OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE UNIT   **G241/01**

MEI STATISTICS

Statistics 1 (Z1)

TUESDAY 5 JUNE 2007   Afternoon

Time: 1 hour 30 minutes

Additional Materials:
        Answer booklet (8 pages)
        Graph paper
        MEI Examination Formulae and Tables (MF2)

INSTRUCTIONS TO CANDIDATES

- Write your name, centre number and candidate number in the spaces provided on the answer booklet.
- 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.

ADVICE TO CANDIDATES

- Read each question carefully and make sure you know what you have to do before starting your answer.
- 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 **7** printed pages and **1** blank page.

Section A (36 marks)

1 A girl is choosing tracks from an album to play at her birthday party. The album has 8 tracks and she selects 4 of them.

    (i) In how many ways can she select the 4 tracks? [2]

    (ii) In how many different orders can she arrange the 4 tracks once she has chosen them? [1]

2 The histogram shows the amount of money, in pounds, spent by the customers at a supermarket on a particular day.
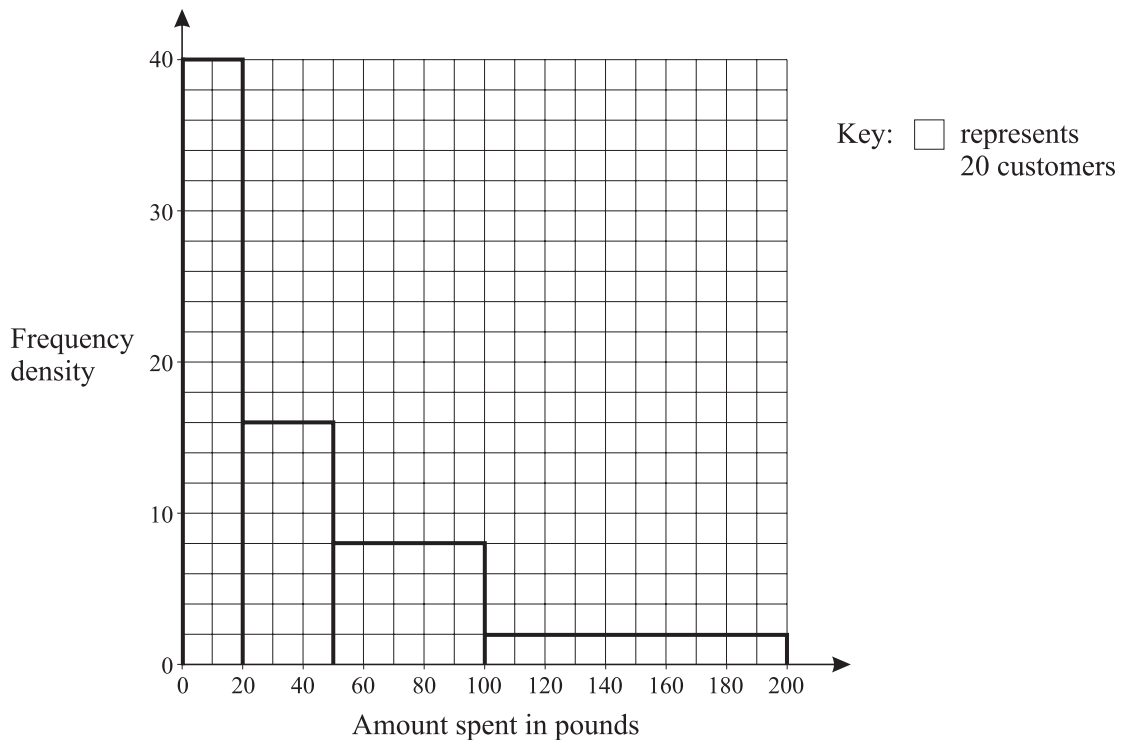
    (i) Express the data in the form of a grouped frequency table. [2]

    (ii) Use your table to estimate the total amount of money spent by customers on that day. [2]

3 The marks x scored by a sample of 56 students in an examination are summarised by

$$n = 56, \quad \Sigma x = 3026, \quad \Sigma x^2 = 178\,890.$$

    (i) Calculate the mean and standard deviation of the marks. [3]

    (ii) The highest mark scored by any of the 56 students in the examination was 93. Show that this result may be considered to be an outlier. [2]

    (iii) The formula $y = 1.2x - 10$ is used to scale the marks. Find the mean and standard deviation of the scaled marks. [3]

4 A local council has introduced a recycling scheme for aluminium, paper and kitchen waste. 50 residents are asked which of these materials they recycle. The numbers of people who recycle each type of material are shown in the Venn diagram.
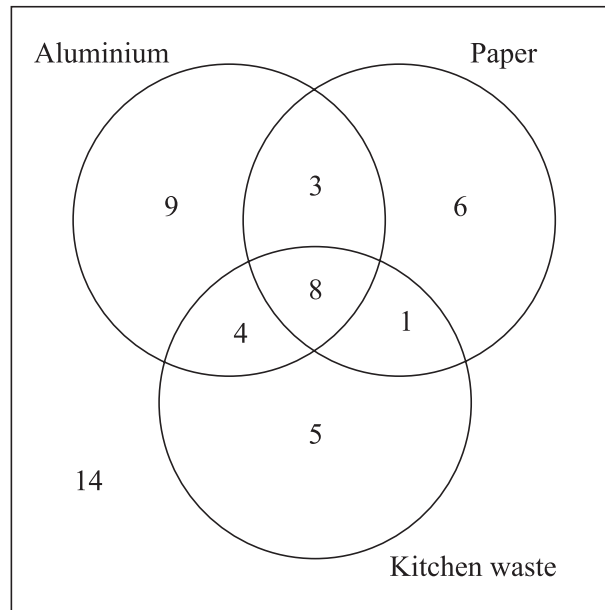
One of the residents is selected at random.

  (i) Find the probability that this resident recycles

     (*A*) at least one of the materials, [1]

     (*B*) exactly one of the materials. [2]

 (ii) Given that the resident recycles aluminium, find the probability that this resident does not recycle paper. [2]

Two residents are selected at random.

(iii) Find the probability that exactly one of them recycles kitchen waste. [3]

5 A GCSE geography student is investigating a claim that global warming is causing summers in Britain to have more rainfall. He collects rainfall data from a local weather station for 2001 and 2006. The vertical line chart shows the number of days per week on which some rainfall was recorded during the 22 weeks of summer 2001.

