Chapter-3 (SOS - 4-th edition)

Measures of central tendency :-

Arithmetic mean :- $\bar{X} = \dfrac{\sum_{i=1}^{N} X_i}{N}$ — (1)

$\boxed{\text{Example-4}}$ → see → Page → 62

If $X_1, X_2, \ldots, X_k$ occur $f_1, f_2, \ldots f_k$ times, (frequencies),

arithmetic mean, $\bar{X} = \dfrac{\sum_{i=1}^{k} f_i X_i}{\sum_{i=1}^{k} f_i} = \dfrac{\sum fX}{N}$ — (2)

see → $\boxed{\text{example-5}}$ → Page → 62

when f_i's represents weighting factor, (2) is called a "weighted arithmetic mean".

see → $\boxed{\text{Example-6 \& 7}}$ → 62 & 63

If A is any gussed or assumed arithmetic mean,

~~and if, $d_j = X_j - A$ — and if~~

and if the deviations $d_j = X_j - A$, (1) & (2) becomes respectively,

$$\bar{X} = A + \frac{\sum_{j=1}^{N} d_j}{N} \quad \text{— (3)} \rightarrow \text{Un-grouped data}$$

and $$\bar{X} = A + \frac{\sum_{j=1}^{k} f_j d_j}{\sum_{j=1}^{k} f_j} = A + \frac{\sum fd}{N} \quad \text{— (4)}$$

→ Grouped data

see → $\boxed{\text{Examples: 3.13, 3.14, 3.15, 3.20,}}$ → Pages 72, 73 & 75.

If all class intervals have equal size c,

$$d_j = X_j - A = c u_j \Rightarrow u_j = \frac{X_j - A}{c} \; ; \text{ where, } u_j = 0, \pm 1, \pm 2, \ldots$$

$$\therefore (4) \Rightarrow \bar{X} = A + \left(\frac{\sum f u}{N}\right) c. \quad \text{——— (5)}$$

See → $\boxed{\text{Example : 3.22, 3.23 \& 3.24}}$ → page → 76 & 77.

Using (2) or (4) is called "<u>long method</u>".

& using (5) is called "<u>coding / short-cut method</u>"!

<u>Median</u> :- The median of a set of numbers arranged in order of magnitude (i.e, in array) is either the middle value (for odd numbers of ~~observa~~ data) ($\frac{n+1}{2}$th data; n = total no. of data)

OR,

the arithmetic mean of the two middle values.

[For even number of data,

$$\frac{1}{2}\left[\frac{n}{2}\text{th} + \left(\frac{n}{2}+1\right)\text{th}\right]$$]

~~For~~ ~~See~~ See → [Example - 8 & 9] → Page → 64

For grouped data,

$$\text{Median} = L_1 + \left(\frac{\frac{N}{2} - (\Sigma f)_1}{f_{median}} \right) C$$

where,

L_1 = lower class boundary of the median.
(i.e, the class containing the median)

N = Total number ~~no.~~ of data (i.e, total frequency).

$(\Sigma f)_1$ = sum of frequencies of all classes lower than the median class

f_{median} = frequency of the median class.

C = size of the median class, interval.

Geometrically the median is the value of x (abscissa) corresponding to the vertical line which divides a histogram into two parts having equal areas. (Sometimes denoted by $\tilde{x}$, x-tilde).

see → $\boxed{\text{Example} - 3.28 \ \& \ 3.30}$ → Page → 79 & 80.

~~Mode~~ **Mode** :- The mode of a set of numbers is that value which occurs with the greatest frequency; that is, it is the most common value. The mode may not exist, and even if it does exist it may not be unique.

see → $\boxed{\text{Examples :- 10, 11 \& 12}}$ → page → 64

3·31 → page → 80

A distribution having only one mode is called unimodal

The mode will be the value(s) of X corresponding to the maximum point (s) on the curve.

(Sometimes denoted by $\hat{X}$, X-cap).

For grouped data :-

$$\text{Mode} = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right) c$$

where, L_1 = Lower class boundary of the modal class.

Δ_1 = excess of modal frequency over frequency of next-lower class.

Δ_2 = " " " " " " of next-higher class

c = size of the modal class interval.

see → $\boxed{\text{Example – 3·3, SOS → 3rd edition book page → 76 / 89}}$

$\Delta_1 = 16 - 10 = 6$
$\Delta_2 = 16 - 14 = 2$

The empirical relation between the Mean, Median and Mode.

For unimodal frequency curves that are moderately skewed (asymmetrical),

$$\text{Mean} - \text{Mode} = 3(\text{Mean} - \text{Median})$$

page → 64, 65 & 83 (3·34)

For normal distribution curve, (symmetrical shape)

$$\text{Mean} = \text{Median} = \text{Mode}.$$

Geometric Mean, G :- → $\boxed{\text{page} \to 65}$ is the N-th root of ~~N~~ ~~positive~~ the product of N-positive numbers $X_1, X_2, \cdots, X_N$.

i.e, $G = \sqrt[N]{X_1 \cdot X_2 \cdots X_N}$

Example-13 ~~:- see~~ & 3.35 → see your self.

Harmonic Mean, H :- is the reciprocal (inverse) of the arithmetic mean of the reciprocal of the N numbers $X_1, X_2, \cdots, X_N$.

i.e, $H = \dfrac{1}{\dfrac{\sum \frac{1}{X}}{N}} \Rightarrow \dfrac{1}{H} = \dfrac{\sum \frac{1}{X}}{N}$

see → Example - 14, 3.39

The relation between the arithmetic, geometric & harmonic means.

$$H \leq G \leq \bar{X}$$ → see page → 66

Example → 15 → see

The Root mean square :- (RMS),

$$RMS = \sqrt{\overline{X^2}} = \sqrt{\frac{\Sigma X^2}{N}}$$

Example – 16 → see

Quartiles :- The values that divide the data into 4-equal parts. (Q_1, Q_2 & Q_3)

Deciles :- The values that divide the data into 10-equal parts. ($D_1, D_2, \ldots, D_9$)

Percentiles :- The values that divide the data into 100-equal parts. ($P_1, P_2, \ldots, P_{99}$).

Page → 66

The fifth decile & 50-th percentile correspond to the median.

The 25-th & 75-th percentiles correspond to the 1st & 3rd quartiles, respectively.

see → Example - 3.44 & 3.45 & 3.46