CHAPTER 14
STATISTICS

NOTES

- **Primary Data:** If an investigator collects data with a specific purpose, then it is called a primary data.
- **Secondary Data:** If a investigator can get a data indirectly from some other source, then the data is called a secondary data.
- **Range** = maximum observation – minimum observation
- **Measures of Central Tendency:**
  - **Mean of ungrouped data:** If $x_1, x_2, x_3, \dots\dots\dots\dots\dots\dots, x_n$ are different observations with respective frequencies $f_1, f_2, f_3, \dots\dots\dots\dots\dots\dots, f_n$, then

    $$\text{Mean} = \frac{1}{N}(x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots\dots\dots\dots\dots + x_n f_n)$$

    $$= \frac{1}{N}\sum_{i=1}^{n} x_i f_i, \text{ where N= total number of observations}$$

    $$= f_1 + f_2 + f_3 + \dots\dots\dots\dots\dots + f_n$$

  - **Median of an ungrouped data:**

    The observations are first arranged in ascending or descending order.

    If the number of observations N is odd, then $\left(\frac{N+1}{2}\right)^{th}$ observation is the median.

    If the number of observations N is even, then $\frac{\left(\frac{N}{2}\right)^{th}\text{ term} + \left(\frac{N}{2}+1\right)^{th}\text{ term}}{2}$ is the median.

  - **Mode of an ungrouped data:** The observation with the maximum frequency is the mode.

- **Karl Pearson's Empirical Formula:**

  Mean – Mode = 3(Mean – Median)

Note: In case of a distribution known as the normal distribution, the three measures of central tendency mean, median and mode are coincident.

Histogram

Representation of a classified data by a graph consisting of rectangles is called the histogram of the data.

Steps to draw the histogram of data:

1. The class limits must be represented along the x-axis on a suitable scale.
2. We construct rectangles whose bases are the class limits and whose areas are proportional to the frequency of the class.

Frequency Polygon

The frequency polygon of a classified data is obtained first by plotting the points whose abscissa are the class marks of the various classes and the ordinates are the frequencies of the corresponding classes and then joining them by line segments.

If we are to draw both histogram and the frequency polygon we get the latter by joining the mid-points of the tops of all rectangles by line segments.

Cumulative Frequency Curve or the Ogive

For a classified data the free-hand smooth curve obtained by joining the points which are plotted by taking the upper limits of the various classes as the abscissa and the cumulative frequencies of the corresponding classes as the ordinates is called the cumulative frequency curve or the ogive of the data.

An ogive is a never descending curve.

****