

Descriptive Statistics – Practice Problems (99-04)

1. One thousand candidates sit an examination. The distribution of marks is shown in the following grouped frequency table.

Marks	1–10	11–20	21–30	31–40	41–50	51–60	61–70	71–80	81–90	91–100
Number of candidates	15	50	100	170	260	220	90	45	30	20

- (a) **Copy** and complete the following table, which presents the above data as a cumulative frequency distribution.

Mark	≤10	≤20	≤30	≤40	≤50	≤60	≤70	≤80	≤90	≤100
Number of candidates	15	65					905			

- (b) Draw a cumulative frequency graph of the distribution, using a scale of 1 cm for 100 candidates on the vertical axis and 1 cm for 10 marks on the horizontal axis.

- (c) Use your graph to answer parts (i)–(iii) below,

- (i) Find an estimate for the median score. (2)
- (ii) Candidates who scored less than 35 were required to retake the examination. How many candidates had to retake? (3)
- (iii) The highest-scoring 15% of candidates were awarded a distinction. Find the mark above which a distinction was awarded. (3)

(Total 16 marks)

2. At a conference of 100 mathematicians there are 72 men and 28 women. The men have a mean height of 1.79 m and the women have a mean height of 1.62 m. Find the mean height of the 100 mathematicians.

<i>Working:</i>	<i>Answer:</i>

(Total 4 marks)

3. The mean of the population x_1, x_2, \dots, x_{25} is m . Given that $\sum_{i=1}^{25} x_i = 300$ and

$$\sum_{i=1}^{25} (x_i - m)^2 = 625, \text{ find}$$

- (a) the value of m ;
- (b) the standard deviation of the population.

<i>Working:</i>	<i>Answers:</i>
	(a)
	(b)

(Total 4 marks)

4. A supermarket records the amount of money d spent by customers in their store during a busy period. The results are as follows:

Money in \$ (d)	0–20	20–40	40–60	60–80	80–100	100–120	120–140
Number of customers (n)	24	16	22	40	18	10	4

- (a) Find an estimate for the mean amount of money spent by the customers, giving your answer to the nearest dollar (\$).

(2)

- (b) Copy and complete the following cumulative frequency table and use it to draw a cumulative frequency graph. Use a scale of 2 cm to represent \$20 on the horizontal axis, and 2 cm to represent 20 customers on the vertical axis.

(5)

Money in \$ (d)	<20	<40	<60	<80	< 100	< 120	< 140
Number of customers (n)	24	40					

- (c) The time t (minutes), spent by customers in the store may be represented by the equation

$$t = 2d^{\frac{2}{3}} + 3.$$

- (i) Use this equation and your answer to part (a) to estimate the mean time in minutes spent by customers in the store.

(3)

- (ii) Use the equation and the cumulative frequency graph to estimate the number of customers who spent more than 37 minutes in the store.

(5)

(Total 15 marks)

5. The table shows the scores of competitors in a competition.

Score	10	20	30	40	50
Number of competitors with this score	1	2	5	k	3

The mean score is 34. Find the value of k .

<i>Working:</i>	<i>Answer:</i>
-----------------	-------------------------

(Total 4 marks)

6. A survey is carried out to find the waiting times for 100 customers at a supermarket.

waiting time (seconds)	number of customers
0–30	5
30– 60	15
60– 90	33
90 –120	21
120–150	11
150–180	7
180–210	5
210–240	3

- (a) Calculate an estimate for the mean of the waiting times, by using an appropriate approximation to represent each interval.

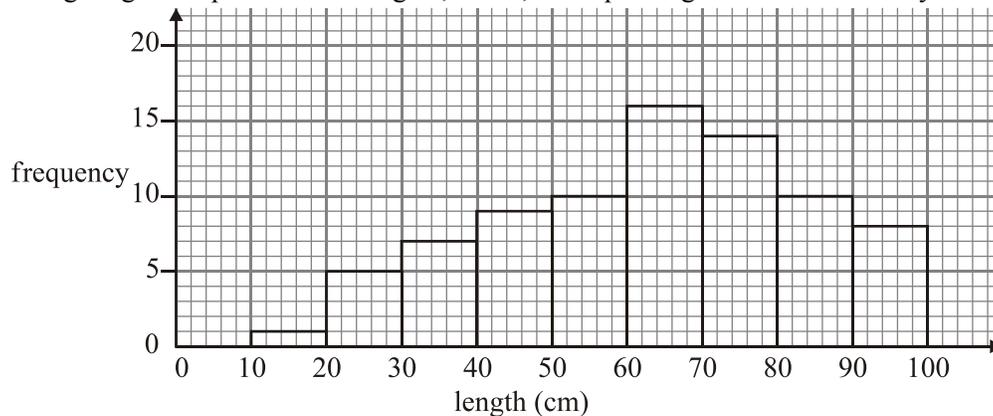
(2)

- (b) Construct a cumulative frequency table for these data.

