## STATISTICS (Q 7, PAPER 2)

# 1997

7 (a) The table shows the distribution of ages of a group of 100 people.

| Age (in years) | 0 – 10 | 10 – 20 | 20 – 30 | 30 – 50 | 50 – 80 |
|---|---|---|---|---|---|
| Number of people | 10 | 19 | 25 | 30 | 16 |

[Note that 10 – 20 means that 10 is included but 20 is not, etc.]

Taking 5, 15, etc. as mid-interval values, estimate the mean age of the people in the group.

(b) $\{2, 5, 6, 4.5, 2.5\}$

Show that 4 is the mean of this set of numbers.
Then, calculate the standard deviation, correct to one place of decimals.

(c) A new shop opened at 0900 hours. During the first hour of trading, customers were counted as they entered the shop. The following cumulative frequency table shows the number of customers who has entered before the given times:

| Time | 0910 | 0920 | 0930 | 0940 | 0950 | 1000 |
|---|---|---|---|---|---|---|
| No. of customers | 45 | 69 | 95 | 120 | 144 | 250 |

(i) Draw a cumulative frequency curve.

(ii) A photograph was taken of the 100th. customer as he or she entered the shop. Use your curve to estimate the time at which the photograph was taken.

(iii) Use your curve to estimate the number of customers who entered the shop during the 15 minutes immediately after the photograph was taken.

## SOLUTION

**7 (a)**

Draw up a frequency table using the mid-interval values. To get a mid-interval value add the two numbers together and divide by 2.
**Ex**. Class interval: 30 – 50

Mid-interval value: $\dfrac{30+50}{2} = 40$

$$\boxed{\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \text{.......} + f_N x_N}{f_1 + f_2 + \text{.........} + f_N} = \frac{\sum fx}{\sum f}} \quad \text{.......} \; ❷$$

| x | f | f x |
|---|---|---|
| 5 | 10 | 50 |
| 15 | 19 | 285 |
| 25 | 25 | 625 |
| 40 | 30 | 1200 |
| 65 | 16 | 1040 |
|  | 100 | 3200 |

Mean price: $\bar{x} = \dfrac{\sum fx}{\sum f} = \dfrac{3200}{100} = 32$

**7 (b)**

> **STEPS**
> **1**. Find the mean.
> **2**. Draw up a table of $x$, $d$ and $d^2$.
> **3**. Apply the standard deviation formula.

**1.**

> The **MEAN** or **AVERAGE** of a set of numbers is calculated by adding the numbers together and dividing by the number of numbers.
>
> $$\text{Mean} = \frac{\text{Sum of the numbers}}{\text{Number of numbers}}$$
>
> The mean is denoted by $\overline{x}$.

$$\overline{x} = \frac{2+5+6+4.5+2.5}{5}$$

$$\overline{x} = \frac{x_1 + x_2 + \dots\dots\dots + x_N}{N} = \frac{\text{Sum of the Numbers}}{\text{Number of Numbers}} = \frac{\sum x}{N} \quad \dots\dots \quad \textbf{1}$$

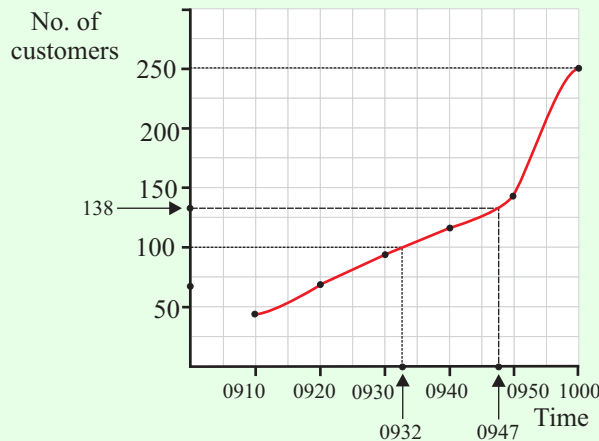$$\Rightarrow \overline{x} = \frac{20}{5} = 4$$

**2.**

> The deviation, $d$, is given by the formula:
> $d = (x - \overline{x}) = (\text{Number} - \text{Mean})$.
> To work out $d$, get the difference between each number, $x$, and the mean, $\overline{x}$.

| $x$ | $d$ | $d^2$ |
|-----|------|-------|
| 2 | $-2$ | 4 |
| 5 | 1 | 1 |
| 6 | 2 | 4 |
| 4.5 | 0.5 | 0.25 |
| 2.5 | $-1.5$ | 2.25 |
| | | 11.5 |

**3.** $\sigma = \sqrt{\dfrac{11.5}{5}} = 1.5$

$$\sigma = \sqrt{\frac{\text{Sum of (Deviations)}^2}{\text{Number of numbers}}} = \sqrt{\frac{\sum d^2}{N}} \quad \dots\dots \quad \textbf{4}$$

**7 (c) (i)**



**7 (c) (ii)**

Go to 100 on the vertical axis. The corresponding value on the horizontal axis is a time of 0932.

**7 (c) (iii)**

15 minutes after 0932 is 0947. Go to 0947 on the horizontal axis. The corresponding value on the vertical axis is 138 customers. Therefore, the number of customers who entered the shop between 0932 and 0947 is 138 – 100 = 38 customers.