$	$9.7 < x \leqslant 9.9$	$9.9 < x \leqslant 10.1$
Frequency	5	7	15	16	7

(i) Draw a cumulative frequency diagram to illustrate these data. **[5]**

(ii) Use the diagram to estimate the median and interquartile range of the data. **[3]**

(iii) Show that there are no outliers in the sample. **[3]**

(iv) Three of these 50 sacks are selected at random. Find the probability that

(*A*) in all three, the heating quality x is more than 9.5, **[3]**

(*B*) in at least two, the heating quality x is more than 9.5. **[4]**