

UNIVERSITY OF CAMBRIDGE INTERNATIONAL EXAMINATIONS
General Certificate of Education Ordinary Level

STATISTICS

4040/01

Paper 1

October/November 2004

2 hours 15 minutes

Additional Materials: Answer Booklet/Paper
Graph paper (3 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.
At the end of the examination, fasten all your work securely together.

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.

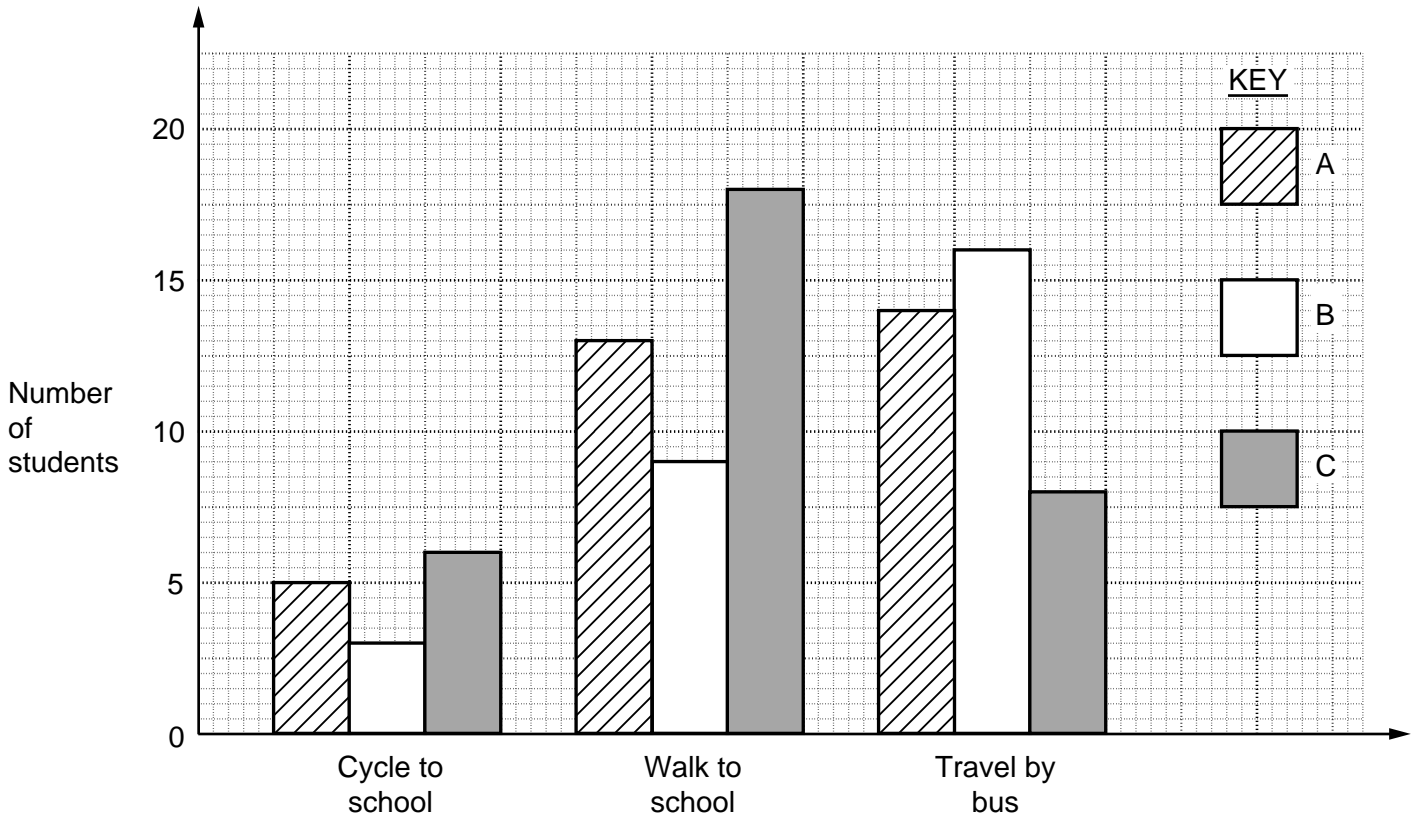
This document consists of **11** printed pages and **1** blank page.