Number of days per week with rain recorded in summer 2001
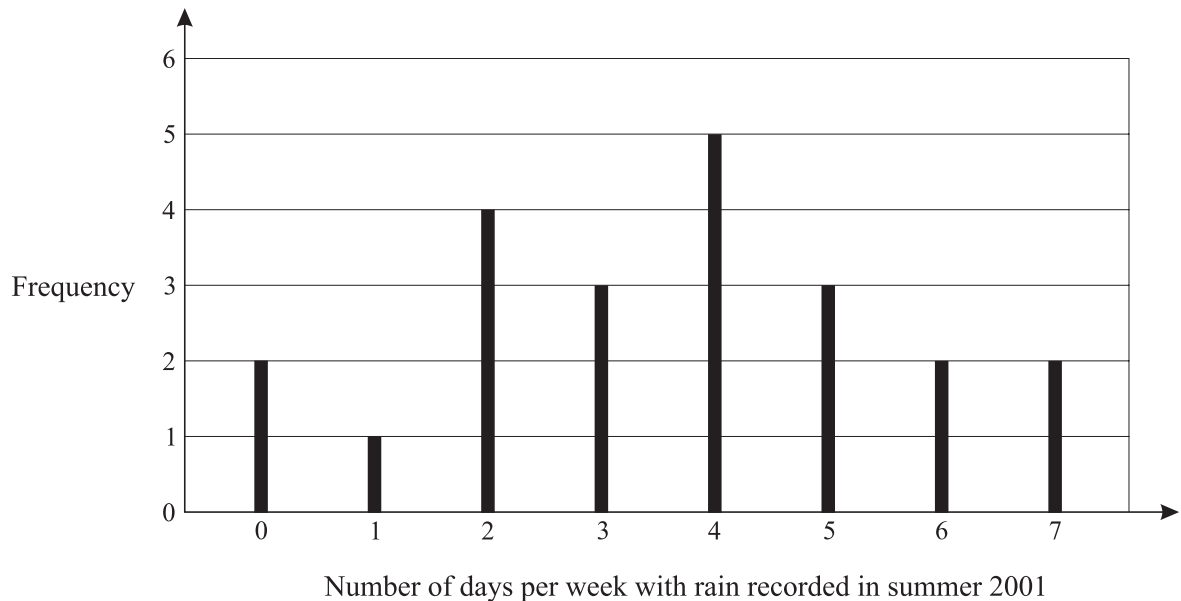
   (i) Show that the median of the data is 4, and find the interquartile range. [3]

   (ii) For summer 2006 the median is 3 and the interquartile range is also 3. The student concludes that the data demonstrate that global warming is causing summer rainfall to decrease rather than increase. Is this a valid conclusion from the data? Give two brief reasons to justify your answer. [3]

6 In a phone-in competition run by a local radio station, listeners are given the names of 7 local personalities and are told that 4 of them are in the studio. Competitors phone in and guess which 4 are in the studio.

   (i) Show that the probability that a randomly selected competitor guesses all 4 correctly is $\frac{1}{35}$. [2]

   Let X represent the number of correct guesses made by a randomly selected competitor. The probability distribution of X is shown in the table.

r	0	1	2	3	4
$P(X = r)$	0	$\frac{4}{35}$	$\frac{18}{35}$	$\frac{12}{35}$	$\frac{1}{35}$

   **(ii)** Find the expectation and variance of X. [5]

Section B (36 marks)

7   A screening test for a particular disease is applied to everyone in a large population. The test classifies people into three groups: 'positive', 'doubtful' and 'negative'. Of the population, 3% is classified as positive, 6% as doubtful and the rest negative.

In fact, of the people who test positive, only 95% have the disease. Of the people who test doubtful, 10% have the disease. Of the people who test negative, 1% actually have the disease.

People who do not have the disease are described as 'clear'.

   (i) Copy and complete the tree diagram to show this information.   [4]
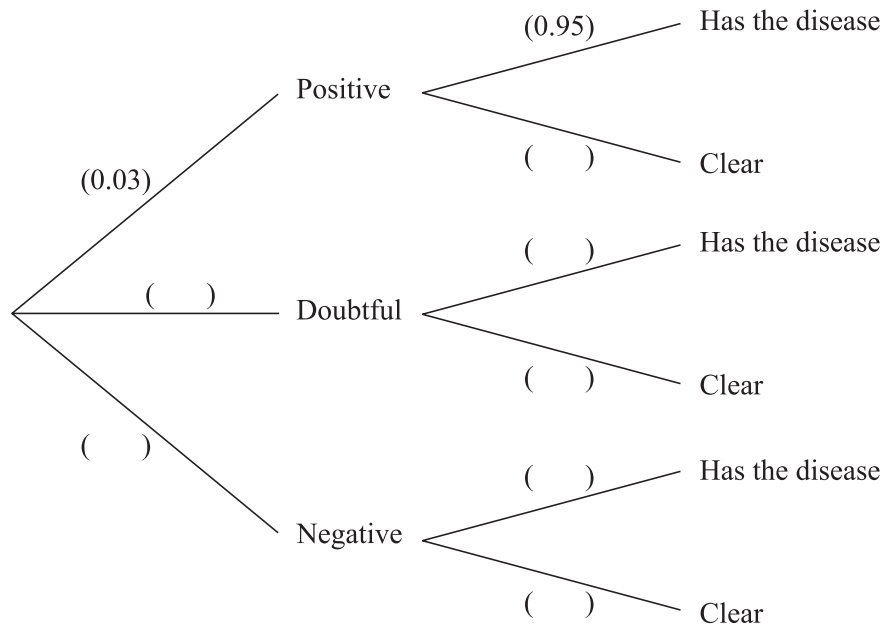
   (ii) Find the probability that a randomly selected person tests negative and is clear.   [2]

   (iii) Find the probability that a randomly selected person has the disease.   [3]

   (iv) Find the probability that a randomly selected person tests negative **given** that the person has the disease.   [3]

   (v) Comment briefly on what your answer to part **(iv)** indicates about the effectiveness of the screening test.   [2]

Once the test has been carried out, those people who test doubtful are given a detailed medical examination. If a person has the disease the examination will correctly identify this in 98% of cases. If a person is clear, the examination will always correctly identify this.

   (vi) A person is selected at random. Find the probability that this person either tests negative originally or tests doubtful and is then cleared in the detailed medical examination.   [4]

8 A multinational accountancy firm receives a large number of job applications from graduates each year. On average 20% of applicants are successful.

A researcher in the human resources department of the firm selects a random sample of 17 graduate applicants.

  (i) Find the probability that at least 4 of the 17 applicants are successful. [3]

 (ii) Find the expected number of successful applicants in the sample. [2]

(iii) Find the most likely number of successful applicants in the sample, justifying your answer. [3]

It is suggested that mathematics graduates are more likely to be successful than those from other fields. In order to test this suggestion, the researcher decides to select a new random sample of 17 mathematics graduate applicants. The researcher then carries out a hypothesis test at the 5% significance level.

