## The Frequency Polygon

An easy way to think of a frequency polygon is to imagine placing a dot in the middle of the top of each bar in a histogram and connecting the dots with straight lines (and then getting rid of the bars), as shown in Figure 2.3a

The polygon is connected to the $X$ axis at either end to form a closed figure.

distribution should be same. If no are not same than frequency of both group has to be converted into percentage (\%)

find the total no of student
Thus, the area enclosed by the polygon is clearly defined and can be used in ways to be described later.

## The Cumulative Frequency Polygon (ogive)

The graph of a cumulative frequency distribution is known as cumulative frequency curve or ogive. Since there are two types of cumulative frequency distribution e.g., " less than" and " more than" cumulative frequencies, we can have two types of ogives
i) 'Less than' Ogive: In 'less than' ogive, the less than cumulative frequencies are plotted against the upper real limit of each value (e.g., 65.5, instead of 65 inches). The
rationale is that to make sure you have accumulated, for instance, all of the heights labelled 65 inches, you have to include all measurements up to 65.5 inches

It is an increasing curve having slopes upwards from left to right. It can stay the same, if the next value has a zero frequency.


If you convert the cumulative frequencies to cumulative percentages, the ogive can be used to estimate percentiles and PRs (Percentile Rank) not in your original table (falls between two). Move straight up from a score until you hit the curve and then horizontally to the left until you hit the $Y$ axis to find the PR of the score. Reversing this procedure allows you to estimate percentiles.

Find the percentile rank for the score 6 and find the score at the 70th percentile,

Drawing the graph to a larger scale tends to increase accuracy.
ii) 'More than' Ogive: In more than ogive, the more than cumulative frequencies are plotted against the lower class boundaries of the respective classes. It is decreasing curve and slopes downwards from left to right

| $\boldsymbol{X}$ | $\boldsymbol{f}$ | $\boldsymbol{c f}$ (less than) | $\boldsymbol{c p f}$ | $\boldsymbol{c f}($ more than) | $\boldsymbol{c p f}$ |
| :--- | :--- | :---: | :---: | :---: | :---: |
| 10 | 2 | 25 | $100 \%$ | 2 |  |
| 9 | 2 | 23 | 92 | 4 | $8 \%$ |
| 8 | 5 | 21 | 84 | 9 | 16 |
| 7 | 3 | 16 | 64 | 12 | 36 |
| 6 | 7 | 13 | 52 | 19 | 48 |
| 5 | 1 | 6 | 24 | 20 | 76 |
| 4 | 4 | 5 | 20 | 24 | 80 |
| 3 | 0 | 1 | 4 | 24 | 96 |

Plot the ogive (more than type) by the data given above

| table 2.7 | 98 | 96 | 93 | 92 | 92 | 89 | 89 | 88 | 86 | 86 | 86 | 85 | 85 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  | 84 | 83 | 81 | 81 | 81 | 81 | 79 | 75 | 75 | 72 | 68 | 64 |  |

There are too many different values; the simple frequency distribution obviously would not be very helpful in this case.

The solution is to group the possible score values into equal-sized ranges, called class intervals. A table that shows the frequency for each class interval is called a grouped frequency distribution.

| table 2.8 |  |  |  |
| :--- | :---: | :---: | :---: |
| Class |  | Class |  |
| Interval |  | Interval |  |
| $\boldsymbol{X}$ | $\boldsymbol{f}$ | $\boldsymbol{X}$ | $\boldsymbol{f}$ |
| $95-99$ | 2 | $75-79$ | 3 |
| $90-94$ | 3 | $70-74$ | 1 |
| $85-89$ | 8 | $165-69$ | 1 |
| $80-84$ | 6 | $60-64$ | 1 |

By focusing on just one class interval from Table 2.8—for example, 80-84. The interval is defined by its apparent limits

In particular, any height above 79.5 cm would be rounded to 80 cm and included in the interval $80-84$. Similarly, any height below 84.5 cm would be rounded to 84 and also included in the interval 80-84. Therefore, the real limits of the class interval are 79.5 (lower real limit) and 84.5 (upper real limit)

Bear in mind that even though there appears to be a gap when you look at the apparent limits (80-84, 85-89), the gap disappears when you look at the real limits (79.5-84.5, 84.5-89.5) and yet there is still no overlap.

What happens if a score turns out to be exactly 84.5. First, when dealing with a continuous scale, the probability of any particular exact value (e.g., 84.500) arising is considered to be too small to worry about

