

UNIVERSITY OF CAMBRIDGE LOCAL EXAMINATIONS SYNDICATE  
Joint Examination for the School Certificate  
and General Certificate of Education Ordinary Level

**STATISTICS**

**4040/1**

PAPER 1

**OCTOBER/NOVEMBER SESSION 2001**

2 hours 15 minutes

Additional materials:

- Answer paper
- Graph paper (3 sheets)
- Mathematical tables
- Pair of compasses
- Protractor

**TIME** 2 hours 15 minutes

**INSTRUCTIONS TO CANDIDATES**

Write your name, Centre number and candidate number in the spaces provided on the answer paper/answer booklet.

Answer **all** questions in Section A and not more than **four** questions from Section B.

Write your answers on the separate answer paper provided.

If you use more than one sheet of paper, fasten the sheets together.

All working must be clearly shown.

**INFORMATION FOR CANDIDATES**

The number of marks is given in brackets [ ] at the end of each question or part question.

The use of an electronic calculator is expected in this paper.

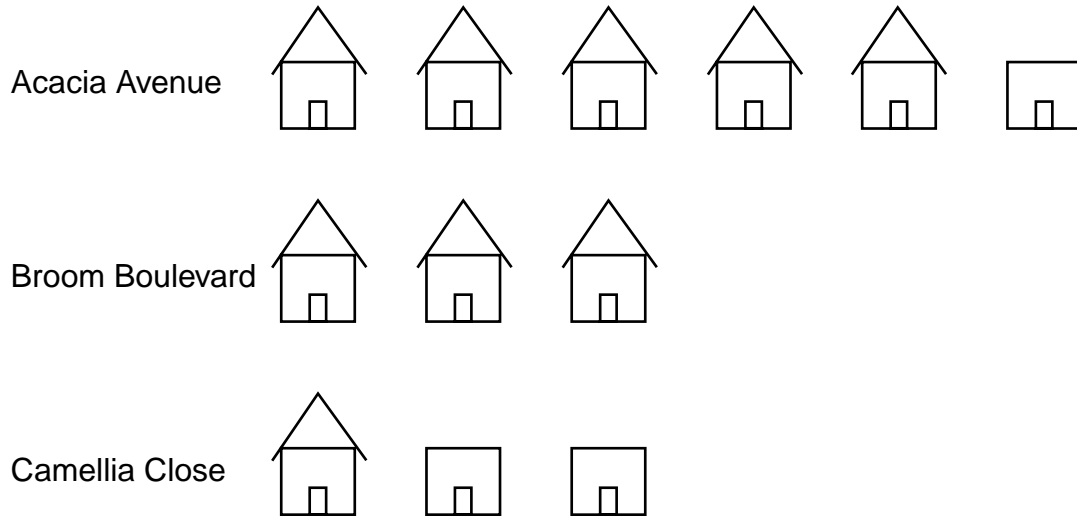
---

This question paper consists of 11 printed pages and 1 blank page.

## Section A [36 marks]

Answer **all** of the questions 1 to 6.

- 1 The pictogram below illustrates the numbers of houses in three streets in a village.



The symbol  represents 4 houses.


- (i) Calculate the number of houses in Broom Boulevard. [1]

In the three streets there are 45 houses altogether. One house is chosen at random from the three streets.

- (ii) Calculate, as a fraction in its lowest terms, the probability that the house is in Broom Boulevard. [1]

Two houses are chosen at random from the three streets.

- (iii) Calculate, as a fraction in its lowest terms, the probability that both houses are in Broom Boulevard. [2]

- (iv) Calculate how many houses the symbol  represents. [2]

- 2 A class of 30 pupils includes 12 boys and 18 girls.

The class teacher wanted to select 5 pupils to represent the class on a Sports committee.

- (i) The teacher decided to use a random sample. Describe how she could have selected this sample. [2]

The following week the teacher selected 5 pupils to represent the class on a Library committee.