(iv) (*A*) Write down suitable null and alternative hypotheses for the test.

     (*B*) Give a reason for your choice of the alternative hypothesis. [4]

 (v) Find the critical region for the test at the 5% level, showing all of your calculations. [4]

(vi) Explain why the critical region found in part **(v)** would be unaltered if a 10% significance level were used. [2]

ADVANCED SUBSIDIARY GCE   **G241/01**

MEI STATISTICS

Statistics 1 (Z1)

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.

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 $P(W) = 0.20$, $P(C) = 0.17$ and $P(W \cap C) = 0.06$.

   (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 $P(W \mid C)$. [2]

   (iv) Given that $P(W \mid C') = 0.169$, 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
$P(X = r)$	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]

4

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.
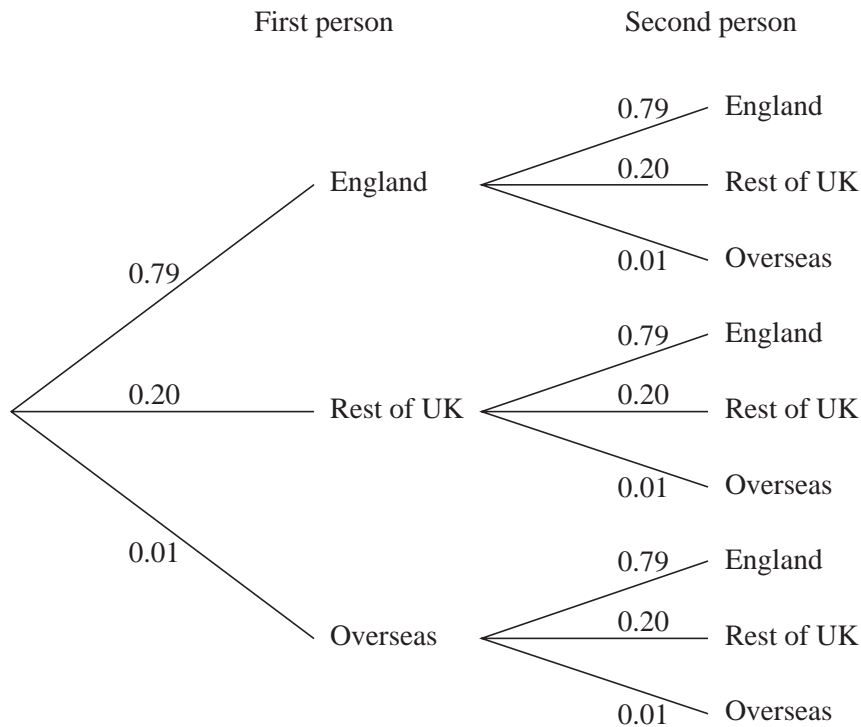
   (i) Find the probability that

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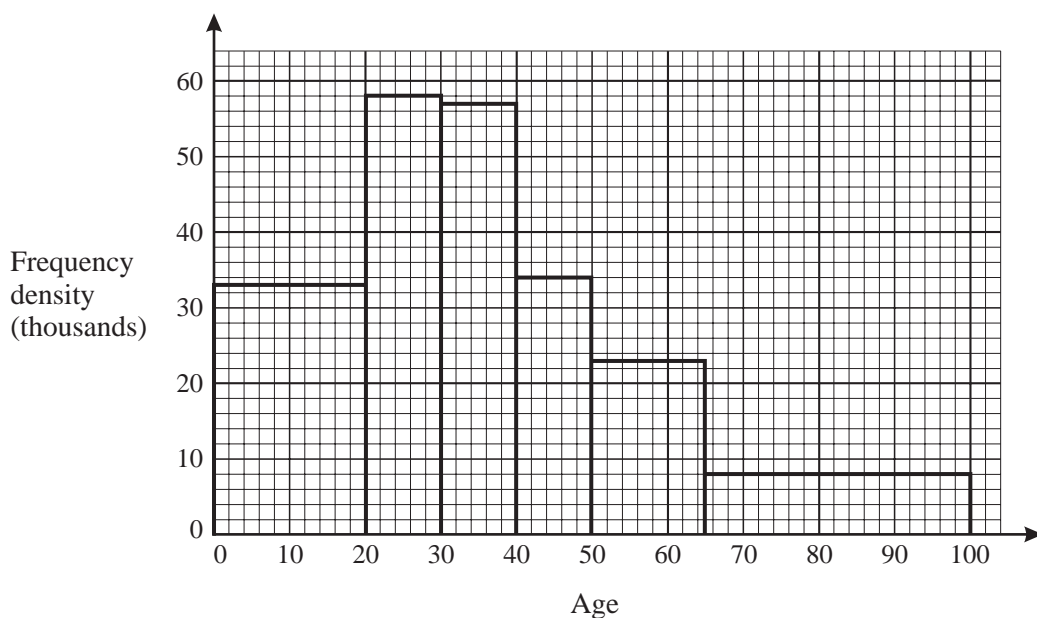
Data sourced from the 2001 Census,
www.statistics.gov.uk

  (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]

ADVANCED SUBSIDIARY GCE

MEI STATISTICS

Statistics 1 (Z1)

G241

Monday 15 June 2009
Afternoon

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

* Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
* Use black ink. Pencil may be used for graphs and diagrams only.
* Read each question carefully and make sure that 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 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.
* 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 document consists of **8** pages. Any blank pages are indicated.

Section A (36 marks)

1   In a traffic survey, the number of people in each car passing the survey point is recorded. The results are given in the following frequency table.

Number of people	1	2	3	4
Frequency	50	31	16	5

    (i) Write down the median and mode of these data.   **[2]**

    (ii) Draw a vertical line diagram for these data.   **[2]**

    (iii) State the type of skewness of the distribution.   **[1]**

2   There are 14 girls and 11 boys in a class. A quiz team of 5 students is to be chosen from the class.

    (i) How many different teams are possible?   **[2]**

    (ii) If the team must include 3 girls and 2 boys, find how many different teams are possible.   **[3]**

3   Dwayne is a car salesman. The numbers of cars, x, sold by Dwayne each month during the year 2008 are summarised by

$$n = 12, \qquad \Sigma x = 126, \qquad \Sigma x^2 = 1582.$$

    (i) Calculate the mean and standard deviation of the monthly numbers of cars sold.   **[3]**

    (ii) Dwayne earns £500 each month plus £100 commission for each car sold. Show that the mean of Dwayne's monthly earnings is £1550. Find the standard deviation of Dwayne's monthly earnings.   **[3]**

    (iii) Marlene is a car saleswoman and is paid in the same way as Dwayne. During 2008 her monthly earnings have mean £1625 and standard deviation £280. Briefly compare the monthly numbers of cars sold by Marlene and Dwayne during 2008.   **[2]**

