# Measures Of Central Tendency & Dispersion

## # Central Tendency

⇒ Meaning of central tendency
- → Central value of all observation
- → Representative of entire series

**Character of Good Central Tendency**

- → Easy to calculate → Mean & Mode
- → Easy to understand → Mean
- → Based on all observation → Mean, GM, HM
- → Rigidly defined → Mean
- → Least affected by extreme values = Median.
- → Have mathematical properties = Mean, GM & HM.

**Note:-** AM is Best Central Tendency
→ Median is best for open ended series.

## # Arithmetic Mean ($\overline{x}$)

$$AM = \frac{\text{Sum of all observation}}{\text{Total no. of observation}}$$

**Direct Method :-**

$\overline{x} = \Sigma f_i x_i / N$ [∵ where $N = \Sigma f_i$]

**Shortcut Method :-**

**Assumed mean method**

If $x_i - A = d_i$

Then $\overline{x} = A + \frac{\Sigma f_i d_i}{N}$

**Step-Deviation Method**

If $\frac{x_i - A}{h} = u_i$

Then $\overline{x} = A + \frac{\Sigma f_i u_i}{N} \times h$

- **Properties :**

1) If all observation are same (let K) then Mean also K ; Mean of 5, 5, 5, 5, 5 = 5.

2) Sum of Deviations from their arithmetic mean is always zero
i.e $\Sigma (x_i - \overline{x}) = 0$

3) Sum of the squares of deviation is minimum when deviations are taken from their arithmetic mean
i.e $\Sigma (x_i - K)^2$ is minimum, when $K = \overline{x}$

4) Combines mean

$$\overline{X}_{12} = \frac{N_1 \overline{X}_1 + N_2 \overline{X}_2}{N_1 + N_2}$$

5.) Arithmetic mean changes with change of origin & change of scale
i.e If $y_i = a x_i + b$
Then $\overline{y} = a\overline{x} + b$

$1 + 2 + 3 + \dots + n = n(n+1)/2$

$1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$

$1^3 + 2^3 + 3^3 + \dots + n^3 = \left[\frac{n(n+1)}{2}\right]^2$

A.M of first 'n' natural numbers $= \frac{n+1}{2}$

# Median

A no. which divides entire distribution in two parts is known as median. It represents half (50%) of total numbers.

50% median

It's not based on all observations. So, it can be used for open ended series.

**Methods**

For individual & descrete Series

Median = $\left(\frac{n+1}{2}\right)^{th}$ term

For Continuous series

→ Makes cf column
→ locate $\frac{N}{2}$ in cf
→ Select median class
→ Use formulas

$$\boxed{\text{Median} = l + \left[\frac{\frac{N}{2} - cf}{f}\right] \times h}$$

i.e $l$ = lower limit
$h$ = class size
$cf$ = Cumulative frequency

→ Median changes with change of origin & change in scale i.e if $\boxed{y = a + bx}$. Then median of $y = a + b$ (med. of $x$)

→ Sum of absolute deviation is minimum when deviations are taken from the median i.e $\sum |x_i - k|$ is minimum (when $k$ = median)

# Partition values (fractiles)

(1) Quartiles ($Q_1$, $Q_2$ & $Q_3$)

Divide entire series in 4 parts

(25%) $Q_1$ — $N/4$; (50%) $Q_2$ — $N/2$; $Q_3$ (75%) — $3N/4$

For Individual & Descrete series

$Q_1 = \left(\frac{N+1}{4}\right)^{th}$ term ; $Q_2 = \left(\frac{N+1}{2}\right)^{th}$ (Median) ; $Q_3 = \left[3\left(\frac{n+1}{4}\right)\right]^{th}$ term

For Continuous Series

For $Q_1$

→ Locate $\frac{N}{4}$ in cf
→ Select '$Q_1$' class
→ $Q_1 = l + \left[\frac{\frac{N}{4} - cf}{f}\right] \times h$

For $Q_3$

→ locate $\frac{3N}{4}$ in cf
→ Select $Q_3$ class
→ $Q_3 = l + \left[\frac{\frac{3N}{4} - cf}{f}\right] \times h$