Section A [36 marks]

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

- 1 The diagram below shows how students in Classes A, B and C travel to school each day.



- (i) Write down how many students walk to school. [1]
- (ii) Write down the number of students in Class A. [1]
- (iii) Calculate how many more students there are in Class C than in Class B. [1]
- (iv) Calculate the probability that a student, chosen at random, is from Class A and walks to school. [1]
- (v) Given that a student travels to school by bus, calculate the probability that the student is a member of Class C. [2]

2 A factory has 450 employees.

- (i) Describe how a simple random sample of size 75 could be selected from these employees. [2]

There are 240 employees in the manufacturing section, 120 employees in the distribution section and 90 employees in management.

A stratified sample of size 75 is to be selected from the 450 employees.

- (ii) Calculate the number of staff that should be chosen from each of the three types of employee. [3]
- (iii) Give one advantage of using a stratified random sample rather than a simple random sample. [1]

3 The owner of a riding school keeps a record of his total expenditure for each year. In 2001 he spent 45% on hay and straw, 36% on pony nuts, 12% on equipment and the rest on veterinary bills.

The data are to be illustrated by a pie chart.

- (i) Calculate, each to the nearest degree, the sector angles of the pie chart. [2]
- (ii) Using a circle of radius 5 cm, draw the pie chart. [2]

In 2002 the total expenditure increased by 10% of the total expenditure for 2001.

- (iii) Calculate, to 2 significant figures, the radius of the comparable pie chart which could be used to represent the expenditure for 2002. [2]

- 4 The first fifty people arriving at a supermarket were asked their ages. The graph below shows the cumulative frequency curve of their ages.



Use the graph to estimate

- (i) the median value of their ages, [1]
- (ii) the interquartile range of their ages, [1]
- (iii) the number of people who were between 20 and 35 years of age, [1]
- (iv) the probability that a person, chosen at random from the fifty people, was at least 18 years of age. [3]

- 5 Each morning Mr Patel drives to work. Sometimes he is late because of traffic problems. When he arrives at work the car park is sometimes full.

The probability that Mr Patel arrives late for work is $\frac{2}{5}$.

The car park was full for 80% of the days when he was late for work.

The car park was full for 40% of the days when he was not late for work.

Calculate the probability that

- (i) the car park was full and he arrived late for work, [2]
(ii) the car park was not full when he arrived for work. [4]

- 6 The data below show the maximum temperatures (in °C) of 12 towns recorded on 24th July 1998.

36 34 37 34 35 37 37 39 35 37 39 35

- (i) Write down the modal maximum temperature. [1]
(ii) Calculate the median maximum temperature. [1]
(iii) Calculate the mean and standard deviation of the maximum temperatures recorded that day in these 12 towns. [4]

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 40 members of a Youth Club took part in a survey to find out their favourite activity. The table below shows the numbers of boys and girls choosing each activity.

<i>Favourite activity</i>	<i>Boys</i>	<i>Girls</i>
Discos	6	8
Quizzes	2	2
Sport	10	4
Films	4	4

For example, 10 boys chose sport as their favourite activity.

A member of the club is selected at random.

Calculate the probability that

- (i) the member is a girl, [2]
 (ii) the member chose discos as their favourite activity. [1]

A boy is selected at random and a girl is selected at random.

- (iii) Calculate the probability that they are a boy who chose films and a girl who chose quizzes as their favourite activities. [2]

- (b) A bag contains 10 white counters, 8 black counters and 7 red counters.

Josie picks out one counter at random.

- (i) Calculate the probability that she picks out a black counter. [1]

Josie then picks out a second counter at random, **without** returning the first counter to the bag.

She wins a prize if she picks out 2 counters of the same colour.

- (ii) Calculate the probability that she wins a prize. [3]

- (c) A spinner with its edges numbered 1, 2, 3 and 4 is biased.

For each spin, the probability of scoring 1 is 0.25, the probability of scoring 3 is 0.35 and the probability of scoring 4 is 0.10.

The spinner is used in a board game in which a player's counter is moved to the right at each turn by the score shown on the spinner.

1	2	3		4	5	6
---	---	---	--	---	---	---

If the player's counter lands on the blank square the player is out of the game.