4   The table shows the probability distribution of the random variable X.

r	10	20	30	40
P($X = r$)	0.2	0.3	0.3	0.2

    (i) Explain why E(X) = 25.   **[1]**

    (ii) Calculate Var(X).   **[3]**

5 The frequency table below shows the distance travelled by 1200 visitors to a particular UK tourist destination in August 2008.

| Distance (d miles)	$0 \leqslant d < 50$	$50 \leqslant d < 100$	$100 \leqslant d < 200$	$200 \leqslant d < 400$
Frequency	360	400	307	133

    (i) Draw a histogram on graph paper to illustrate these data. **[5]**

    (ii) Calculate an estimate of the median distance. **[3]**

6 Whitefly, blight and mosaic virus are three problems which can affect tomato plants. 100 tomato plants are examined for these problems. The numbers of plants with each type of problem are shown in the Venn diagram. 47 of the plants have none of the problems.
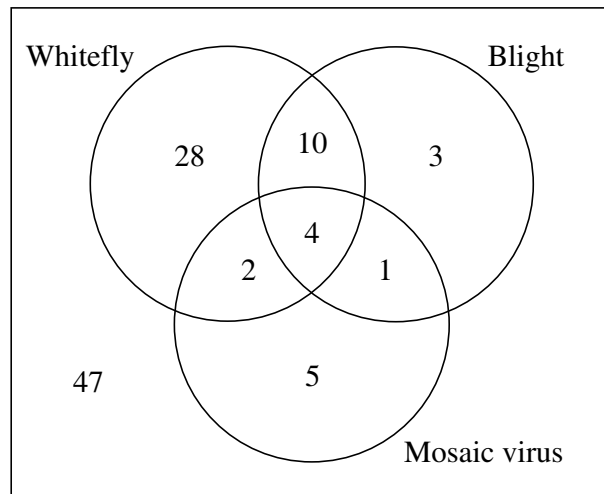
    (i) One of the 100 plants is selected at random. Find the probability that this plant has

        (*A*) at most one of the problems, **[1]**

        (*B*) exactly two of the problems. **[2]**

    (ii) Three of the 100 plants are selected at random. Find the probability that all of them have at least one of the problems. **[3]**

Section B (36 marks)

7   Laura frequently flies to business meetings and often finds that her flights are delayed. A flight may be delayed due to technical problems, weather problems or congestion problems, with probabilities 0.2, 0.15 and 0.1 respectively. The tree diagram shows this information.
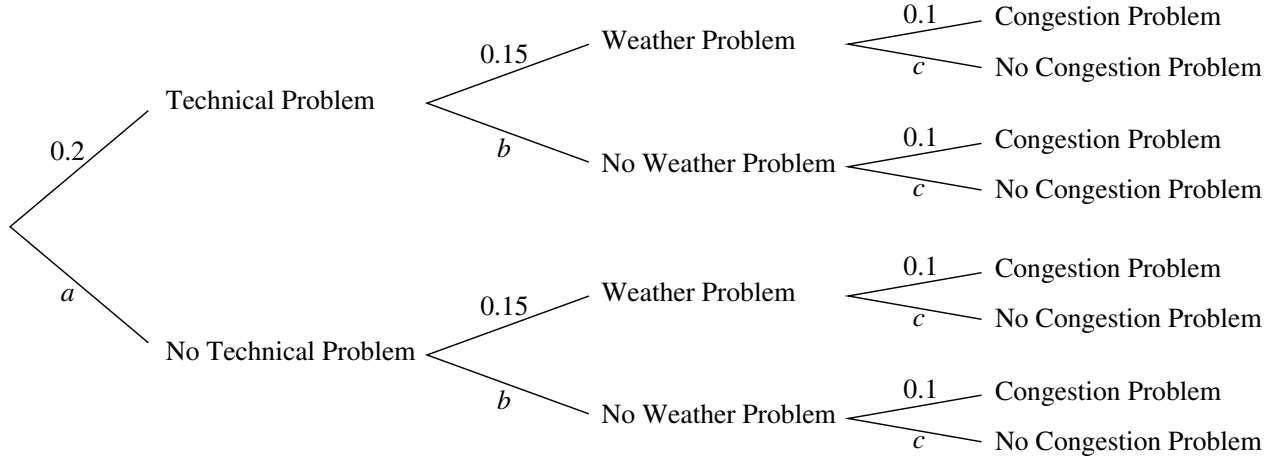
(i)   Write down the values of the probabilities *a*, *b* and *c* shown in the tree diagram.   **[2]**

One of Laura's flights is selected at random.

(ii)   Find the probability that Laura's flight is not delayed and hence write down the probability that it is delayed.   **[4]**

(iii)   Find the probability that Laura's flight is delayed due to just one of the three problems.   **[4]**

(iv)   Given that Laura's flight is delayed, find the probability that the delay is due to just one of the three problems.   **[3]**

(v)   Given that Laura's flight has no technical problems, find the probability that it is delayed.   **[3]**

(vi)   In a particular year, Laura has 110 flights. Find the expected number of flights that are delayed.   **[2]**

8 The Department of Health 'eat five a day' advice recommends that people should eat at least five portions of fruit and vegetables per day. In a particular school, 20% of pupils eat at least five a day.

    (i) 15 children are selected at random.

        (*A*) Find the probability that exactly 3 of them eat at least five a day. **[3]**

        (*B*) Find the probability that at least 3 of them eat at least five a day. **[3]**

        (*C*) Find the expected number who eat at least five a day. **[2]**

A programme is introduced to encourage children to eat more portions of fruit and vegetables per day. At the end of this programme, the diets of a random sample of 15 children are analysed. A hypothesis test is carried out to examine whether the proportion of children in the school who eat at least five a day has increased.

    (ii) (*A*) Write down suitable null and alternative hypotheses for the test.

          (*B*) Give a reason for your choice of the alternative hypothesis. **[4]**

    (iii) Find the critical region for the test at the 10% significance level, showing all of your calculations. Hence complete the test, given that 7 of the 15 children eat at least five a day. **[6]**

OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE

MEI STATISTICS

Statistics 1 (Z1)

G241

Candidates answer on the Answer Booklet

OCR Supplied Materials:
* 8 page Answer Booklet
* Graph paper
* MEI Examination Formulae and Tables (MF2)

Other Materials Required:
None

Monday 25 January 2010
Morning

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

* Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
* Use black ink. Pencil may be used for graphs and diagrams only.
* Read each question carefully and make sure that 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 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.
* 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 document consists of **8** pages. Any blank pages are indicated.

