OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE **4766/01**

MATHEMATICS (MEI)

Statistics 1

FRIDAY 6 JUNE 2008 Afternoon

Time: 1 hour 30 minutes

Additional materials (enclosed): None

Additional materials (required):
Answer Booklet (8 pages)
Graph paper
MEI Examination Formulae and Tables (MF2)

INSTRUCTIONS TO CANDIDATES

- Write your name in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
- Read each question carefully and make sure you know what you have to do before starting your answer.
- 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.
- The total number of marks for this paper is **72**.
- 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.

This document consists of **6** printed pages and **2** blank pages.

[Turn over

Section A (36 marks)

1 In a survey, a sample of 44 fields is selected. Their areas (x hectares) are summarised in the grouped frequency table.

Area (x)	$0 < x \leqslant 3$	$3 < x \leqslant 5$	$5 < x \leqslant 7$	$7 < x \leqslant 10$	$10 < x \leqslant 20$
Frequency	3	8	13	14	6

(i) Calculate an estimate of the sample mean and the sample standard deviation. [4]

(ii) Determine whether there could be any outliers at the upper end of the distribution. [2]

2 In the 2001 census, people living in Wales were asked whether or not they could speak Welsh. A resident of Wales is selected at random.

- W is the event that this person speaks Welsh.
- C is the event that this person is a child.

You are given that $\mathrm{P}(W) = 0.20$, $\mathrm{P}(C) = 0.17$ and $\mathrm{P}(W \cap C) = 0.06$.

(i) Determine whether the events W and C are independent. [2]

(ii) Draw a Venn diagram, showing the events W and C, and fill in the probability corresponding to each region of your diagram. [3]

(iii) Find $\mathrm{P}(W|C)$. [2]

(iv) Given that $\mathrm{P}(W|C') = 0.169$, use this information and your answer to part **(iii)** to comment very briefly on how the ability to speak Welsh differs between children and adults. [1]

3 In a game of darts, a player throws three darts. Let X represent the number of darts which hit the bull's-eye. The probability distribution of X is shown in the table.

r	0	1	2	3
$\mathrm{P}(X = r)$	0.5	0.35	p	q

(i) (A) Show that $p + q = 0.15$. [1]

(B) Given that the expectation of X is 0.67, show that $2p + 3q = 0.32$. [1]

(C) Find the values of p and q. [2]

(ii) Find the variance of X. [3]

4 A small business has 8 workers. On a given day, the probability that any particular worker is off sick is 0.05, independently of the other workers.

(i) A day is selected at random. Find the probability that

(*A*) no workers are off sick, [2]

(*B*) more than one worker is off sick. [3]

(ii) There are 250 working days in a year. Find the expected number of days in the year on which more than one worker is off sick. [2]

5 A psychology student is investigating memory. In an experiment, volunteers are given 30 seconds to try to memorise a number of items. The items are then removed and the volunteers have to try to name all of them. It has been found that the probability that a volunteer names all of the items is 0.35. The student believes that this probability may be increased if the volunteers listen to the same piece of music while memorising the items and while trying to name them.

The student selects 15 volunteers at random to do the experiment while listening to music. Of these volunteers, 8 name all of the items.

(i) Write down suitable hypotheses for a test to determine whether there is any evidence to support the student's belief, giving a reason for your choice of alternative hypothesis. [4]

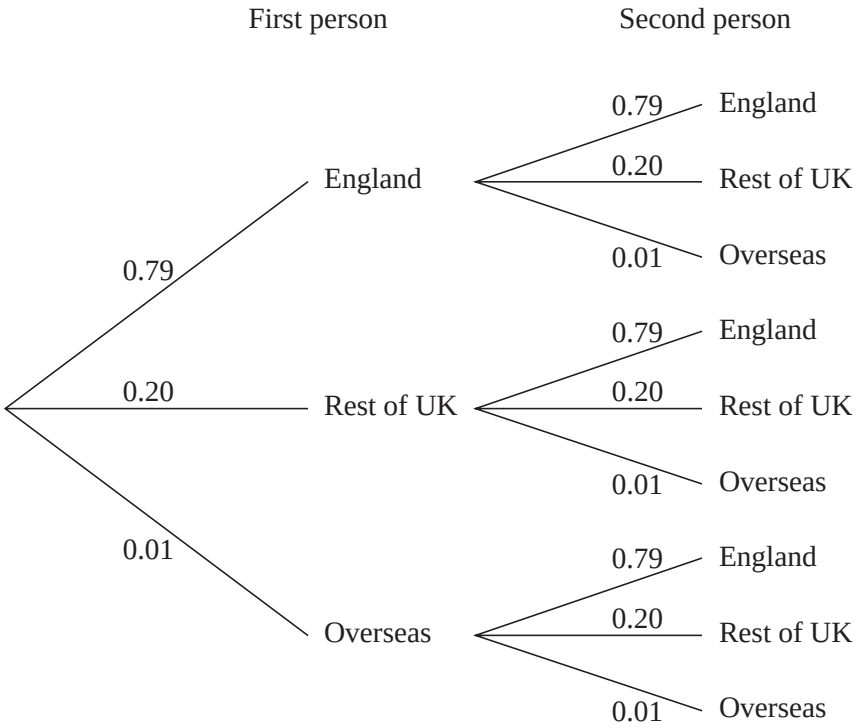
(ii) Carry out the test at the 5% significance level. [4]

[Turn over

Section B (36 marks)

6 In a large town, 79% of the population were born in England, 20% in the rest of the UK and the remaining 1% overseas. Two people are selected at random.