- (c) Use the cumulative frequency table to draw, on graph paper, a cumulative frequency graph, using a scale of 1 cm per 20 seconds waiting time for the horizontal axis and 1 cm per 10 customers for the vertical axis. (1)
- (d) Use the cumulative frequency graph to find estimates for the median and the lower and upper quartiles. (4)

(3)
(Total 10 marks)

7. The following diagram represents the lengths, in cm, of 80 plants grown in a laboratory.



- (a) How many plants have lengths in cm between
 - (i) 50 and 60?
 - (ii) 70 and 90?(2)
- (b) Calculate estimates for the mean and the standard deviation of the lengths of the plants. (4)
- (c) Explain what feature of the diagram suggests that the median is different from the mean. (1)
- (d) The following is an extract from the cumulative frequency table.

length in cm less than	cumulative frequency
50	22
60	32
70	48
80	62

Use the information in the table to estimate the median. Give your answer to **two** significant figures.

(3)
(Total 10 marks)

8. Given the following frequency distribution, find

- (a) the median;
- (b) the mean.

Number (x)	1	2	3	4	5	6
Frequency (f)	5	9	16	18	20	7

Working: _____

--

Answers:

(a)

(b)

(Total 4 marks)

9. The table below represents the weights, W , in grams, of 80 packets of roasted peanuts.

Weight (W)	$80 < W \leq 85$	$85 < W \leq 90$	$90 < W \leq 95$	$95 < W \leq 100$	$100 < W \leq 105$	$105 < W \leq 110$	$110 < W \leq 115$
Number of packets	5	10	15	26	13	7	4

(a) Use the midpoint of each interval to find an estimate for the standard deviation of the weights.

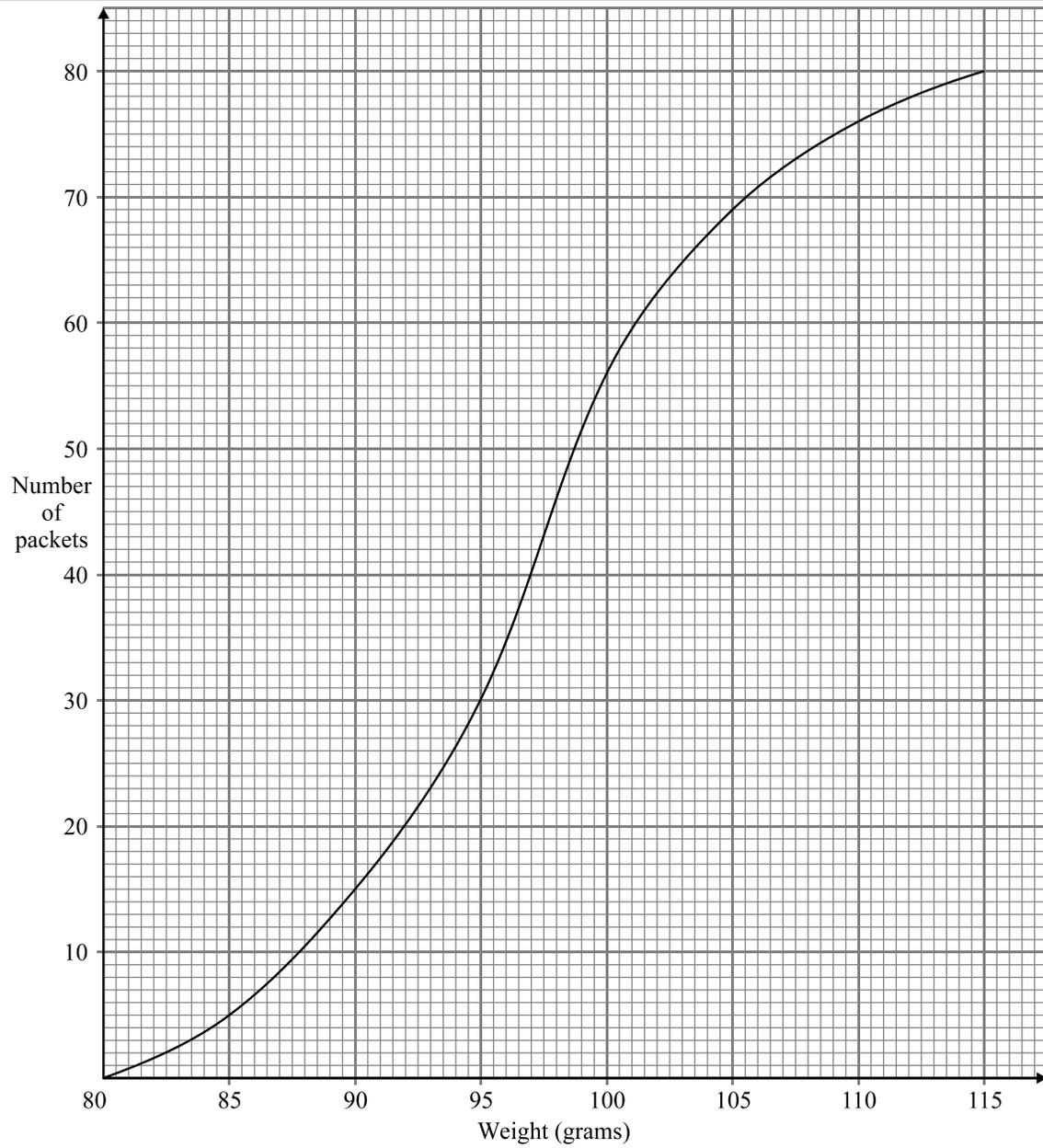
(3)