Section A (36 marks)

1   A camera records the speeds in miles per hour of 15 vehicles on a motorway. The speeds are given below.

73   67   75   64   52   63   75   81   77   72   68   74   79   72   71

   (i) Construct a sorted stem and leaf diagram to represent these data, taking stem values of 50, 60, … . **[4]**

   (ii) Write down the median and midrange of the data. **[2]**

   (iii) Which of the median and midrange would you recommend to measure the central tendency of the data? Briefly explain your answer. **[2]**

2   In her purse, Katharine has two £5 notes, two £10 notes and one £20 note. She decides to select two of these notes at random to donate to a charity. The total value of these two notes is denoted by the random variable £X.

   (i) (*A*) Show that $P(X = 10) = 0.1$. **[1]**

   (*B*) Show that $P(X = 30) = 0.2$. **[2]**

The table shows the probability distribution of X.

r	10	15	20	25	30
$P(X = r)$	0.1	0.4	0.1	0.2	0.2

   (ii) Find $E(X)$ and $Var(X)$. **[5]**

3   In a survey, a large number of young people are asked about their exercise habits. One of these people is selected at random.

   • G is the event that this person goes to the gym.

   • R is the event that this person goes running.

You are given that $P(G) = 0.24$, $P(R) = 0.13$ and $P(G \cap R) = 0.06$.

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

   (ii) Determine whether the events G and R are independent. **[2]**

   (iii) Find $P(R \mid G)$. **[3]**

4   In a multiple-choice test there are 30 questions. For each question, there is a 60% chance that a randomly selected student answers correctly, independently of all other questions.

    (i) Find the probability that a randomly selected student gets a total of exactly 20 questions correct. **[3]**

    (ii) If 100 randomly selected students take the test, find the expected number of students who get exactly 20 questions correct. **[2]**

5   My credit card has a 4-digit code called a PIN. You should assume that any 4-digit number from 0000 to 9999 can be a PIN.

    (i) If I cannot remember any digits and guess my number, find the probability that I guess it correctly. **[1]**

In fact my PIN consists of four different digits. I can remember all four digits, but cannot remember the correct order.

    (ii) If I now guess my number, find the probability that I guess it correctly. **[2]**

6   Three prizes, one for English, one for French and one for Spanish, are to be awarded in a class of 20 students.
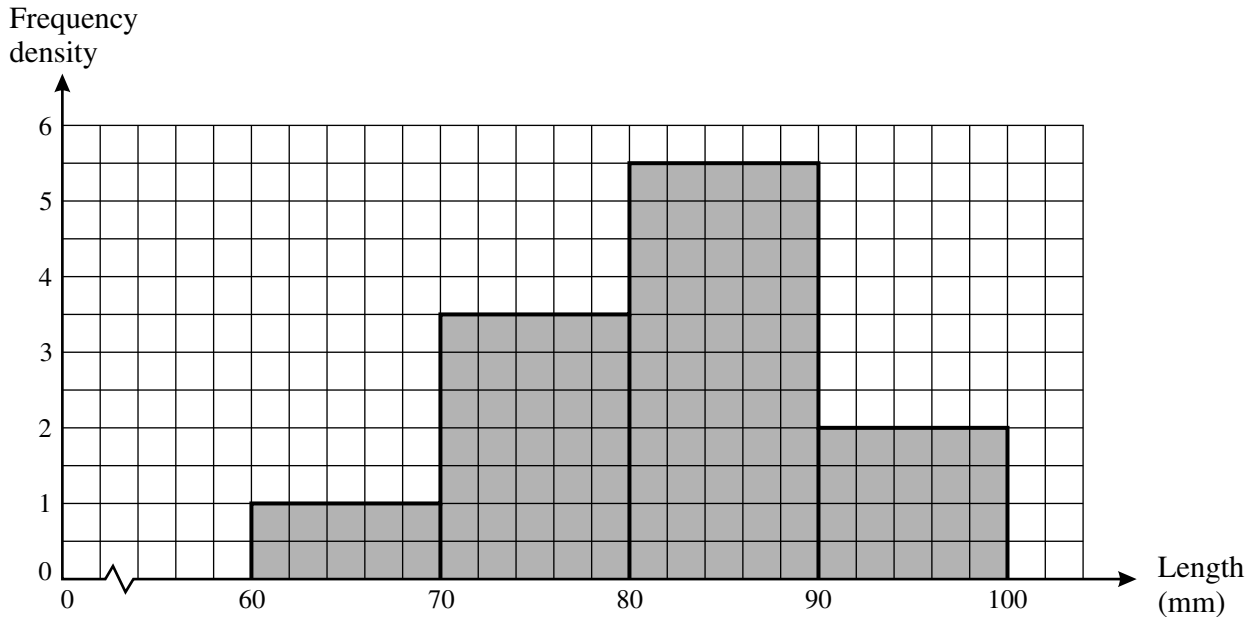
Find the number of different ways in which the three prizes can be awarded if

    (i) no student may win more than 1 prize, **[2]**

    (ii) no student may win all 3 prizes. **[2]**

4

Section B (36 marks)

7 A pear grower collects a random sample of 120 pears from his orchard. The histogram below shows the lengths, in mm, of these pears.

  (i) Calculate the number of pears which are between 90 and 100 mm long. **[2]**

  (ii) Calculate an estimate of the mean length of the pears. Explain why your answer is only an estimate. **[4]**

  (iii) Calculate an estimate of the standard deviation. **[3]**

  (iv) Use your answers to parts **(ii)** and **(iii)** to investigate whether there are any outliers. **[4]**

  (v) Name the type of skewness of the distribution. **[1]**

  (vi) Illustrate the data using a cumulative frequency diagram. **[5]**

8 An environmental health officer monitors the air pollution level in a city street. Each day the level of pollution is classified as low, medium or high. The probabilities of each level of pollution on a randomly chosen day are as given in the table.

Pollution level	Low	Medium	High
Probability	0.5	0.35	0.15

  (i) Three days are chosen at random. Find the probability that the pollution level is

     (*A*) low on all 3 days, **[2]**

     (*B*) low on at least one day, **[2]**

     (*C*) low on one day, medium on another day, and high on the other day. **[3]**

 (ii) Ten days are chosen at random. Find the probability that

     (*A*) there are no days when the pollution level is high, **[2]**

     (*B*) there is exactly one day when the pollution level is high. **[3]**

The environmental health officer believes that pollution levels will be low more frequently in a different street. On 20 randomly selected days she monitors the pollution level in this street and finds that it is low on 15 occasions.