- (i) Andy's counter is on square 2.
Calculate the probability that he will **not** be out of the game after one more turn. [1]
- (ii) Beatriz's counter is on square 2.
Calculate the probability that, after two more turns, her counter will be on square 4. [2]
- (iii) Ceri's counter is on square 1.
Calculate the probability that, after three more turns, his counter will be on square 5. [4]

- 8 A science teacher wants her students to find a relationship between degrees Celsius ($^{\circ}\text{C}$) and degrees Fahrenheit ($^{\circ}\text{F}$). To do this she asks 10 pupils to measure a number of temperatures in $^{\circ}\text{C}$ and also in $^{\circ}\text{F}$. The results of the experiment are shown in the table below.

$x (^{\circ}\text{C})$	10	15	20	30	35	48	56	64	70	82
$y (^{\circ}\text{F})$	50	60	64	88	98	120	130	146	160	184

- (i) Using 2 cm to represent 10°C on the horizontal axis and 2 cm to represent 20°F on the vertical axis, plot on graph paper a scatter diagram of the data given above. [2]
- (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 the temperature in $^{\circ}\text{F}$ corresponding to 40°C . [1]
- (iv) Use your line of best fit to estimate the temperature in $^{\circ}\text{C}$ corresponding to 180°F . [1]
- (v) Use your graph to find the equation of the line of best fit in the form $y = ax + b$, where a and b are constants.

Give your value of a to 1 decimal place and your value of b to the nearest integer. [3]

A "quick" method of conversion is to multiply degrees Celsius by 2.5 to convert to degrees Fahrenheit.

- (vi) Draw on your graph a line representing this "quick" method of conversion. [2]
- (vii) Using your graphs, estimate the temperature in $^{\circ}\text{C}$ for which both methods of conversion give the same answer. [1]

- 9 The grouped frequency table below shows the journey times, in minutes, taken by a group of 90 workers travelling to work on 10th August 2002.

<i>Journey time, t (minutes)</i>	<i>Frequency</i>
$20 \leq t < 30$	8
$30 \leq t < 35$	20
$35 \leq t < 40$	15
$40 \leq t < 55$	27
$55 \leq t < 75$	20

- (i) State the median group of journey times. [1]
- (ii) State the modal group of journey times. [1]
- (iii) Calculate an estimate of the total number of workers who took 45 minutes or less to travel to work. [2]
- (iv) Using 2 cm to represent 10 minutes and a column height of 4 cm to represent the $20 \leq t < 30$ journey time group, draw on graph paper a histogram to illustrate the distribution of the journey times of the workers. [4]
- (v) Calculate, to 3 significant figures, an estimate of the mean journey time. [3]
- (vi) Calculate, to 3 significant figures, an estimate of the standard deviation of the journey times. [3]

For another group of workers who travelled to work on 10th August 2002, the mean journey time was 48.7 minutes and the standard deviation was 10.7 minutes.

- (vii) Make **two** comments on the difference between these figures and the corresponding figures for the first group of workers, making clear which group of workers you are referring to. [2]

10 The annual wages of 150 women working for a company are shown below.

<i>Annual wage \$w</i>	<i>Number of women</i>
$5\,000 \leq w < 10\,000$	24
$10\,000 \leq w < 12\,000$	22
$12\,000 \leq w < 15\,000$	40
$15\,000 \leq w < 20\,000$	30
$20\,000 \leq w < 25\,000$	16
$25\,000 \leq w < 35\,000$	10
$35\,000 \leq w < 60\,000$	8

- (i) Using 2 cm to represent \$10 000 on the horizontal axis and 2 cm to represent 20 women on the vertical axis, draw on graph paper a cumulative frequency **polygon** to illustrate the distribution of the annual wage earned by the women. [4]
- (ii) Estimate the proportion of women with an annual wage of more than \$30 000. [2]
- (iii) Use your cumulative frequency polygon to find, to the nearest \$500,
- (a) the median, [1]
 - (b) the lower quartile, [1]
 - (c) the upper quartile, [1]
 - (d) the 10th percentile, [1]
 - (e) the 90th percentile. [1]

Exactly 40% of the women earned an annual wage of less than \$W.

- (iv) Use your cumulative frequency polygon to estimate the value of W. [1]

The range between the 10th percentile and the 90th percentile for the **men** working for this company was \$25 000.

- (v) (a) Give one advantage of using this range. [1]
- (b) Write down one statement comparing the wages of men and women working for this company. [1]

At Christmas all the women received the same bonus of \$250.