(b) Copy and complete the following cumulative frequency table for the above data.

Weight (W)	$W \leq 85$	$W \leq 90$	$W \leq 95$	$W \leq 100$	$W \leq 105$	$W \leq 110$	$W \leq 115$
Number of packets	5	15					80

(1)

(c) A cumulative frequency graph of the distribution is shown below, with a scale 2 cm for 10 packets on the vertical axis and 2 cm for 5 grams on the horizontal axis.



Use the graph to estimate

- (i) the median;
- (ii) the upper quartile (that is, the third quartile).

Give your answers to the nearest gram.

(4)

- (d) Let W_1, W_2, \dots, W_{80} be the individual weights of the packets, and let \bar{W} be their mean. What is the value of the sum

$$(W_1 - \bar{W}) + (W_2 - \bar{W}) + (W_3 - \bar{W}) + \dots + (W_{79} - \bar{W}) + (W_{80} - \bar{W})?$$

(2)

- (e) One of the 80 packets is selected at random. Given that its weight satisfies $85 < W \leq 110$, find the probability that its weight is greater than 100 grams.

(4)

(Total 14 marks)

10. The speeds in km h^{-1} of cars passing a point on a highway are recorded in the following table.

Speed v	Number of cars
$v \leq 60$	0
$60 < v \leq 70$	7
$70 < v \leq 80$	25
$80 < v \leq 90$	63
$90 < v \leq 100$	70
$100 < v \leq 110$	71
$110 < v \leq 120$	39
$120 < v \leq 130$	20
$130 < v \leq 140$	5
$v > 140$	0

- (a) Calculate an estimate of the mean speed of the cars.

(2)

- (b) The following table gives some of the cumulative frequencies for the information above.

Speed v	Cumulative frequency
$v \leq 60$	0
$v \leq 70$	7
$v \leq 80$	32
$v \leq 90$	95
$v \leq 100$	a
$v \leq 110$	236
$v \leq 120$	b
$v \leq 130$	295
$v \leq 140$	300

- (i) Write down the values of a and b .
- (ii) On graph paper, construct a cumulative frequency **curve** to represent this information. Use a scale of 1 cm for 10 km h^{-1} on the horizontal axis and a scale of 1 cm for 20 cars on the vertical axis.

(5)

- (c) Use your graph to determine

- (i) the percentage of cars travelling at a speed in excess of 105 km h^{-1} ;
- (ii) the speed which is exceeded by 15% of the cars.