(iii) Carry out a test at the 5% level to determine if there is evidence to suggest that she is correct. Use hypotheses $H_0: p = 0.5$, $H_1: p > 0.5$, where p represents the probability that the pollution level in this street is low. Explain why H_1 has this form. **[5]**

ADVANCED SUBSIDIARY GCE

MEI STATISTICS

Statistics 1 (Z1)

G241

QUESTION PAPER

Candidates answer on the Printed Answer Book

OCR Supplied Materials:
*   Printed Answer Book G241
*   MEI Examination Formulae and Tables (MF2)

Other Materials Required:
*   Scientific or graphical calculator

Friday 18 June 2010
Afternoon

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

*   Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Printed Answer Book.
*   **The questions are on the inserted Question Paper.**
*   **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 and make sure that 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 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   A business analyst collects data about the distribution of hourly wages, in £, of shop-floor workers at a factory. These data are illustrated in the box and whisker plot.
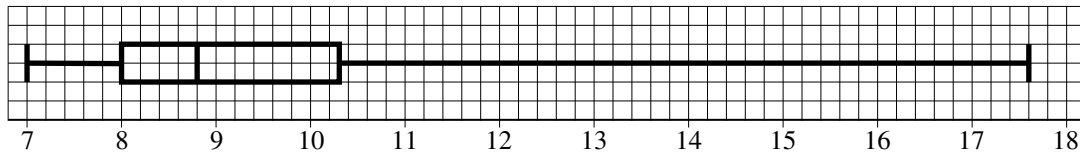
   (i) Name the type of skewness of the distribution. **[1]**

   (ii) Find the interquartile range and hence show that there are no outliers at the lower end of the distribution, but there is at least one outlier at the upper end. **[5]**

   (iii) Suggest possible reasons why this may be the case. **[2]**

2   The probability distribution of the random variable X is given by the formula

$$P(X = r) = kr(5 - r) \text{ for } r = 1, 2, 3, 4.$$

   (i) Show that $k = 0.05$. **[2]**

   (ii) Find $E(X)$ and $Var(X)$. **[5]**

3   The lifetimes in hours of 90 components are summarised in the table.

Lifetime (x hours)	$0 < x \leqslant 20$	$20 < x \leqslant 30$	$30 < x \leqslant 50$	$50 < x \leqslant 65$	$65 < x \leqslant 100$
Frequency	24	13	14	21	18

   (i) Draw a histogram to illustrate these data. **[5]**

   (ii) In which class interval does the median lie? Justify your answer. **[2]**

4   Each packet of *Cruncho* cereal contains one free fridge magnet. There are five different types of fridge magnet to collect. They are distributed, with equal probability, randomly and independently in the packets. Keith is about to start collecting these fridge magnets.

   (i) Find the probability that the first 2 packets that Keith buys contain the same type of fridge magnet. **[2]**

   (ii) Find the probability that Keith collects all five types of fridge magnet by buying just 5 packets. **[3]**

   (iii) Hence find the probability that Keith has to buy more than 5 packets to acquire a complete set. **[1]**

5   A retail analyst records the numbers of loaves of bread of a particular type bought by a sample of shoppers in a supermarket.

Number of loaves	0	1	2	3	4	5
Frequency	37	23	11	3	0	1

   (i) Calculate the mean and standard deviation of the numbers of loaves bought per person.   **[5]**

   (ii) Each loaf costs £1.04. Calculate the mean and standard deviation of the amount spent on loaves per person.   **[3]**

Section B (36 marks)

6   A manufacturer produces tiles. On average 10% of the tiles produced are faulty. Faulty tiles occur randomly and independently.

A random sample of 18 tiles is selected.

   (i) (*A*) Find the probability that there are exactly 2 faulty tiles in the sample.   **[3]**

   (*B*) Find the probability that there are more than 2 faulty tiles in the sample.   **[3]**

   (*C*) Find the expected number of faulty tiles in the sample.   **[2]**

A cheaper way of producing the tiles is introduced. The manufacturer believes that this may increase the proportion of faulty tiles. In order to check this, a random sample of 18 tiles produced using the cheaper process is selected and a hypothesis test is carried out.

   (ii) (*A*) Write down suitable null and alternative hypotheses for the test.

   (*B*) Explain why the alternative hypothesis has the form that it does.   **[4]**

   (iii) Find the critical region for the test at the 5% level, showing all of your calculations.   **[4]**

   (iv) In fact there are 4 faulty tiles in the sample. Complete the test, stating your conclusion clearly.   **[2]**

7   One train leaves a station each hour. The train is either on time or late. If the train is on time, the probability that the next train is on time is 0.95. If the train is late, the probability that the next train is on time is 0.6. On a particular day, the 09 00 train is on time.

   (i) Illustrate the possible outcomes for the 10 00, 11 00 and 12 00 trains on a probability tree diagram.   **[4]**

   (ii) Find the probability that

   (*A*) all three of these trains are on time,   **[2]**

   (*B*) just one of these three trains is on time,   **[4]**

   (*C*) the 12 00 train is on time.   **[4]**

   (iii) Given that the 12 00 train is on time, find the probability that the 10 00 train is also on time.   **[4]**

ADVANCED SUBSIDIARY GCE

MEI STATISTICS

G241

Statistics 1 (Z1)

QUESTION PAPER

Candidates answer on the printed answer book.

OCR supplied materials:
• Printed answer book G241
• MEI Examination Formulae and Tables (MF2)

Other materials required:
• Scientific or graphical calculator

Monday 24 January 2011
Morning

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. 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 **8** 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.

OCR is an exempt Charity

Turn over

Section A (36 marks)

1   The stem and leaf diagram shows the weights, rounded to the nearest 10 grams, of 25 female iguanas.

$$
\begin{array}{r|llllllll}
8 & 3 & 9 \\
9 & 3 & 5 & 6 & 6 & 6 & 8 & 9 & 9 \\
10 & 0 & 2 & 2 & 3 & 4 & 6 & 9 \\
11 & 2 & 4 & 7 & 8 \\
12 & 3 & 4 & 5 \\
13 & 2
\end{array}
$$

Key: 11 │ 2   represents 1120 grams

   (i) Find the mode and the median of the data.   **[2]**

   (ii) Identify the type of skewness of the distribution.   **[1]**

2   The table shows all the possible products of the scores on two fair four-sided dice.

		Score on second die			
		1	2	3	4
Score on first die	1	1	2	3	4
	2	2	4	6	8
	3	3	6	9	12
	4	4	8	12	16

   (i) Find the probability that the product of the two scores is less than 10.   **[1]**

   (ii) Show that the events 'the score on the first die is even' and 'the product of the scores on the two dice is less than 10' are not independent.   **[3]**

3   There are 13 men and 10 women in a running club. A committee of 3 men and 3 women is to be selected.

   (i) In how many different ways can the three men be selected?   **[2]**

   (ii) In how many different ways can the whole committee be selected?   **[2]**

   (iii) A random sample of 6 people is selected from the running club. Find the probability that this sample consists of 3 men and 3 women.   **[2]**

