

- Ref. ch. 1, Sec. 1.4, 1.5.
- Frequency: Number of occurrences of each value of a variable in a set of data is referred to as 'frequency' of that variate value.
- Frequency distribution: A frequency distribution is a listing of classes of possible values for a variable, together with a tabulation of the no. of observations in that class. Frequency distribution table summarizes a large set of data.
- Relative frequency: Relative frequency for a class of values of a variable is the proportion of observations that fall in that class. It is a number between 0 and 1.

$$\text{Relative frequency} = \frac{\text{Frequency}}{\text{Total frequency}} \text{ and Percent frequency} = 100 \times \text{rel. freq.}$$

- Relative frequency distribution: It is a listing of class-intervals of possible values of a variable, together with a tabulation of the proportion of observations in that class.

Ex. Vide Table 1.10, p. 26 and Table 1.12, p. 27.

Example: Considering 'word length' as variable construct the relative frequency distribution table for word-length from the following quotation of John Keats. Also obtain the percent frequency for each variate value: "Beauty is truth, truth beauty, — that is all ye know on earth, and all ye need to know."

- Diagrammatic representation of frequency distribution: Stemplots and Histograms are the most popular diagrams for representing the distribution. Stemplot (or Stem-and-leaf plots) gives a quick idea of the shape of a distribution for a small data set.

Procedure: (1) Select leading digits for the stem-values. The trailing digit or digits become the leaves. (2) List possible stem-values in a vertical column.

(3) Record the leaf-values for every observation beside the corresponding stem-value. Arrange the leaves in order on each line from smallest to largest.

Ex. Vide Ex. 1.7, p. 21, Ex. 1.8, pp. 21-22.

Histogram: A diagrammatic representation of frequency distribution.

Construction of relative frequency histogram for continuous data for equal class-widths is described below:

1. Mark class boundaries on horizontal axis
2. Use relative frequencies on the vertical axis
3. Draw the rectangle for each class directly above the corresponding class interval (so that the edges are at the class boundaries).

Ex. Vide Table 1.10 and Figure 1.15, p. 26.

From the histogram we can have an idea about the shape of the distribution:

- A distribution is symmetric if the left and right sides of the distribution form mirror images, when divided at the middle value.
- A distribution is skewed to the right if greater proportion of observations lie to the right of the peak value.
- A distribution is skewed to the left if greater proportion of observations lie to the left of the peak value.
- A distribution is unimodal if it has one peak.
- A distribution is bimodal if it has two peaks.