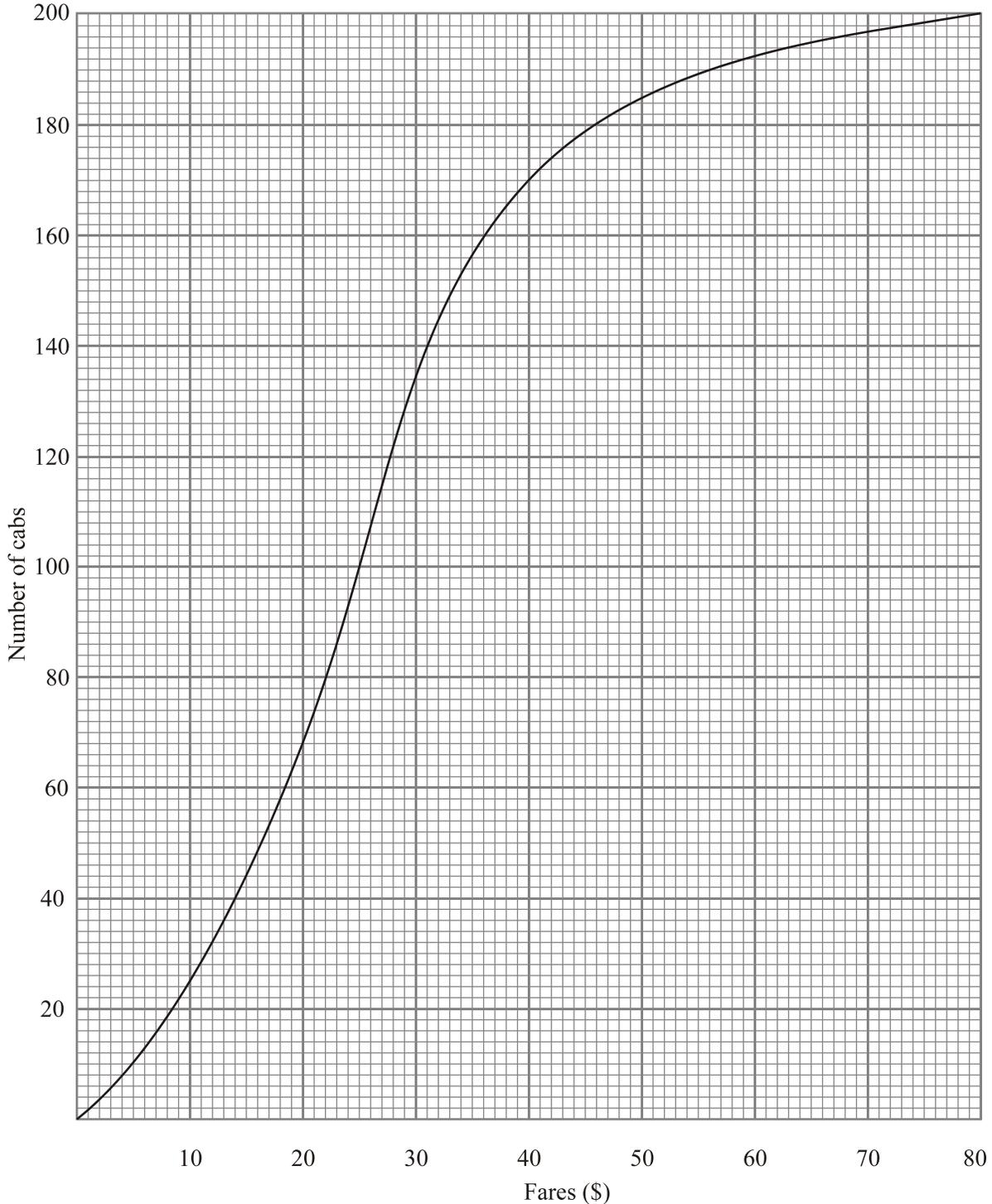
(4)
(Total 11 marks)

11. From January to September, the mean number of car accidents per month was 630. From October to December, the mean was 810 accidents per month.
 What was the mean number of car accidents per month for the whole year?

<i>Working:</i>	<i>Answer:</i>
-----------------	----------------

(Total 6 marks)

12. A taxi company has 200 taxi cabs. The cumulative frequency curve below shows the fares in dollars (\$) taken by the cabs on a particular morning.



- (a) Use the curve to estimate
 - (i) the median fare;
 - (ii) the number of cabs in which the fare taken is \$35 or less.
- (2)

The company charges 55 cents per kilometre for distance travelled. There are no other charges. Use the curve to answer the following.