4 The probability distribution of the random variable X is given by the formula

$$P(X = r) = kr(r + 1) \quad \text{for } r = 1, 2, 3, 4, 5.$$

   (i) Show that $k = \frac{1}{70}$. **[2]**

   (ii) Find E(X) and Var(X). **[5]**

5 Andy can walk to work, travel by bike or travel by bus. The tree diagram shows the probabilities of any day being dry or wet and the corresponding probabilities for each of Andy's methods of travel.
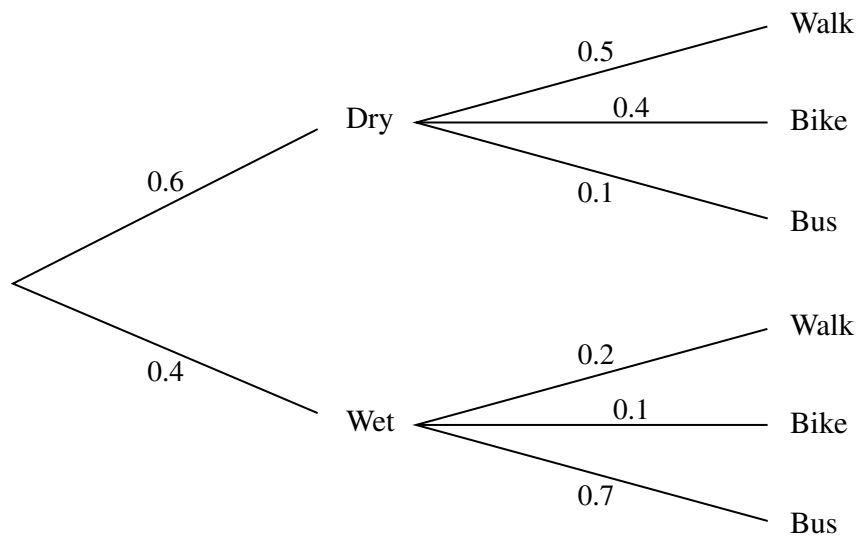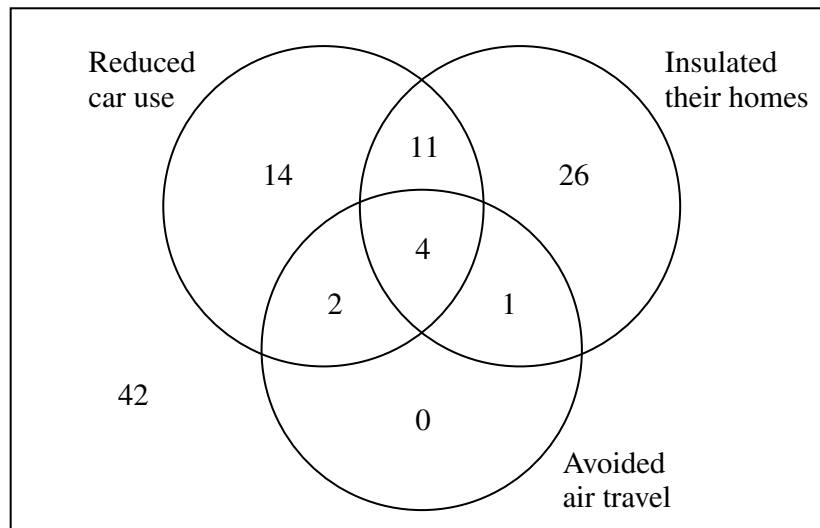

A day is selected at random. Find the probability that

   (i) the weather is wet and Andy travels by bus, **[2]**

   (ii) Andy walks or travels by bike, **[3]**

   (iii) the weather is dry given that Andy walks or travels by bike. **[3]**

6 A survey is being carried out into the carbon footprint of individual citizens. As part of the survey, 100 citizens are asked whether they have attempted to reduce their carbon footprint by any of the following methods.

- Reducing car use
- Insulating their homes
- Avoiding air travel

The numbers of citizens who have used each of these methods are shown in the Venn diagram.

One of the citizens is selected at random.

(i) Find the probability that this citizen

    (*A*) has avoided air travel, **[1]**

    (*B*) has used at least two of the three methods. **[2]**

(ii) Given that the citizen has avoided air travel, find the probability that this citizen has reduced car use. **[2]**

Three of the citizens are selected at random.

(iii) Find the probability that none of them have avoided air travel. **[3]**

Section B (36 marks)

7 The incomes of a sample of 918 households on an island are given in the table below.

Income (x thousand pounds)	$0 \leqslant x \leqslant 20$	$20 < x \leqslant 40$	$40 < x \leqslant 60$	$60 < x \leqslant 100$	$100 < x \leqslant 200$
Frequency	238	365	142	128	45

   (i) Draw a histogram to illustrate the data. **[5]**

  (ii) Calculate an estimate of the mean income. **[3]**

 (iii) Calculate an estimate of the standard deviation of the incomes. **[4]**

 (iv) Use your answers to parts **(ii)** and **(iii)** to show there are almost certainly some outliers in the sample. Explain whether or not it would be appropriate to exclude the outliers from the calculation of the mean and the standard deviation. **[4]**

  (v) The incomes were converted into another currency using the formula $y = 1.15x$. Calculate estimates of the mean and variance of the incomes in the new currency. **[3]**

8 Mark is playing solitaire on his computer. The probability that he wins a game is 0.2, independently of all other games that he plays.

   (i) Find the expected number of wins in 12 games. **[2]**

  (ii) Find the probability that

     (*A*) he wins exactly 2 out of the next 12 games that he plays, **[3]**

     (*B*) he wins at least 2 out of the next 12 games that he plays. **[3]**

 (iii) Mark's friend Ali also plays solitaire. Ali claims that he is better at winning games than Mark. In a random sample of 20 games played by Ali, he wins 7 of them. Write down suitable hypotheses for a test at the 5% level to investigate whether Ali is correct. Give a reason for your choice of alternative hypothesis. Carry out the test. **[9]**

ADVANCED SUBSIDIARY GCE

MEI STATISTICS **G241**

Statistics 1 (Z1)

QUESTION PAPER

Candidates answer on the printed answer book.

