Statitics

Measures of Dispersion

This gives a measure of the dispersion of the observation around the measure of central tendency of the data collected.

Range

Range of a data = Maximum value – Minimum value.

Mean Deviation

Mean deviation of central tendancy 'a' of a data is given by

$$M.D(a) = \frac{\text{Sum of absolute values of deviations from 'a'}}{\text{Number of observations}}$$

- ### Ungrouped data

  Mean deviation for the data $x_1, x_2, x_3, ..., x_n$ is

  $$M.D(a) = \frac{\sum_{i=1}^{n} |x_i - a|}{n}$$

  - $Mean(\bar{x}) = \frac{\sum_{i=1}^{n} x_i}{n}$
  - Median is $\frac{n+1}{2}^{th}$ term when n is odd ,mean of $\frac{n}{2}^{th}$ and $\frac{n}{2} + 1^{th}$ term when n is even

- ### Grouped Discrete frequency distribution

  Mean deviation for the data $x_1, x_2, ..., x_n$ occurring with frequencies $f_1, f_2, ..., f_n$ respectively is

  $$M.D(a) = \frac{\sum_{i=1}^{n} f_i|x_i - a|}{\sum_{i=1}^{n} f_i}$$

  - $Mean(\bar{x}) = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$
  - To find the median form a column for cumulative frequencies

- ### Grouped Continuous frequency distribution

  A continuous frequency distribution is a series in which the data are classified into different class-intervals without gaps alongwith their respective frequencies.

  To find the Mean deviation about a central tendancy 'a' , choose x_i as midpoint of the intervals and convert it into Discrete.

  - $Mean(\bar{x}) = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$
  - median $= l + \frac{c(\frac{N}{2} - f_0)}{f_1}$
    l -lower limit of the median class
    f_0- Cumulative frequency of the class preceding the median class.
    f_1 - Cumulative frequency of the median class.
    c - width of the interval
    $N - \sum_{n}^{i=0} f_i$

Shortcut method for finding the mean of frequency distribution

1. Select an assumed mean 'a' (value close to middle)
2. substract assumed mean from eaach x_i
3. choose appropriate h and divide each $x_i - a$ by h and define it as d_i
4. find mean using the fromula

$$\bar{x} = a + \frac{\sum_{n}^{i=1} f_i d_i}{N} \times h$$

Variance and Standard Deviation

standard deviation= $\sqrt{variance}$

Ungrouped Data

variance for the data $x_1, x_2, x_3, ..., x_n$ is $\sigma^2 = \frac{1}{n}\sum_{n}^{i=1}(x_i - \bar{x})^2$

Grouped frequency distribution

Variance for the data $x_1, x_2, ..., x_n$ occurring with frequencies $f_1, f_2, ..., f_n$ respectively is

$$\sigma^2 = \frac{1}{N}\sum_{n}^{i=1} f_i(x_i - \bar{x})^2$$

Another formula for standard deviation

$$\sigma = \frac{1}{N}\sqrt{N\sum_{n}^{i=1} f_i x_i^2 - (\sum_{n}^{i=1} f_i x_i)^2}$$

Shortcut method to find variance and standard deviation

Let the assumed mean be 'A' and the scale be reduced to $\frac{1}{h}$ times (h being the width of class-intervals). Let the step-deviations or the new values be y_i.

$$y_i = \frac{x_i - A}{h}$$

then

$$\sigma = \frac{h}{N}\sqrt{N\sum_{n}^{i=1} f_i y_i^2 - (\sum_{n}^{i=1} f_i y_i)^2}$$