- (ii) The teacher decided to use the method of systematic sampling. Describe this method. [2]
- (iii) The teacher later decided that the Library committee should represent boys and girls fairly. Calculate the number of boys and the number of girls required for this sample. [1]

- 3 The marks scored in a test by a group of eleven students were

8 7 4 19 9 15 7 7 1 25 8.

Three different measures of average can be found from the data above.

One measure has the value 8.

- (i) Give the name of this statistical measure. [1]
- (ii) Explain briefly how it is obtained. [1]

A second measure has the value 10.

- (iii) Give the name of this statistical measure. [1]
- (iv) Explain briefly how it is obtained. [1]

The third measure has the value 7.

- (v) Give the name of this statistical measure. [1]
- (vi) Explain briefly how it is obtained. [1]

- 4 Hamel took part in a traffic survey outside her school. She kept a record of the different types of vehicles passing the school in periods of 5 minutes. During one particular 5 minute period, she recorded 30 vehicles passing the school as follows:

lorry car van lorry car van car bus car car  
 van van car car lorry car bus car car van  
 car van lorry car van car bus car lorry van.

- (i) Copy and complete the following tally chart for these data.

<i>Type of vehicle</i>	<i>Tally</i>	<i>Frequency</i>
Car Bus Van Lorry		

[2]

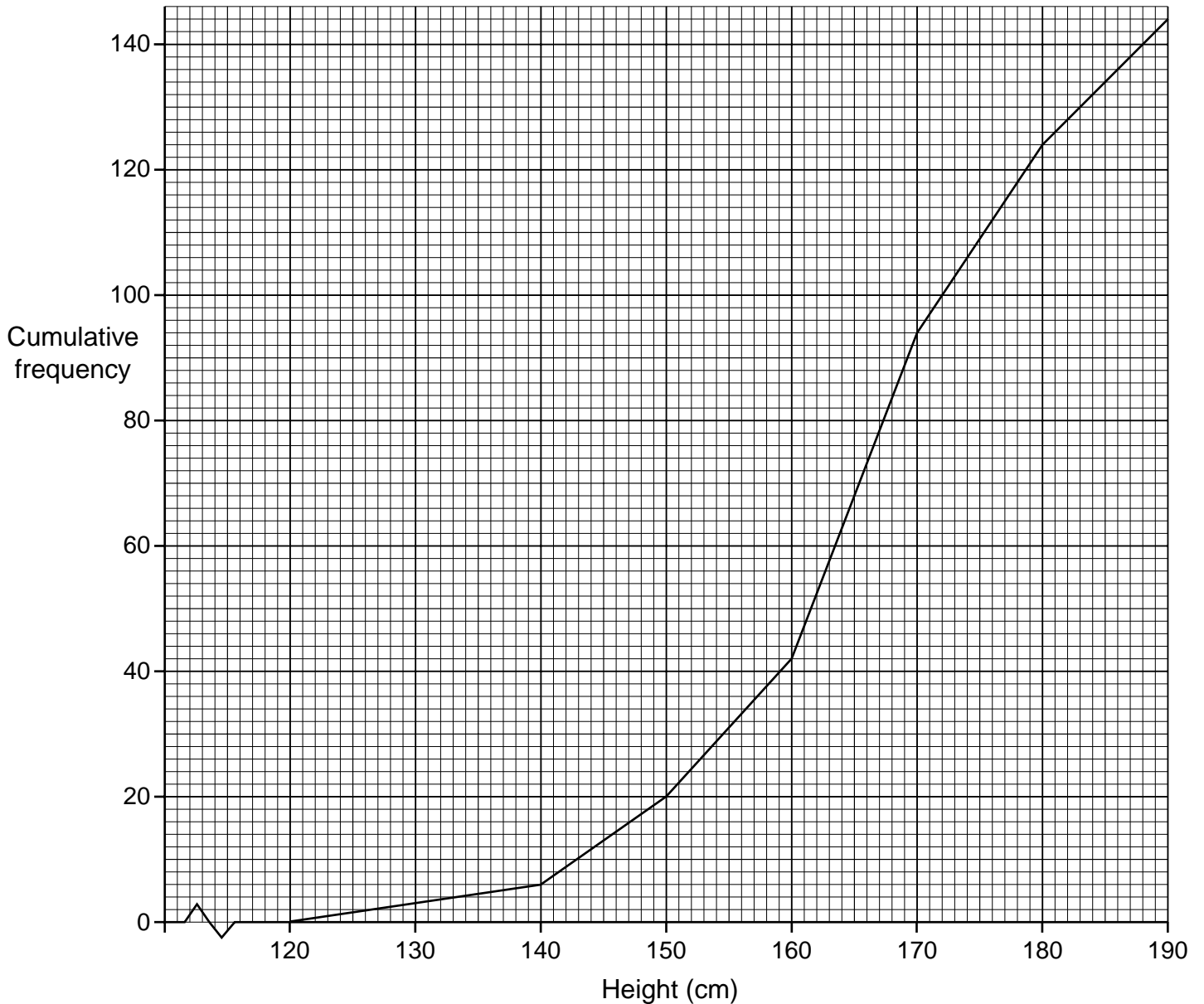
The data is to be illustrated by a pie chart.

