

UNIVERSITY OF CAMBRIDGE INTERNATIONAL EXAMINATIONS  
General Certificate of Education Ordinary Level

**STATISTICS**

**4040/01**

Paper 1

October/November 2006

**2 hours 15 minutes**

Additional Materials: Answer Booklet/Paper  
Graph paper (2 sheets)  
Mathematical tables  
Pair of compasses  
Protractor

**READ THESE INSTRUCTIONS FIRST**

If you have been given an Answer Booklet, follow the instructions on the front cover of the Booklet.  
Write your Centre number, candidate number and name on all the work you hand in.  
Write in dark blue or black pen on both sides of the paper.  
You may use a soft pencil for any diagrams or graphs.  
Do not use staples, paper clips, highlighters, glue or correction fluid.

Answer **all** questions in Section A and not more than **four** questions from Section B.  
Write your answers on the separate Answer Booklet/Paper provided.  
All working must be clearly shown.  
The use of an electronic calculator is expected in this paper.

The number of marks is given in brackets [ ] at the end of each question or part question.  
At the end of the examination, fasten all your work securely together.

## Section A [36 marks]

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

- 1 The mean mass of 7 pears is 112 grams. The masses of 4 of the pears are 109, 115, 124 and 130 grams respectively, and the remaining 3 pears all have equal masses.
- (i) Calculate the mass of each of the remaining 3 pears. [4]
- (ii) State the modal mass of the 7 pears. [1]
- (iii) State the median mass of the 7 pears. [1]

- 2 The table below gives two pieces of information about each of the houses in a village: the number of adults who live in the house, and the number of cars owned by those adults.

For example, there are 25 houses in which there are 3 adults, and 2 cars owned by those 3 adults.

		Number of adults in a house				TOTAL
		1	2	3	4	
Number of cars owned	0	12	5	1	3	21
	1	22	46	14	9	91
	2	0	$y$	25	17	$z$
	3	0	0	7	15	22
TOTAL		34	$x$	47	44	

There are twice as many houses in which two adults live as there are houses in which one adult lives.

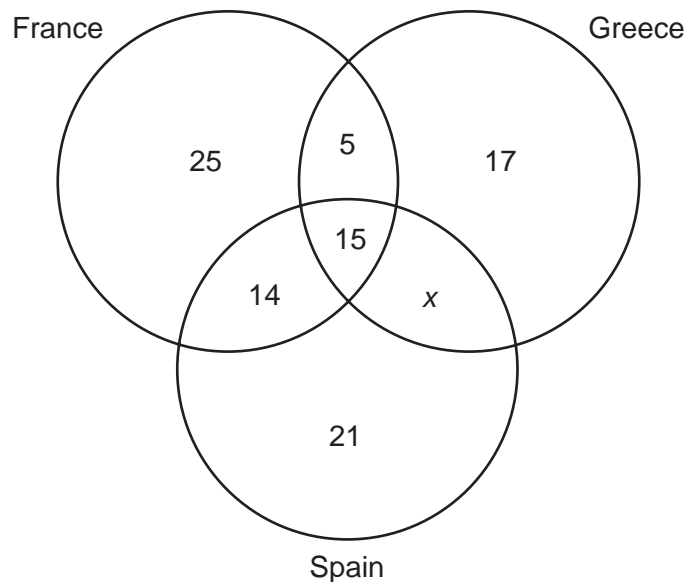
- (i) State the value of  $x$ . [1]
- (ii) Find the value of  $y$ . [1]
- (iii) Find the value of  $z$ . [1]

All the other 7 houses in the village are empty, and so have no adults living in them.

- (iv) State the total number of houses in the village. [1]
- (v) Calculate the total number of adults living in the village. [2]

- 3 (i) Give one advantage and one disadvantage of each of the following as a method of collecting information from the residents of a village:
- (a) postal questionnaires, [2]
- (b) quota sampling. [2]
- (ii) When sampling from a population of pupils at a school, a stratified random sample is often more appropriate than a simple random sample.
- Briefly explain why this is so. [2]

- 4 The diagram below summarises information which has been collected about a sample of 100 people who have all visited at least one of France, Spain and Greece.