OCR supplied materials:
*   Printed answer book G241
*   MEI Examination Formulae and Tables (MF2)

Other materials required:
*   Scientific or graphical calculator

Thursday 26 May 2011
Morning

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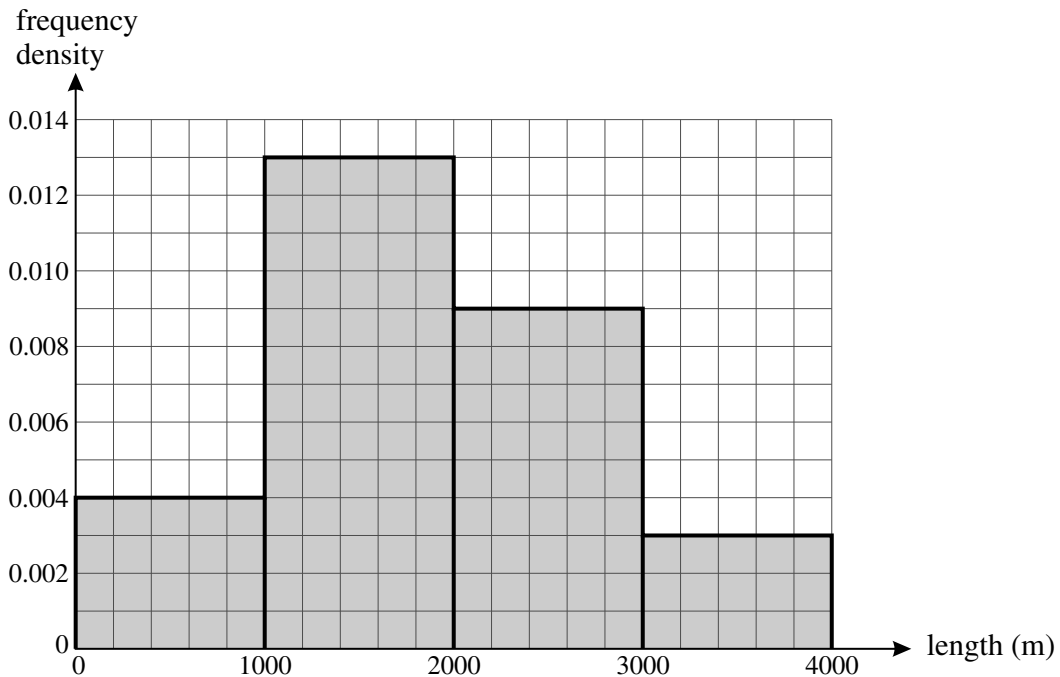
   (i) Find the number of sections which are between 1000 and 2000 metres in length.   **[2]**

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

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

  (ii) Use a probability argument to show that

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

     (B) $P(X = 0) = \frac{1}{6}$. **[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, with part-time being 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 $P(W) = 0.14$, $P(F) = 0.41$ and $P(W \cap F) = 0.11$.

   (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 $P(W \mid F)$ 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* ≤ 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]**

OCR
RECOGNISING ACHIEVEMENT

Thursday 24 May 2012 – Morning

AS GCE MEI STATISTICS

G241 Statistics 1 (Z1)

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:
* Printed Answer Book G241
* 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**.
* 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.

Turn over

Section A (36 marks)

1   At a garden centre there is a box containing 50 hyacinth bulbs. Of these, 30 will produce a blue flower and the remaining 20 will produce a red flower. Unfortunately they have become mixed together so that it is not known which of the bulbs will produce a blue flower and which will produce a red flower.

Karen buys 3 of these bulbs.

   (i) Find the probability that all 3 of these bulbs will produce blue flowers.   [3]

   (ii) Find the probability that Karen will have at least one flower of each colour from her 3 bulbs.   [3]

2   An examination paper consists of two sections. Section A has 5 questions and Section B has 9 questions. Candidates are required to answer 6 questions.

   (i) In how many different ways can a candidate choose 6 questions, if 3 are from Section A and 3 are from Section B?   [3]

   (ii) Another candidate randomly chooses 6 questions to answer. Find the probability that this candidate chooses 3 questions from each section.   [3]

3   At a call centre, 85% of callers are put on hold before being connected to an operator. A random sample of 30 callers is selected.

   (i) Find the probability that exactly 29 of these callers are put on hold.   [3]

   (ii) Find the probability that at least 29 of these callers are put on hold.   [3]

   (iii) If 10 random samples, each of 30 callers, are selected, find the expected number of samples in which at least 29 callers are put on hold.   [2]

4   It is known that 8% of the population of a large city use a particular web browser. A researcher wishes to interview some people from the city who use this browser. He selects people at random, one at a time.

   (i) Find the probability that the first person that he finds who uses this browser is

      (A) the third person selected,   [3]

      (B) the second or third person selected.   [2]

   (ii) Find the probability that at least one of the first 20 people selected uses this browser.   [3]

5   A manufacturer produces titanium bicycle frames. The bicycle frames are tested before use and on average 5% of them are found to be faulty. A cheaper manufacturing process is introduced and the manufacturer wishes to check whether the proportion of faulty bicycle frames has increased. A random sample of 18 bicycle frames is selected and it is found that 4 of them are faulty. Carry out a hypothesis test at the 5% significance level to investigate whether the proportion of faulty bicycle frames has increased.   [8]

Section B (36 marks)

6 The engine sizes $x\,\mathrm{cm}^3$ of a sample of 80 cars are summarised in the table below.

Engine size x	$500 \leqslant x \leqslant 1000$	$1000 < x \leqslant 1500$	$1500 < x \leqslant 2000$	$2000 < x \leqslant 3000$	$3000 < x \leqslant 5000$
Frequency	7	22	26	18	7

(i) Draw a histogram to illustrate the distribution. [5]

(ii) A student claims that the midrange is $2750\,\mathrm{cm}^3$. Discuss briefly whether he is likely to be correct. [1]

(iii) Calculate estimates of the mean and standard deviation of the engine sizes. Explain why your answers are only estimates. [5]

(iv) Hence investigate whether there are any outliers in the sample. [3]

(v) A vehicle duty of £1000 is proposed for all new cars with engine size greater than $2000\,\mathrm{cm}^3$. Assuming that this sample of cars is representative of all new cars in Britain and that there are 2.5 million new cars registered in Britain each year, calculate an estimate of the total amount of money that this vehicle duty would raise in one year. [3]

(vi) Why in practice might your estimate in part (v) turn out to be too high? [1]

7 Yasmin has 5 coins. One of these coins is biased with P(heads) = 0.6. The other 4 coins are fair. She tosses all 5 coins once and records the number of heads, X.

(i) Show that $P(X = 0) = 0.025$. [2]

(ii) Show that $P(X = 1) = 0.1375$. [4]

The table shows the probability distribution of X.

r	0	1	2	3	4	5
$P(X=r)$	0.025	0.1375	0.3	0.325	0.175	0.0375

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

(iv) Comment on the skewness of the distribution. [1]

(v) Find $E(X)$ and $Var(X)$. [5]

(vi) Yasmin tosses the 5 coins three times. Find the probability that the total number of heads is 3. [4]

THERE ARE NO QUESTIONS WRITTEN ON THIS PAGE.