- (ii) Calculate, showing your method, the sector angles of the pie chart. [2]
- (iii) Draw the pie chart, using a circle of radius 5 cm, labelling each sector with the type of vehicle it represents. [2]

In another 5 minute period, Hamel recorded 35 vehicles passing the school.

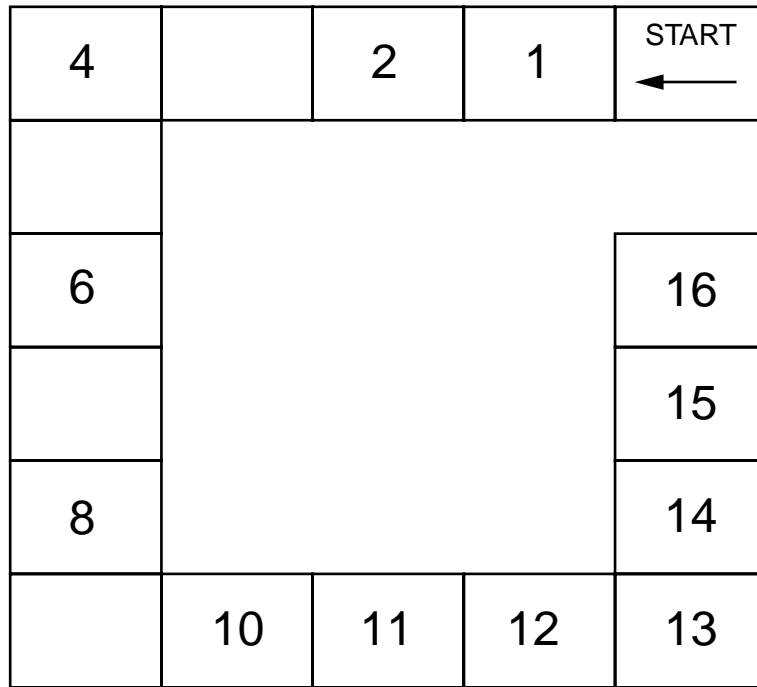
- (iv) Calculate, to 2 significant figures, the radius of a comparable pie chart which could be used to represent this increased number of vehicles. [2]

- 5 The cumulative frequency polygon shown below represents the heights of a group of 144 students.



Use the diagram to estimate

- (i) the median height of the students, [1]
- (ii) the interquartile range of the height of the students, [2]
- (iii) the number of students less than 165 cm in height. [1]



A group of boys are playing a game called BLANKO. To play the game players take it in turns to throw a fair die once and then move their counter the number of places shown uppermost on the die.

Before the first turn each counter is placed on the START square. If a counter lands on a blank square it is then moved back to the START square. The first player whose counter lands on square 16 is the winner. The player must score the exact number for his counter to land on square 16, otherwise he is not allowed to move his counter and must wait for his next turn.