- (vi) (a) Write down their median wage after the bonus was paid. [1]
- (b) Write down the range between the 10th percentile and the 90th percentile after the bonus was paid. [1]

- 11 (a) The following table shows the death rates for Seeton in 1990, together with its population and the standard population.

<i>Age group</i>	<i>Population</i>	<i>Death rate per 1000</i>	<i>Standard population</i>
0 – 25	19 000	2	45%
26 – 55	17 000	5	30%
56 – 69	14 000	8	17%
70 and over	10 000	28	8%

Calculate

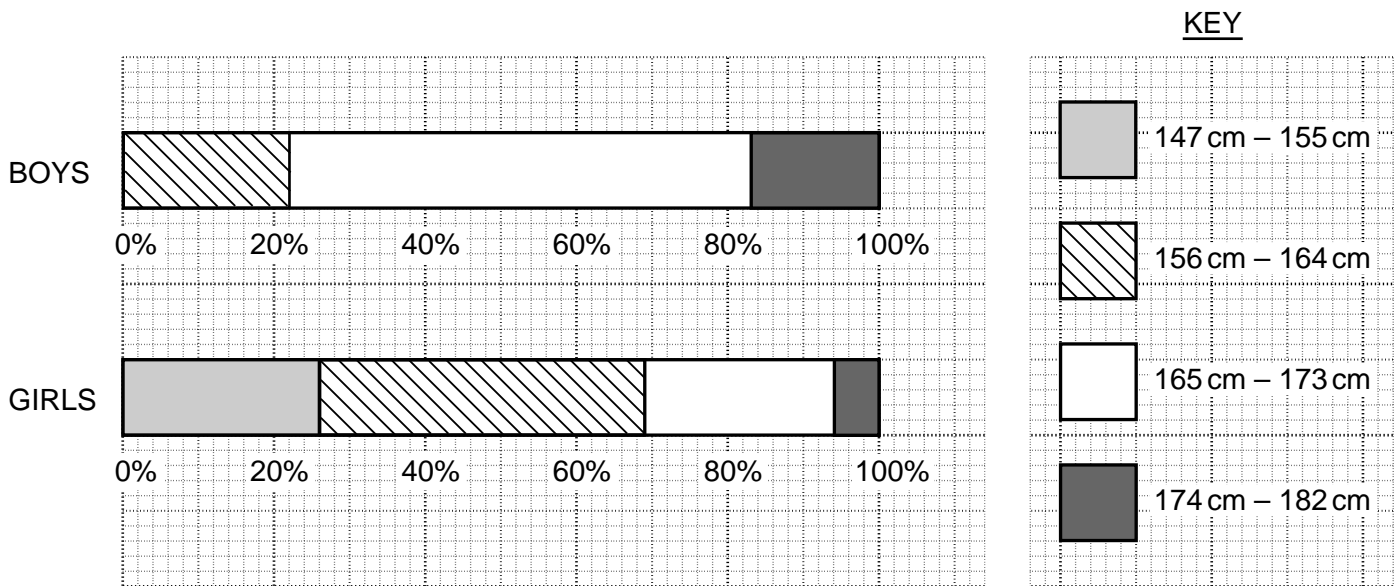
- (i) the total number of deaths in 1990, [2]
 (ii) the crude death rate, giving your answer to 1 decimal place, [3]
 (iii) the standardised death rate. [3]

The crude death rate of another town, Exton, in 1990 was 7.8 per thousand.
 The standardised death rate of Exton was 7.2 per thousand.

- (iv) State, with a reason, which of the two towns offers the better chance of a longer life. [2]

[Question 11 continues on the next page.]

- (b) The percentage charts below show information about the heights, measured to the nearest centimetre, of boys and girls in a school.



- (i) Estimate what percentage of boys have heights between 156 cm and 164 cm. [1]
- (ii) Estimate what percentage of girls do **not** have heights between 156 cm and 164 cm. [2]

Using the percentage charts above, state whether each of the following statements is true, or is false, or there is not enough information to decide whether it is true or false. Give a clear reason for each of your answers.

- (iii) The shortest boy is taller than the shortest girl. [1]
- (iv) The tallest boy is taller than the tallest girl. [1]
- (v) There are more boys than girls having heights between 165 cm and 173 cm. [1]