# Deciles ($D_1, D_2, D_3, \ldots, D_9$) (Divides entire in 10 parts)

$\boxed{D_5 = \text{Median (5}^{th}\text{ Decile)}}$

For $D_1$

→ Locate $\frac{N}{10}$ in cf
→ Select $D_1$ class
→ $D_1 = l + \left[\frac{\frac{N}{10} - cf}{f}\right] \times h$

For $D_3$

→ Locate $\frac{3N}{10}$ in cf
→ Select $D_3$ class
→ $D_3 = l + \left[\frac{\frac{3N}{10} - cf}{f}\right] \times h$

# Percentiles ($P_1, P_2, \ldots, P_{99}$) [Divide entire into 100 parts]

$P_{50}$ = Median

<u>For Individual & Descrete Series</u>

$P_1 = \left(\frac{n+1}{100}\right)^{th}$ term ; $P_k = \left[k\left(\frac{n+1}{100}\right)\right]^{th}$ term

<u>For Continuous Series</u>

For $P_1$
→ Locate $N/100$ in cf
→ Select $P_1$ class
→ $P_1 = l + \left[\frac{\frac{N}{100} - cf}{f}\right] \times h$

For $P_7$
→ Locate $\frac{7N}{100}$ in cf
→ Select $P_7$ class
→ $P_7 = l + \left[\frac{\frac{7N}{100} - cf}{f}\right] \times h$

# MODE (An observation with highest frequency)

<u>Individual series</u> :- Eg:- 2, 1, 3, 1, 4, 3, 2, 3, 5, 2, 3, 1, 3, 1, 3, 5, 3, 4, 3
Mode = 3.

<u>Descrete series</u> :-

| xi | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| fi | 6 | 12 | 3 | 5 |

— Mode = 3

<u>For Continuous Series</u> → Check the modal class i.e class having highest frequency.

$$\text{Mode} = l + \left[\frac{f_1 - f_0}{2f_1 - f_2 - f_0}\right] \times h$$

[Note: Mode also changes with change of origin or change in scale i.e $y = a + bx$
Then mode of $y = a + b$(mode of $x$)]

# $\boxed{\text{Geometric Mean}}$

$N^{th}$ root of the product of n observations.

<u>For individual series</u>

$GM = (x_1 \times x_2 \times \ldots x_3)^{1/n}$

[Used for :-
→ Average Dep → Average of %]

<u>For descrete & continuous series</u>

$$GM = \left[(x_1)^{f_1} \times (x_2)^{f_2} \times \ldots \times x_n^{f_n}\right]^{1/N}$$

## $\boxed{\text{Properties}}$

1) If all items are same (let K) then $GM = K$
2) $\log(GM) = \Sigma \log x_i / N$

$$GM = AL\left[\frac{\Sigma \log x_i}{N}\right]$$

3) $GM(xy) = GM(x) \times GM(y)$
4) $GM\left(\frac{x}{y}\right) = \frac{GM(x)}{GM(y)}$
5) Combined GM

$$G = \left[(G_1)^{N_1} \times (G_2)^{N_2}\right]^{1/N_1+N_2}$$

# Harmonic Mean (Used for Avg speed, Avg of rates)

Reciprocal of average of reciprocall of 'N' observation

→ find reciprocal of all itmes → find avg of these rec → find reciprocal of average

For Individual Series

$$HM = \frac{N}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{N}{\Sigma\left(\frac{1}{x_i}\right)}$$

For discrete & Continuous series

$$HM = \frac{N}{\Sigma\left(\frac{f_i}{x_i}\right)}$$

Properties : 1) If all items are same (K) then HM = K

(2) Combined HM $= N_1 + N_2 / \frac{N_1}{H_1} + \frac{N_2}{H_2}$

$$\boxed{AM \geq GM \geq HM}$$

If all items are different then AM > GM > AM

For any two items a & b

$AM \times HM = GM^2$

- $GM = \sqrt{AM \times HM}$

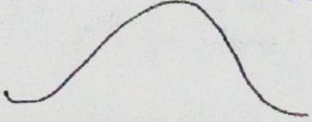
Weighted AM = $\Sigma W_i x_i / \Sigma W_i$

Weighted HM = $\Sigma W_i / \Sigma (W_i / x_i)$

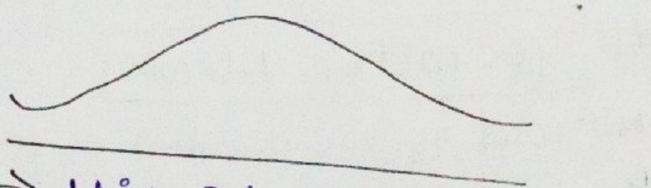
Weighted GM = $AL\left[\frac{\Sigma W_i \log x_i}{\Sigma W_i}\right]$

# Dispersion

Statistical Technique to find degree consistency (or variability) in all the Observation.