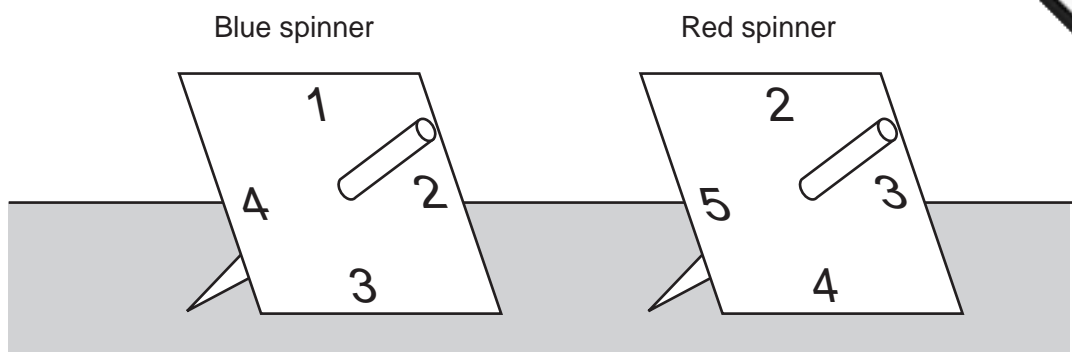


For example, 25 people have visited France only,

5 people have visited both France and Greece, but not Spain.

- (i) Find the value of  $x$ . [2]
- (ii) State what the value of  $x$  represents. [1]
- (iii) Calculate the number of people in the sample who have not visited Greece. [1]
- (iv) Calculate the total number of people in the sample who have visited Spain. [2]

5



An unbiased blue spinner has edges numbered 1, 2, 3 and 4. An unbiased red spinner has edges numbered 2, 3, 4 and 5. When each spinner is spun, the **result** of the spin is the number on the edge which comes to rest on the surface. In a game, the red and blue spinners are both spun, and the **game score** is obtained by multiplying the **results** on the two spinners.

(i) Copy and complete the following table by inserting the **game score** in each cell. [2]

		Result of blue spinner			
		1	2	3	4
Result of red spinner	2				
	3				
	4				
	5				

(ii) Calculate the probability that both spinners show the same **result**. [2]

(iii) State the probability that the **game score** is odd. [1]

(iv) If the **game score** is 4, state the probability that the **result** on the blue spinner is 2. [1]

- 6 A manufacturer of electric light bulbs wished to market a new type of bulb. A random sample of 100 of these bulbs was tested, and the lifetime of each bulb was recorded to the nearest hour. The following table summarises the data collected.

<i>Lifetime (to nearest hour)</i>	<i>Number of bulbs</i>
700 – 709	3
710 – 719	7
720 – 729	15
730 – 739	38
740 – 749	21
750 – 759	16

- (i) Using 2 cm to represent 10 hours on the horizontal axis, starting at 690 hours, and using 2 cm to represent 10 light bulbs on the vertical axis, draw on graph paper a cumulative frequency **polygon** to represent these data. [5]
- (ii) Use your graph to estimate the number of hours below which the lifetimes of 5% of the bulbs fell. [1]

**Section B** [64 marks]

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

Each question in this section carries 16 marks.

**7 In this question calculate all death rates per thousand, and to 2 decimal places.**

The first table below gives certain information about the population and deaths in town *A* for the year 2004, together with the standard population of the area in which town *A* is situated.

<i>Age group</i>	<i>Deaths</i>	<i>Population in age group</i>	<i>Standard population (%)</i>
0 – 19	16	3500	25
20 – 39	$x$	6000	30
40 – 59	38	5000	25
60 and over	77	6000	20

- (i) The death rate for the 20 – 39 age group is 2.50 per thousand. Show that  $x = 15$ . [1]
- (ii) Calculate the crude death rate for town *A*. [4]
- (iii) Calculate the death rates for the other three age groups. [2]
- (iv) Using the given rate for the 20 – 39 age group, and the rates you have calculated in (iii), calculate the standardised death rate for town *A*. [4]

**[Question 7 continues on the next page.]**

The table below gives information about town *B*, another town in the same area, for the year 2004. The crude death rate for town *B* in 2004 was 6.84 per thousand.

<i>Age group</i>	<i>Death rate per thousand</i>	<i>Population in age group (%)</i>
0 – 19	3	32
20 – 39	2	28
40 – 59	4	24
60 and over	27	16

- (v) Calculate the standardised death rate for town *B* in the year 2004, using the same standard population as for town *A*. [2]

One of the two towns has a higher crude death rate, but the other has a higher standardised death rate.