- (i) Calculate the probability that a counter lands on a blank square on a player's first turn. [1]
- (ii) A counter is moved to square 4 on a player's first turn. Calculate the probability that his counter lands on a blank square on his second turn. [1]
- (iii) A player's counter is on square 13 after three turns. Write down, in the order thrown, the different scores the player could have had. [3]
- (iv) Calculate the probability that a player moves his counter from START to square 16 after exactly three turns. [2]

**Section B** [64 marks]

Answer not more than **four** of the questions 7 to 11.

Each question in this section carries 16 marks.

- 7 (a) The data below relates to the region Exshire in the year 1982.

<i>Age group</i>	<i>Population</i>	<i>Number of deaths</i>	<i>Death rate per 1000</i>	<i>Standard Population</i>
0 - 21	25 000	275	<i>P</i>	25 %
22 - 48	20 000	<i>Q</i>	8	35 %
49 and over	<i>R</i>	525	15	40 %

- (i) Calculate each of the numbers represented by the letters *P*, *Q* and *R*. [3]
- (ii) Calculate the crude death rate of Exshire. [3]
- (iii) Calculate the standardised death rate of Exshire. [3]
- (b) The data below refer to the number of children per family for a particular class in a school in 1984.

<i>Number of children per family</i>	<i>Number of families</i>
1	3
2	20
3	6
4	3

- (i) Explain briefly why the number 0 is not included in the “number of children per family” column. [1]
- (ii) Calculate the range in the number of children per family. [1]
- (iii) Calculate, to 2 significant figures, the mean number of children per family for the class. [3]

In 1958 a similar class survey showed that the number of children per family ranged from 1 to 8 and the mean number of children per family was 2.8.

- (iv) Given that the class surveys are representative of the numbers of children per family in 1958 and in 1984, state two changes that have taken place in the number of children per family in 1984 compared with 1958. [2]

- 8 A sample of 120 bolts was taken from the output of a machine. The diameter of each bolt in the sample was measured and the results are given below.

<i>Diameter of bolt <math>d</math> (mm)</i>	<i>Number of bolts</i>
$2.01 \leq d < 2.04$	24
$2.04 \leq d < 2.05$	36
$2.05 \leq d < 2.06$	26
$2.06 \leq d < 2.07$	22
$2.07 \leq d < 2.09$	12

- (i) State the median diameter of the bolts. [1]
- (ii) State the modal class of the distribution of the diameters of the bolts. [1]
- (iii) Using 2 cm to represent 0.01 mm of the diameter on the horizontal axis, starting at 2.00 mm, and a column of height 18 cm to represent the  $2.04 \leq d < 2.05$  diameter group, draw, on graph paper, a histogram to illustrate the distribution of the diameters of the bolts. [4]
- (iv) Write down the mid-values of the five groups of diameters. [2]
- (v) Calculate, to 3 significant figures, an estimate of the mean diameter of the bolts. [2]
- (vi) Calculate, to 3 significant figures, an estimate of the standard deviation of the diameters of the bolts. [3]

Bolts are only accepted by the purchaser if their diameter falls within the limits  $2.02 \leq d \leq 2.08$ .

- (vii) Estimate, to the nearest integer, the percentage of bolts which are not accepted by the purchaser. [3]



- 9 The number of hours worked by the employees of a large factory during one particular week is shown below.

<i>Number of hours <math>h</math></i>	<i>Number of employees</i>
$25 \leq h < 30$	50
$30 \leq h < 40$	150
$40 \leq h < 45$	400
$45 \leq h < 55$	150
$55 \leq h < 60$	50

- (i) Using 2 cm to represent 5 hours on the horizontal axis, starting at 25 hours, and 2 cm to represent 100 employees on the vertical axis, draw, on graph paper, a cumulative frequency **polygon** to illustrate the distribution of the number of hours worked by the employees. [4]
- (ii) Use your cumulative frequency polygon to find
- (a) the lower quartile, [1]
  - (b) the median, [1]
  - (c) the upper quartile, [1]
  - (d) the 10th percentile, [1]
  - (e) the 90th percentile. [1]
- (iii) Use your cumulative frequency polygon to estimate the number of employees who worked more than 54 hours. [2]

An employee is selected at random from the employees in the survey.