- (b) On that morning, 40% of the cabs travel less than a km. Find the value of a .
- (4)
- (c) What percentage of the cabs travel more than 90 km on that morning?
- (4)

(Total 10 marks)

13. Three positive integers a , b , and c , where $a < b < c$, are such that their median is 11, their mean is 9 and their range is 10. Find the value of a .

<i>Working:</i>	<i>Answer:</i>

(Total 6 marks)

14. In a suburb of a large city, 100 houses were sold in a three-month period. The following **cumulative frequency table** shows the distribution of selling prices (in thousands of dollars).

Selling price P (\$1000)	$P \leq 100$	$P \leq 200$	$P \leq 300$	$P \leq 400$	$P \leq 500$
Total number of houses	12	58	87	94	100

- (a) Represent this information on a cumulative frequency **curve**, using a scale of 1 cm to represent \$50000 on the horizontal axis and 1 cm to represent 5 houses on the vertical axis.
- (4)
- (b) Use your curve to find the interquartile range.
- (3)

The information above is represented in the following frequency distribution.

Selling price P (\$1000)	$0 < P \leq 100$	$100 < P \leq 200$	$200 < P \leq 300$	$300 < P \leq 400$	$400 < P \leq 500$
Number of houses	12	46	29	a	b

- (c) Find the value of a and of b .
- (2)
- (d) Use mid-interval values to calculate an estimate for the mean selling price.
- (2)
- (e) Houses which sell for more than \$350000 are described as *De Luxe*.
 - (i) Use your graph to estimate the number of *De Luxe* houses sold. Give your answer to the nearest integer.
 - (ii) Two *De Luxe* houses are selected at random. Find the probability that **both** have a selling price of more than \$400000.

(4)
(Total 15 marks)

15. The number of hours of sleep of 21 students are shown in the frequency table below.

Hours of sleep	Number of students
4	2
5	5
6	4
7	3
8	4
10	2
12	1

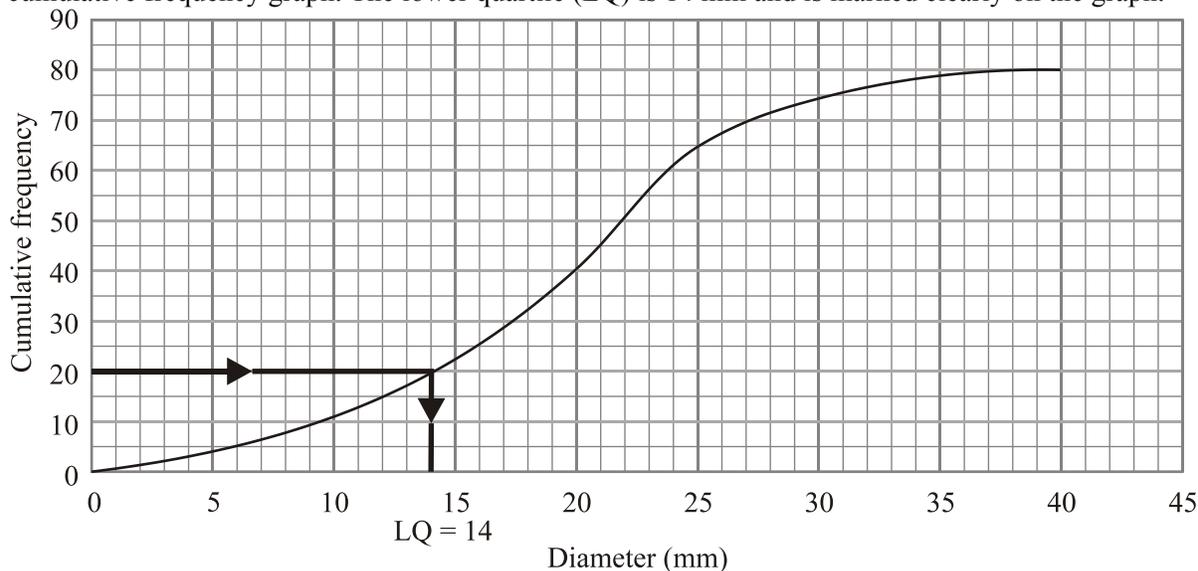
Find

- (a) the median;
- (b) the lower quartile;
- (c) the interquartile range.

