STATISTICS (Q 7, PAPER 2)

2011

7. (a) Calculate the mean of the numbers 8, 6, 1, 3, 7, 8, 2.

(b) An information evening was held at a school. The number of people who entered the school during 20 minute intervals, beginning at 18:00, is given in the following table:

Time	Number of people
18:00 – 18:20	35
18:20 – 18:40	55
18:40 – 19:00	190
19:00 – 19:20	140
19:20 – 19:40	110
19:40 – 20:00	70

[Note: 18:20 – 18:40 means 18:20 or later, but before 18:40, etc.]

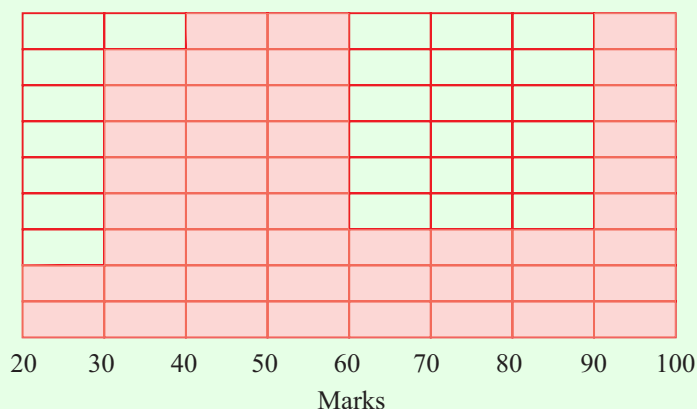**(i)** Copy and complete the following cumulative frequency table:

Time	Number of people
Before 18:20	
Before 18:40	
Before 19:00	
Before 19:20	
Before 19:40	
Before 20:00	

(ii) Draw the cumulative frequency curve (ogive).

(iii) Use your curve to estimate the interquartile range.

PTO

(c) The histogram represents the marks obtained by candidates in an examination.

20   30   40   50   60   70   80   90   100

Marks

(i) Copy and complete the following frequency table:

Marks	20 – 30	30 – 40	40 – 60	60 – 90	90 – 100
Number of candidates	4				

(ii) The mean mark was 60. Taking the mid-interval values of the completed frequency table, find the standard deviation, correct to the nearest integer.

(iii) Find the maximum possible number of candidates whose marks were within one standard deviation of the mean.

SOLUTION

7 (a)

$$\overline{x} = \frac{x_1 + x_2 + \ldots\ldots\ldots + x_N}{N} = \frac{\text{Sum of the Numbers}}{\text{Number of Numbers}} = \frac{\sum x}{N}$$

$$\overline{x} = \frac{8 + 6 + 1 + 3 + 7 + 8 + 2}{7} = \frac{35}{7} = 5$$

7 (b) (i)

Construct a cumulative frequency table:

There are 35 people who arrived before 18.20.

There are 90 people $(35 + 55)$ who arrived before 18.40.

There are 280 people $(35 + 55 + 190)$ who arrived before 19.00. And so on.....

Time	Number of people
Before 18:20	35
Before 18:40	90
Before 19:00	280
Before 19:20	420
Before 19:40	530
Before 20:00	600

7 (b) (ii), (iii)

No. of people (vertical axis) vs Times (horizontal axis)
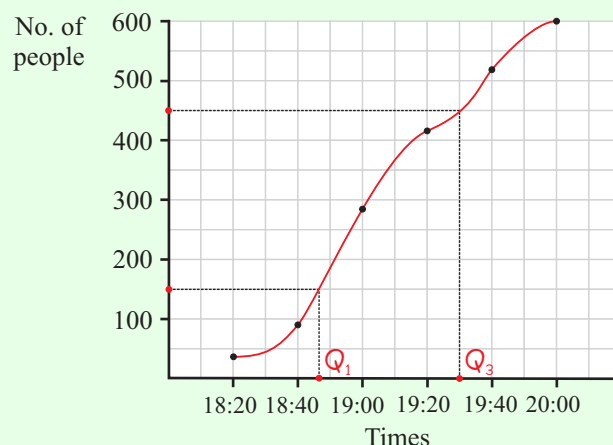
The lower quartile (Q_1): Go to 150 on the vertical axis (one-quarter of the people). The lower quartile is 18:47.

The upper quartile (Q_3): Go to 450 on the vertical axis (three-quarters of the people). The upper quartile is 19:30.

The interquartile range: $Q_3 - Q_1 = 19:30$ hours $- 18:47$ hours $= 43$ minutes.

7 (c) (i)

The first number is filled in for the range $20 - 30$ marks. You can see from the histogram that each rectangular block represents 2 candidates.

Marks	20 – 30	30 – 40	40 – 60	60 – 90	90 – 100
Number of candidates	**4**	16	36	18	18

7 (c) (ii)

Fill out the table below as shown. x represents mid-interval values. For example, in the range $20 - 30$ add the values and divide by 2 to get a value of 25.

The mean $\bar{x} = 60$.

The deviation, d, is given by the formula: $d = (x - \bar{x}) = (\text{Number} - \text{Mean})$.

To work out d, get the difference between each number x and the mean $\bar{x}$.

x	f	fx	d	d^2	fd^2
25	4	100	-35	1225	4900
35	16	560	-25	625	10000
50	36	1800	-10	100	3600
75	18	1350	15	225	4050
95	18	1710	35	1225	22050
	92	5520			44600

$$\sigma = \sqrt{\frac{\sum fd^2}{\sum f}}$$

f: Frequency
σ: Standard deviation

$$\sigma = \sqrt{\frac{44600}{92}} = 22$$

7 (c) (iii)

Marks	20 – 30	30 – 40	40 – 60	60 – 90	90 – 100
Number of candidates	4	**16**	**36**	**18**	18

A standard deviation of 22 added to and subtracted from 60 puts you in the range of marks highlighted in the table.

Maximum possible number of students $= 16 + 36 + 18 = 70$