You may use the tree diagram below in answering this question.



(i) Find the probability that

(*A*) both of these people were born in the rest of the UK, [2]

(*B*) at least one of these people was born in England, [3]

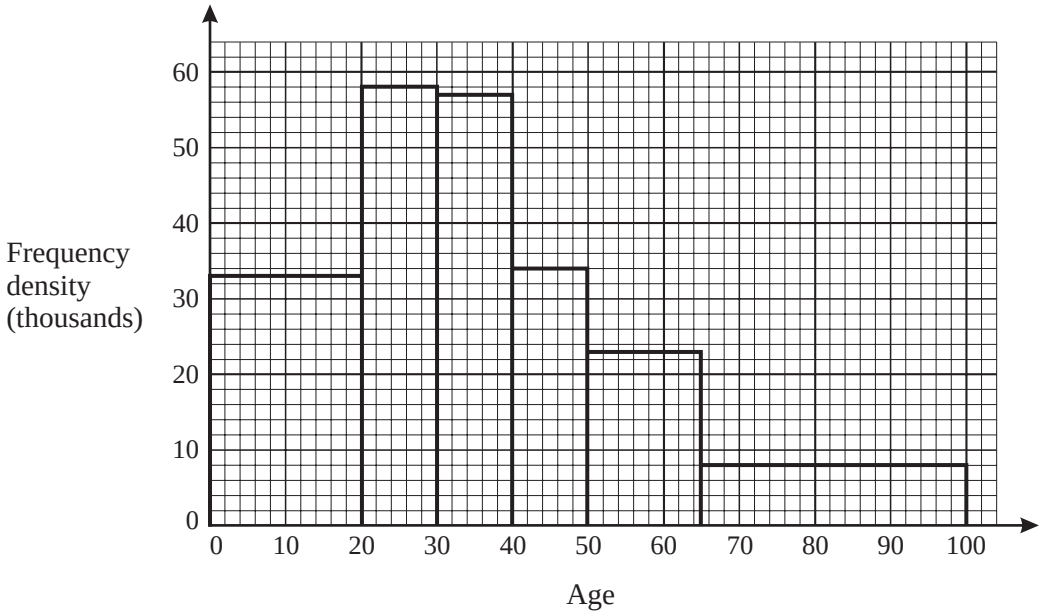
(*C*) neither of these people was born overseas. [2]

(ii) Find the probability that both of these people were born in the rest of the UK given that neither was born overseas. [3]

(iii) (*A*) Five people are selected at random. Find the probability that at least one of them was not born in England. [3]

(*B*) An interviewer selects n people at random. The interviewer wishes to ensure that the probability that at least one of them was not born in England is more than 90%. Find the least possible value of n. You must show working to justify your answer. [3]

7 The histogram shows the age distribution of people living in Inner London in 2001.



Data sourced from the 2001 Census, www.statistics.gov.uk

(i) State the type of skewness shown by the distribution. [1]

(ii) Use the histogram to estimate the number of people aged under 25. [3]

(iii) The table below shows the cumulative frequency distribution.

Age	20	30	40	50	65	100
Cumulative frequency (thousands)	660	1240	1810	a	2490	2770

(*A*) Use the histogram to find the value of a. [2]

(*B*) Use the table to calculate an estimate of the median age of these people. [3]

The ages of people living in Outer London in 2001 are summarised below.

Age (x years)	$0 \leqslant x < 20$	$20 \leqslant x < 30$	$30 \leqslant x < 40$	$40 \leqslant x < 50$	$50 \leqslant x < 65$	$65 \leqslant x < 100$
Frequency (thousands)	1120	650	770	590	680	610

(iv) Illustrate these data by means of a histogram. [5]

(v) Make two brief comments on the differences between the age distributions of the populations of Inner London and Outer London. [2]

(vi) The data given in the table for Outer London are used to calculate the following estimates.

Mean 38.5, median 35.7, midrange 50, standard deviation 23.7, interquartile range 34.4.

The final group in the table assumes that the maximum age of any resident is 100 years. These estimates are to be recalculated, based on a maximum age of 105, rather than 100. For each of the five estimates, state whether it would increase, decrease or be unchanged. [4]