<p><i>Working:</i></p>	<p><i>Answers:</i></p> <p>(a)</p> <p>(b)</p> <p>(c)</p>
------------------------	---

(Total 6 marks)

16. A student measured the diameters of 80 snail shells. His results are shown in the following cumulative frequency graph. The lower quartile (LQ) is 14 mm and is marked clearly on the graph.

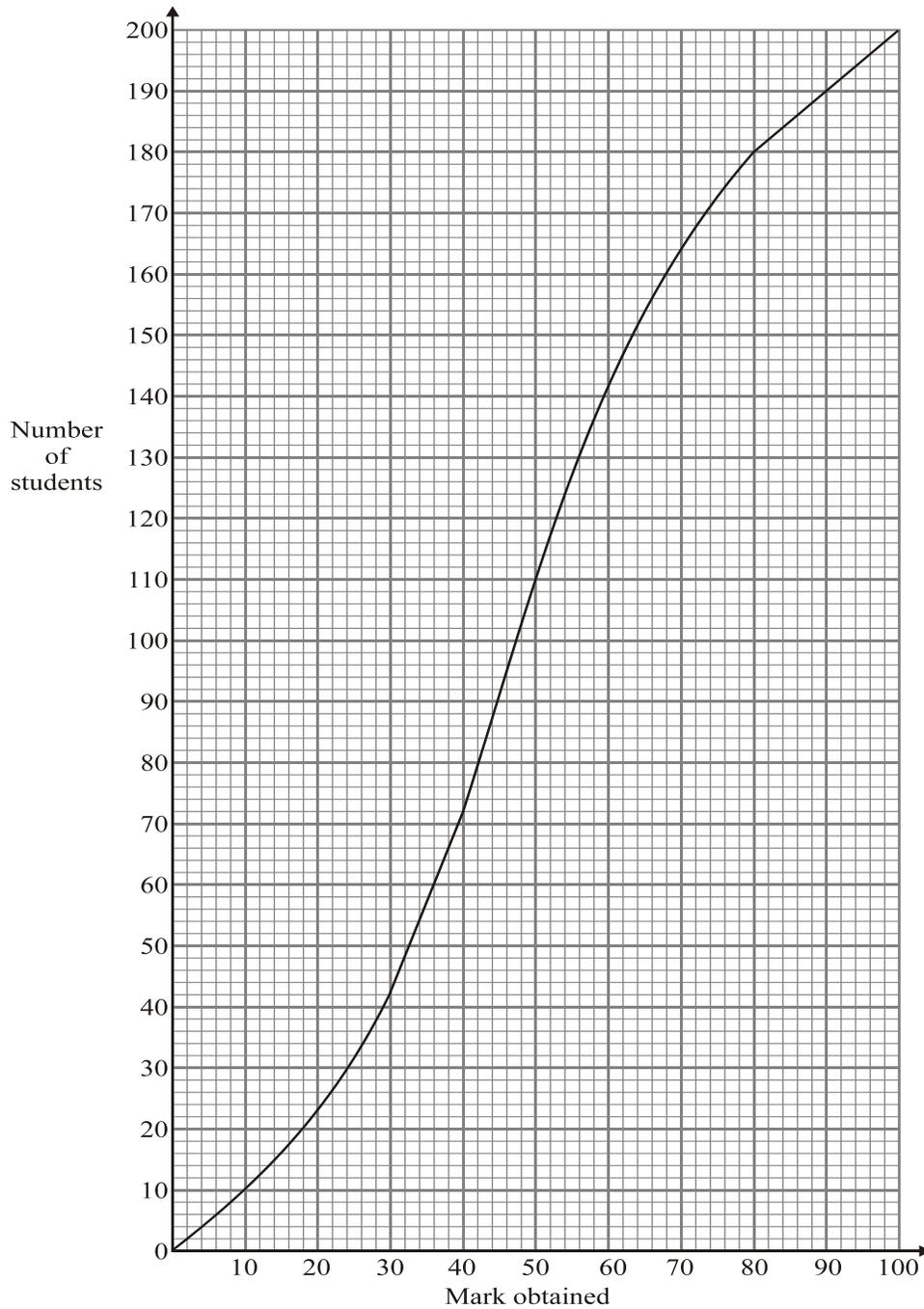


- (a) On the graph, mark clearly in the same way and write down the value of
 - (i) the median;
 - (ii) the upper quartile.
- (b) Write down the interquartile range.

