CHAPTER 22

Statistics

ARITHMETIC MEAN

(*i*) **Arithmetic Mean for Unclassified (Ungrouped or Raw) Data:** If there are n observations, $x_1, x_2, x_3, \ldots, x_n$, then their arithmetic mean

$$A \text{ or } \bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

(*ii*) **Arithmetic Mean for Discrete Frequency Distribution or Ungrouped Frequency Distribution:** Let $f_1, f_2, \ldots, f_n$ be corresponding frequencies of $x_1, x_2, \ldots, x_n$. Then, arithmetic mean

$$A = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_n f_n}{f_1 + f_2 + \ldots + f_n} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i}$$

(*iii*) **Arithmetic Mean for Classified (Grouped) Data or Grouped Frequency Distribution:** For a classified data, we take the class marks $x_1, x_2, \ldots, x_n$ of the classes, then arithmetic mean by

$$A = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i}$$

Combined Mean: If $A_1, A_2, \ldots, A_r$ are means of $n_1, n_2, \ldots, n_r$ observations respectively, then arithmetic mean of the combined group is called the combined mean of the observation.

$$A = \frac{n_1 A_1 + n_2 A_2 + \ldots + n_r A_r}{n_1 + n_2 + \ldots + n_r} = \frac{\sum_{i=1}^{r} n_i A_i}{\sum_{i=1}^{r} n_i}$$

MEDIAN

Median for Simple Distribution or Raw Data

Firstly, arrange the data in ascending or descending order and then find the number of observations n.

(*a*) If n is odd, then $\left(\frac{n+1}{2}\right)$ th term is the median.

(*b*) If n is even, then there are two middle terms namely $\left(\frac{n}{2}\right)^{\text{th}}$ and $\left(\frac{n}{2}+1\right)^{\text{th}}$ terms, median is mean of these terms.

Median for Classified (Grouped) Data or Grouped Frequency Distribution

For a continuous distribution, median

$$M_d = l + \frac{\frac{N}{2} - C}{f} \times h$$

where, l = lower limit of the median class

f = frequency of the median class

N = total frequency = $\sum_{i=1}^{n} f_i$

C = cumulative frequency of the class just before the median class

h = length of the median class

Mode for Classified (Grouped) Distribution or Grouped Frequency Distribution

$$M_o = l + \frac{f_0 - f_1}{2f_0 - f_1 - f_2} \times h$$

where, l = lower limit of the modal class

f_0 = frequency of the modal class

f = frequency of the pre-modal class

f = frequency of the post-modal class

h = length of the class interval

Relation Between Mean, Median and Mode

(*i*) Mean – Mode = 3 (Mean – Median)

(*ii*) Mode = 3 Median – 2 Mean

MEAN DEVIATION (MD)

(*i*) For simple (raw) distribution, $\delta = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$

where, n = number of terms, $\bar{x} = A$ or M_d or M_o

(*ii*) For unclassified frequency distribution, $\delta = \dfrac{\sum_{i=1}^{n} f_i |x_i - \overline{x}|}{\sum_{i=1}^{n} f_i}$

(*iii*) For classified distribution, $\delta = \dfrac{\sum_{i=1}^{n} f_i |x_i - \overline{x}|}{\sum_{i=1}^{n} f_i}$

where, x_i is the class mark of the interval.

STANDARD DEVIATION AND VARIANCE

(*i*) For simple distribution

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}} = \frac{1}{n}\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

where, n is a number of observations and $\bar{x}$ is mean.

(*ii*) For discrete frequency distribution

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} f(x_i - \overline{x})^2}{N}} = \frac{1}{N}\sqrt{N\sum_{i=1}^{n} f_i\, x_i^2 - \left(\sum_{i=1}^{n} f_i x_i\right)^2}$$

(*iii*) For continuous frequency distribution

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

where, x_i is class mark of the interval.

Standard Deviation of the Combined Series

If n_1, n_2 are the sizes, $\overline{X}_1, \overline{X}_2$ are the means and σ_1, σ_2, are the standard deviation of the series, then the standard deviation of the combined series is

$$\sigma = \sqrt{\frac{n_1(\sigma_1^2 + d_1^2) + n_2\left(\sigma_2^2 + d_2^2\right)}{n_1 + n_2}}$$

where, $d_1 = \overline{X}_1 - \overline{X}, d_2 = \overline{X}_2 - \overline{X}$

and $\overline{X} = \dfrac{n_1\overline{X}_1 + n_2\overline{X}_2}{n_1 + n_2}$

IMPORTANT POINTS TO BE REMEMBERED

(*i*) The ratio of SD (σ) and the AM ($\overline{x}$) is called the coefficient of standard deviation $\left(\dfrac{\sigma}{\overline{x}}\right)$

(*ii*) The percentage form of coefficient of SD i.e. $\left(\dfrac{\sigma}{\overline{x}}\right)\times 100$ is called coefficient of variation.

(*iii*) The distribution for which the coefficient of variation is less is more consistent.

(*iv*) Standard deviation of first n natural numbers is $\sqrt{\dfrac{n^2 - 1}{12}}$.

(*v*) Standard deviation is independent of change of origin, but it depends on change of scale.