→ low dispersion
→ Concentrate Data
→ Less variability
→ More consistency



→ High Dispersion
→ Scattered data
→ More variability
→ Less consistency

Measures

Absolute

→ Range - Q.D - M.D - S.D

Relative

- Coffecient of range
- Coefficient of Q.D
- Coefficient of M.D
- Cofficient of variable

# Range

Difference b/w Largest & Smallest observation

$R_x = L - S$ ; Coefficient of Range $= \frac{L-S}{L+S} \times 100$

→ Range doesn't change with Origin

→ Range changes with scale

If $y = a + bx$ then $R_y = |b| * R_x$.

## Quartile Deviations

$\Rightarrow$ Interquartile Range $= Q_3 - Q_1$

$\Rightarrow$ Semi Quartile Range (Quartile Deviation) $= \frac{Q_3 - Q_1}{2}$

$\Rightarrow$ Coefficient of Q.D $= \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

Or

$\frac{QD}{median} \times 100$ (Only for symmetrical distribution)

→ QD doesn't change with anything

→ QD changes with scale

$y_i = a + bx \Rightarrow$ QD of $y = |b| \times$ QD of $x$.

## Mean Deviation (MD)

Average of Absolute deviations taken from mean, median or mode

About Mean (MD)

$$MD_{\bar{x}} = \frac{\sum f_i |x_i - \bar{x}|}{N}$$

About Median (MD)

$$MD_{M} = \frac{\sum f_i |x_i - m|}{N}$$

Coefficient of MD $= \frac{MD}{mean} \times 100 = \frac{MD}{Median} \times 100$

$\Rightarrow$ MD doesn't change with origin

$\Rightarrow$ MD changes with scale

[If $y = a + bx$ then $MD_y = |b| \times MD_x$

## # Standard Deviation & Variance

Variance $= \sigma^2$ ; S.D $= \sigma$ S.D $= \sqrt{\text{variance}}$

Variance $= (S.D)^2$

# S.D $(\sigma) = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{N}}$

# S.D $= \sqrt{\frac{\sum f_i x_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2}$

where $d_i = x_i - A$

# S.D $= \sqrt{\frac{f_i u_i^2}{N} - \left(\frac{\sum f_i u_i}{N}\right)^2} \times h$ where $\frac{x_i - A}{h} = u_i$

$\Rightarrow$ S.D doesn't change with origin

$\Rightarrow$ S.D change with change of scale

$$\boxed{\begin{array}{l}\text{If } y_i = a + bx \text{ ; S.D of } y_i = |b| \times \text{S.D of } x \\ \text{Variance of } y_i = b^2 \times \text{variance of } x_i\end{array}}$$

# Coefficient of Variance (C.V) $= \frac{\sigma}{\bar{x}} \times 100$

will be used for consistency & variable higher C.V = Higher variabi[lity]

Lesser CV = Higher consistency.

# S.D of first 'n' natural number = $\sqrt{\frac{n^2-1}{12}}$

# S.D of first 'n' even natural no. = $\sqrt{\frac{n^2-1}{3}}$

# S.D of first 'n' odd natural no. = $\sqrt{\frac{n^2-1}{3}}$

# S.D of two numbers a & b = $\frac{|a-b|}{2}$

# Q.D : MD : SD = 10 : 12 : 15

$\frac{QD}{MD} = \frac{10}{12}$ | $\frac{MD}{SD} = \frac{12}{15}$ | $\frac{QD}{SD} = \frac{10}{15}$

# Combined S.D

$$= \sqrt{\frac{N_1(\sigma_1^2 + d_1^2) + N_2(\sigma_2^2 + d_2^2)}{N_1 + N_2}}$$ where $d_1 = \overline{X}_{12} - \overline{X}_1$

$d_2 = \overline{X}_{12} - \overline{X}_2$

& $$\overline{X}_{12} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2}{N_1 + N_2}.$$