- (vi) Give a brief explanation of why such a situation can occur. [1]
- (vii) State, with a reason, which of the two towns would appear to have a healthier environment. [2]

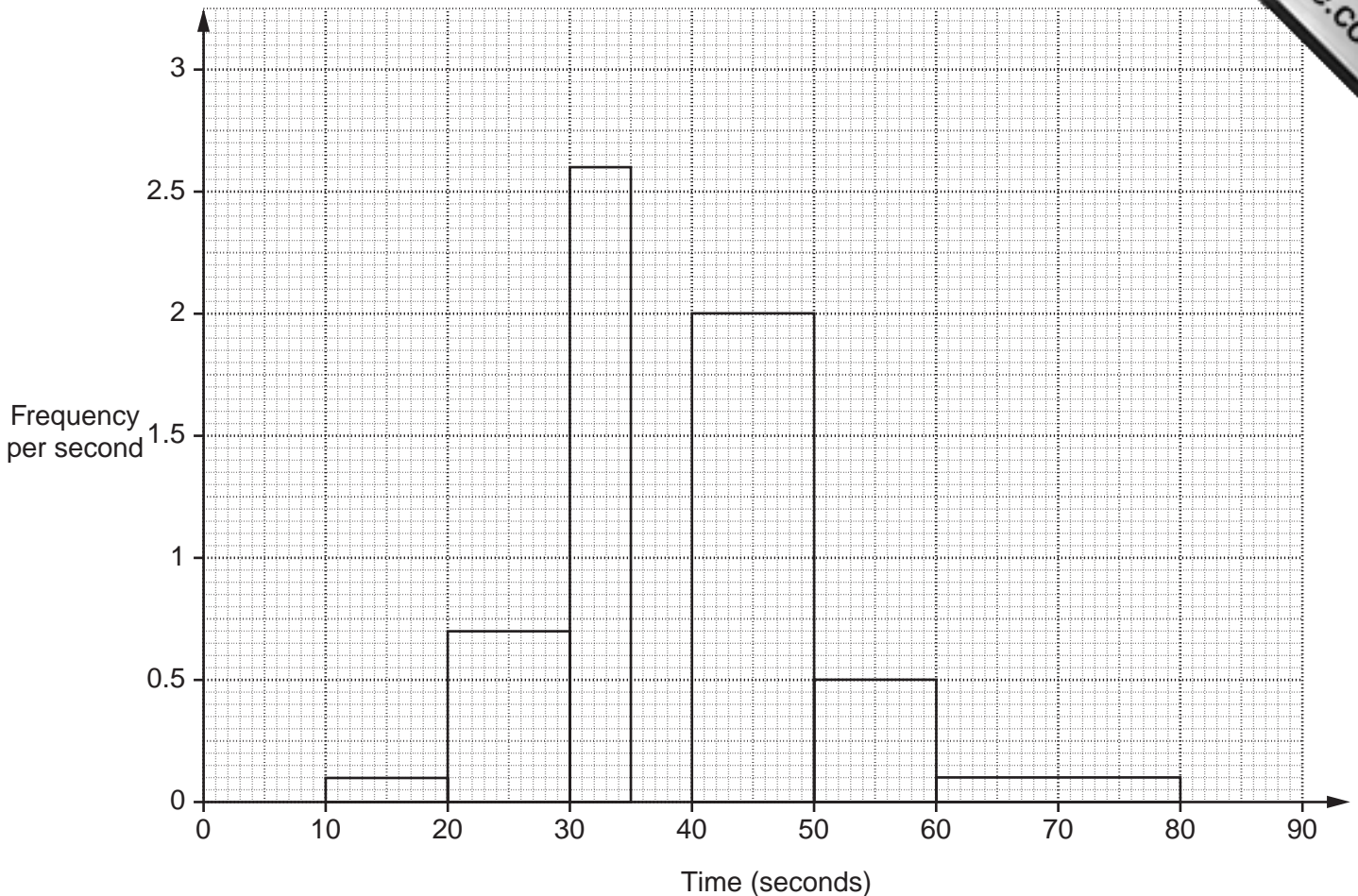
- 8 Each week, for ten weeks, a large bookshop kept a record of the amount of money,  $x$ , it had spent advertising a particular book, and the number of copies of the book,  $y$ , which were sold. The following table gives the figures collected.