<p><i>Working:</i></p>	<p><i>Answer:</i></p> <p>(b)</p>
------------------------	--

(Total 6 marks)

17. The cumulative frequency curve below shows the marks obtained in an examination by a group of 200 students.



- (a) Use the cumulative frequency curve to complete the frequency table below.

Mark (x)	$0 \leq x < 20$	$20 \leq x < 40$	$40 \leq x < 60$	$60 \leq x < 80$	$80 \leq x < 100$
Number of students	22				20

- (b) Forty percent of the students fail. Find the pass mark.

<i>Working:</i>	<i>Answer:</i>
	(b)

(Total 6 marks)

18. The table below shows the marks gained in a test by a group of students.

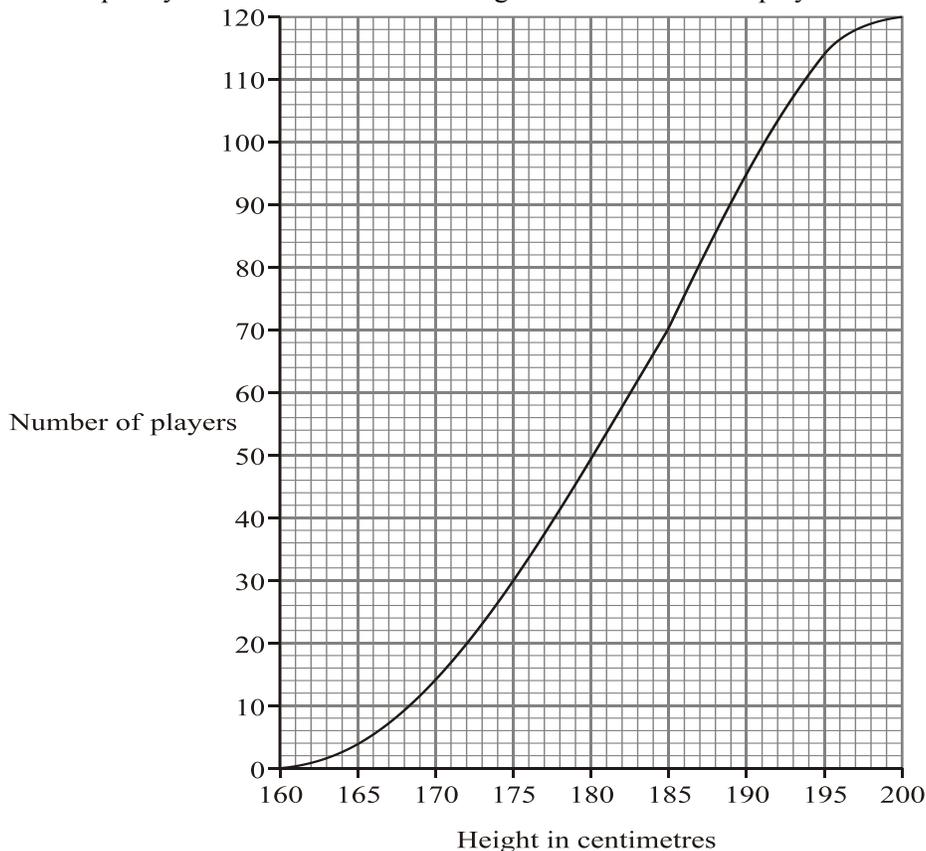
Mark	1	2	3	4	5
Number of students	5	10	p	6	2

The median is 3 and the mode is 2. Find the **two** possible values of p .

<i>Working:</i>	<i>Answer:</i>
-----------------	-------------------------

(Total 6 marks)

19. The cumulative frequency curve below shows the heights of 120 basketball players in centimetres.



Use the curve to estimate

- (a) the median height;
- (b) the interquartile range.

<i>Working:</i>	<i>Answers:</i> (a) (b)
-----------------	---

(Total 6 marks)

20. Let a, b, c and d be integers such that $a < b, b < c$ and $c = d$.
 The mode of these four numbers is 11.
 The range of these four numbers is 8.
 The mean of these four numbers is 8.
 Calculate the value of each of the integers a, b, c, d .

<i>Working:</i>	<i>Answers:</i> $a =$, $b =$ $c =$, $d =$
-----------------	---

(Total 6 marks)