- (iv) Use your cumulative frequency polygon to calculate, to 1 decimal place, the probability that this employee worked between 37.5 hours and 47.5 hours during the week. [2]

Exactly 35 % of the employees worked less than  $H$  hours.

- (v) Use your cumulative frequency polygon to estimate the value of  $H$ . [2]

The data could be illustrated by a histogram.

- (vi) Describe the shape that the histogram would have. [1]

- 10** The manageress of a boarding kennel in England feeds the dogs in her care on Puppygrow complete dog food. The instructions on the bags in which the dog food is supplied give the following guidance regarding the mass of food to be given.

<i>Body mass of dog to the nearest 5 kg (x kg)</i>	<i>Mass of food per day (y g)</i>
5	150
10	200
15	350
20	400
25	500
35	550
40	650
50	800

- (i) Using 2 cm to represent 5 kg on the horizontal axis and 2 cm to represent 100 g on the vertical axis, plot, on graph paper, a scatter diagram of the data given above. [3]
- (ii) By calculating and plotting  $(\bar{x}, \bar{y})$  and the two semi-averages, draw the line of best fit. [6]
- (iii) Use your line of best fit to estimate, to 3 significant figures, the mass of food per day required by a dog having a body mass of 35 kg. [1]
- (iv) State, giving a reason, whether you think it would be correct to extend the line of best fit to the  $y$ -axis. [1]

The manageress decided to try another complete dog food, Doggybix, which recommended that the mass of food per day to be given to dogs should be based on the simple rule: “100 g for every 5 kg of body mass of the dog”.

- (v) Draw a line on your graph to represent the relationship between mass of food and body mass of dog for Doggybix complete dog food. [2]
- (vi) Use your graphs to find the body mass of the dog for which both Puppygrow and Doggybix recommend the same amount of food. [1]

Given that 100 g of Puppygrow costs the same as 100 g of Doggybix,

- (vii) state, with a reason, which brand of dog food is more economical for
- (a) dogs having a body mass greater than 20 kg, [1]
- (b) dogs having a body mass less than 15 kg. [1]

- 11 (a) In this part of the question all probabilities should be given as fractions in their lowest terms.

Michael bought 6 tins of beans, 4 tins of spaghetti and 2 tins of tomatoes. Unfortunately, all the labels came off the tins, leaving them looking the same in size and shape.

Michael wished to open a tin of tomatoes and a tin of beans for tea.

- (i) Michael took a tin at random and opened it. Calculate the probability that it was a tin of beans. [1]
- (ii) The first tin opened contained beans. Michael opened a second tin. Calculate the probability that it was a tin of tomatoes. [1]
- (iii) Calculate the probability of Michael opening any 2 tins and getting 1 tin of tomatoes and 1 tin of beans. [2]

Michael opened a tin of tomatoes and a tin of beans.

- (iv) Calculate the probability that the next 2 tins that he opens will both contain spaghetti. [2]
- (b) An office secretary carried out a survey, over a period of time, to compare the delivery times of letters and parcels. The results of the survey are shown below.

<i>Days to deliver</i>	1	2	3	4
<i>Letters</i>	89%	8%	2%	1%
<i>Parcels</i>	6%	54%	33%	7%

For example, 54% of parcels take exactly 2 days to deliver.

Use the information to calculate the probability of

- (i) a parcel taking more than 2 days to be delivered, [1]
- (ii) 2 letters each taking exactly 2 days to be delivered, [2]
- (iii) a letter and a parcel taking the same number of days to be delivered. [3]
- (c) A circular dart board of diameter 50 cm is mounted centrally on a square cork board of side 80 cm. A dart is thrown in the direction of the dart board and sticks in either the dart board or the cork surrounding the dart board. The dart is equally likely to stick in any point within the square. Taking  $\pi = 3.14$ , calculate the probability, to 2 significant figures, that the dart sticks in the cork surround. [4]