<i>Amount spent (\$x)</i>	<i>Number of copies sold (y)</i>
540	695
490	635
455	560
260	340
265	340
405	550
175	300
365	465
300	425
145	230

- (i) Using 2 cm to represent \$100 on the horizontal axis, and 2 cm to represent 100 copies on the vertical axis, starting both axes from the origin, plot on graph paper a scatter diagram illustrating these figures. [5]
- (ii) Using the method of semi-averages, calculate and plot three appropriate points, and use them to draw the line of best fit. [5]
- (iii) Use your line of best fit to estimate the number of copies of the book sold in a week if \$400 was spent on advertising it. [2]
- (iv) Interpret the point at which the line of best fit meets the vertical axis, and give a reason why it would be unwise to use this point for estimating sales of the book. [2]
- (v) For another book, the line of best fit is given by the equation  $y = x$ . Draw this line on your graph, and use the two lines of best fit to compare the effectiveness of the amounts spent on advertising the two books. [2]



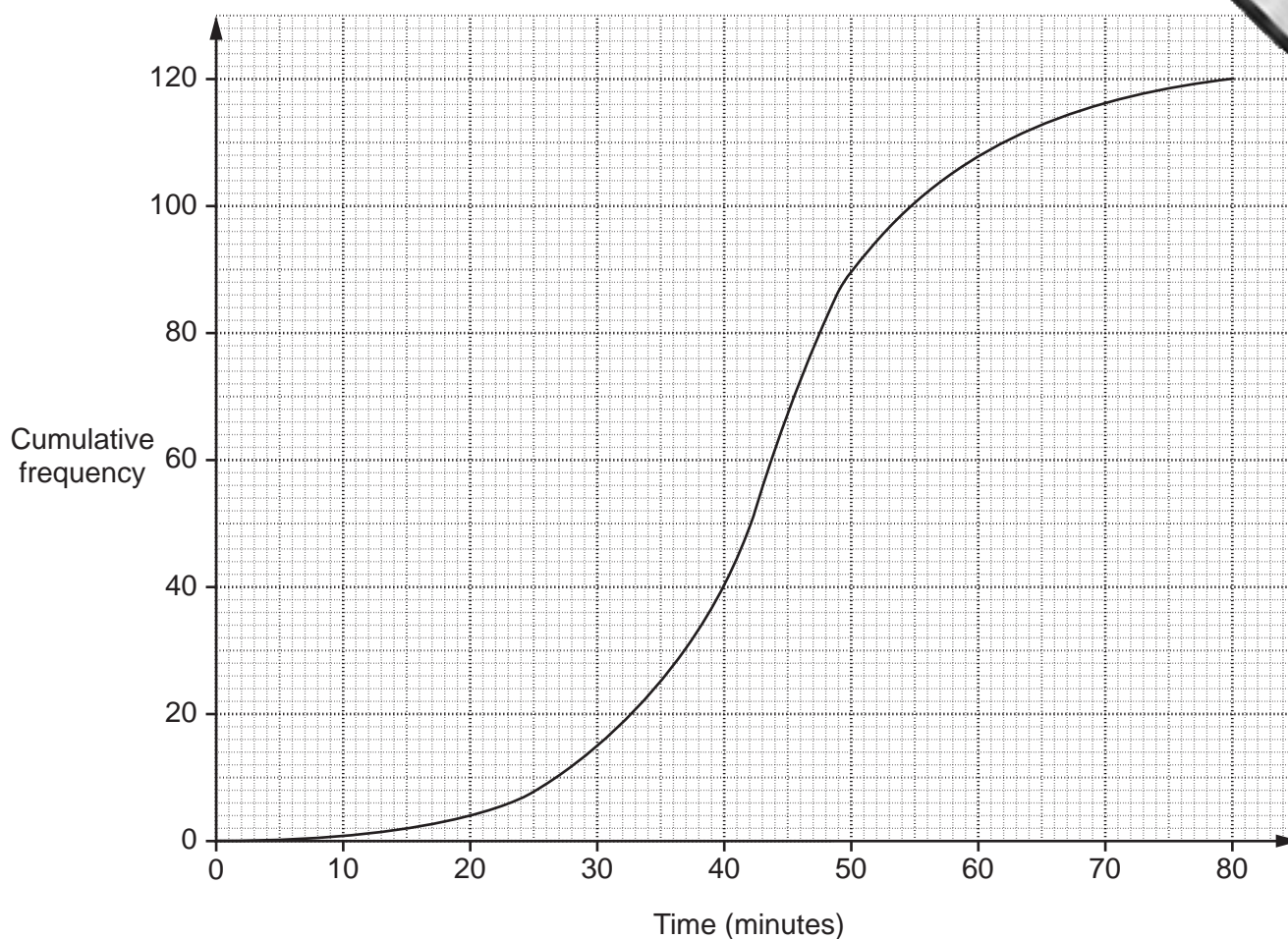
- 9 Sixty children were given a puzzle to solve. The times, in seconds, taken by the children to solve the puzzle were grouped into a frequency distribution, and represented by the following histogram. One rectangle, representing the '35 – under 40' seconds class, has been omitted.



The '10 – under 20' seconds class contained one child.

- (i) Show that 2 children required '60 – under 80' seconds to solve the puzzle. [2]
- (ii) Calculate the number of children represented by each of the other rectangles shown in the histogram. [5]
- (iii) Calculate the value on the vertical axis of the height of the rectangle representing the '35 – under 40' seconds class. [4]
- (iv) State the modal class, giving a reason for your choice of class. [2]
- (v) The two classes '20 – under 30' seconds and '30 – under 35' seconds are to be combined into a single class. Calculate the value on the vertical axis of the height of the rectangle which would represent this combined class. [3]

- 10 The cumulative frequency graph below shows the distribution of the times, in minutes, for the journeys home from work on a Monday of the 120 workers employed at a factory.



- (i) Use the cumulative frequency graph to estimate
- (a) the median, [1]
  - (b) the lower quartile, [1]
  - (c) the upper quartile, [1]
  - (d) the 10th percentile, [1]
  - (e) the 90th percentile. [1]
- (ii) Estimate the number of workers whose journey time was longer than 43 minutes. [3]
- (iii) Determine which percentile corresponds to a journey time of 52 minutes. [2]

Each day, all the workers leave the factory at 1715 (5.15 pm).

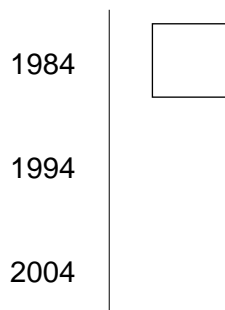
- (iv) Calculate the probability that on this Monday one worker, chosen at random, arrived home by 1811 (6.11 pm). [2]

On the next day, Tuesday, due to bad weather, the journey time of every worker was longer than on the Monday.

- (v) State the effect which this would have on
- (a) the median journey time,
  - (b) the interquartile range of journey times,
- compared to the Monday. [2]
- (vi) Calculate the probability that on the Tuesday one worker, chosen at random, arrived home by 1811 (6.11 pm). [2]

[Question 11 is printed on the next page.]

- 11 (i) A pictogram is to be used to represent the amount of cloth a manufacturer produced in the years 1984, 1994 and 2004.  
 10 000 square metres of cloth was produced in 1984, and this is represented in the pictogram by a square of side 1 cm.  
 20 000 square metres of cloth was produced in 1994.  
 35 000 square metres of cloth was produced in 2004.



- (a) Copy and complete the above pictogram. [2]
- (b) If the amount of cloth produced in 2004 is to be represented in a different pictogram by one square, calculate the length of a side of this square. [2]
- (ii) A random sample of 50 people who attended a cricket match were asked to rate how entertaining they had found it on a five-point scale, *A*, *B*, *C*, *D*, *E*, where *A* represents 'thoroughly entertaining' and *E* represents 'not at all entertaining'. The results are given in the following table.

<i>Rating</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>
<i>Number of people</i>	5	17	14	9	5

- (a) Give a reason why the arithmetic mean would not be an appropriate measure to use for these data. [1]
- (b) Represent these data by a fully-labelled pie chart of radius 5 cm. [4]
- A second sample of people who attended this cricket match were asked the same question, and their results were represented by a second pie chart, which had a radius of 6 cm.  
 [You are not required to draw the second pie chart.]
- (c) Calculate how many people there were in this second sample. [2]
- (d) In the second pie chart the angle of the sector representing a rating of *B* is 110 degrees. Calculate the number of people in the second sample who gave the match a rating of *B*. [2]
- (e) The angles in the second pie chart representing ratings of *A*, *C*, *D* and *E* are in the ratio 1:2:1:1. Calculate the number of people in each of these categories in the second sample. [3]