Measures of Central Tendency

Practically the condensation of data into a frequency distribution and the visual Presentation are not enough, when two or more different sets of data are to be compared.

⇒ A data set can be summerized in a single value at which the data have a tendency to concentrate. The tendency of the observations to cluster in the central part of the data is called central tendency, and the summary value as a measure of central tendency.

⇒ It indicates the location or general Position of the distribution.

⇒ Also known as Averages.

Types of Averages :-

⇒ Arithematic Mean

⇒ Geometric Mean

⇒ Harmonic Mean

⇒ Median

⇒ Mode

Arithmetic Mean:-

- ⇒ Most familiar Average.
- ⇒ It is defined as a value obtained by dividing the sum of all the observations by their number.

$$\text{Mean} = \frac{\text{Sum of all observations}}{\text{Number of observations}}$$

- ⇒ Population Mean is denoted by μ (Mu)
- ⇒ Sample Mean denoted by $\bar{x}$.

$$\bar{x} = \frac{\Sigma x_i}{n} \quad (i = 1, 2 \ldots n)$$

where $\bar{x}$ is mean and n is sample size.

ungroup data

$$\bar{x} = \frac{\Sigma x}{n}$$

Grouped data

$$\bar{x} = \frac{\Sigma f x}{\Sigma f}$$

Example 3-1 (pg 56)

The marks obtained by 9 students are given 45, 32, 37, 46, 39, 36, 41, 48, 36 calculate A.M.

$$\bar{x} = \frac{\Sigma x}{n} = \frac{45+32+37+46+39+36+41+48+36}{9}$$

$$= \frac{360}{9}$$

$$\boxed{\bar{x} = 40}$$

Question:- calculate the A.M from the following data-

weights	f	x	fx
65 - 84	9	74.5	670.5
85 - 104	10	94.5	94.5
105 - 124	17	114.5	1946.5
125 - 144	10	134.5	1345
145 - 164	5	154.5	772.5
165 - 184	4	174.5	698
185 - 204	5	194.5	972.5
	60		7350

$$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{7350}{60} = 122.5$$

Ans

Properties of Arithmatic Mean:-

1) Sum of deviations taken from their mean is zero. i.e $\Sigma(x - \bar{x}) = 0$

ii) Sum of squared deviations taken from their mean is least.

i.e $\Sigma (x - \bar{x})^2 \leq \Sigma (x - a)^2$

This property is known as minimal property of mean.

iii) If $Y = ax + b$, where a and b are any two numbers then $\bar{Y} = a\bar{x} + b$

(iv) It is independent of origin and scale.

Question:-

The sum of deviations of certain number of observations measured from 6 is 90 and the sum of deviations of observation measured from 9 is -9. Find the number of observations.

Sol:-

$$\Sigma(x-6) = 90. \quad \text{——} \quad (1)$$

$$\Sigma(x-9) = -9 \quad \text{——} \quad (2)$$

By Subtracting eq (2) from (1)

$$\cancel{\Sigma x} - 6n = 90$$

$$\cancel{\Sigma x} - 9n = -9$$

$$- \quad + \quad\quad +$$

$$3n = 99$$

$$n = \frac{99}{3}$$

$$\boxed{n = 33}$$

Ans

Geometric Mean:-

The Geometric Mean G, of a set of n positive values $x_1, x_2 ... x_n$ is defined as the positive n^{th} root of their product.

$$G = \sqrt[n]{x_1 \cdot x_2 \cdots x_n} \quad \text{where} \quad x > 0$$

⇒ G.M is appropriate to Average ratios.

⇒ G.M is used when data contains values with different units. e.g some measures are height, some are dollers and some are miles etc.

$$G = \text{antilog}\left[\frac{1}{n} \Sigma \log x_i\right]$$ (Example 3.7)

Group data

$$G = \text{antilog}\left[\frac{\Sigma f \log x_i}{\Sigma f}\right]$$ (Example 3.8)

Harmonic Mean:-

The harmonic mean H, of a set of n values $x_1, x_2 ... x_n$ is defined as the reciprocal of the A.M of the reciprocal of values.

$$H = \frac{n}{\Sigma(1/x_i)} \quad \text{where} \quad x \neq 0$$ (ungroup data)

$$H = \frac{\Sigma f}{\Sigma f(1/x_i)}$$ (Group data)

(Example 3.9)

Suppose a car is running at the rate of 15 km/hr during the first 30km; at 20 km/hr during the second 30km; and at 25 km/hr during the third 30km. The distance is constant but times are variable. So H.M is appropirate Average.

$$H = \text{Reciprocal of } \frac{\frac{1}{15}+\frac{1}{20}+\frac{1}{25}}{3}$$

$$= \frac{3}{0.0667+0.5000+0.04000}$$

$$= \frac{3}{0.15667}$$

$= 19.15$ km/hr approximately.

⇒ Harmonic mean is appropirate mean if the data is comprised of rates.

Median:-

Median is defined as a value which divides an arranged data set into two equal parts.

OR

Middle value of an arranged data set.

⇒ when n is odd

$\frac{1}{2}(n+1)^{th}$ item or value

⇒ when n is even

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

Grouped data

Median $= l + \frac{h}{f}\left(\frac{\Sigma f}{2} - c\right)$ (Example 3.13 Pg 71)

Example

Find Median from given data

1 5 3 9 7

1 3 5 7 9

Med $= \frac{1}{2}(5+1)^{th}$ value

$= 3^{rd}$ value

So

$\boxed{\text{Median} = 5}$

Quantiles:-

When number of observations is large, the principle according to which a distribution or data set is divided into two equal parts, may be extended to any number of divisions.

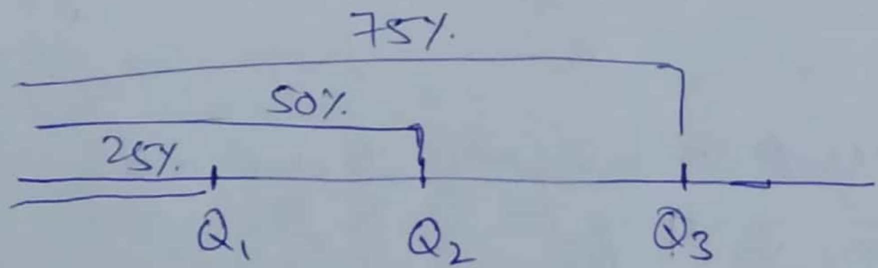
Quartiles:-

Quartiles are the values which divide an arranged data set into four equal parts.

Q_1 : first or Lower Quartile

Q_3 : third or upper Quartile

Q_2 : Second Quartile



ungroup data:-

odd

$$Q_1 = \left(\frac{n+1}{4}\right)^{Th} \text{ value}$$

$$Q_3 = \frac{3(n+1)}{4}^{Th} \text{ value}$$

Even

$$Q_1 = \frac{1}{2}\left[\frac{n}{4} + \left(\frac{n}{4}\right) + 1\right]^{th} \text{ value}$$

$$Q_3 = \frac{1}{2}\left[\frac{3n}{4} + \left(\frac{3n}{4}\right) + 1\right]^{th} \text{ value}$$

Grouped data:-

$$Q_1 = l + \frac{h}{f}\left(\frac{\Sigma f}{4} - c\right), \quad Q_3 = l + \frac{h}{f}\left(\frac{3\Sigma f}{4} - c\right)$$

OR

Finding Quartiles

First Arrange the data.

Even number of values

⇒ Split the data in upper half and lower half.

⇒ Median of upper half is Q_3 and median of lower half is Q_1.

e.g

7 8 9 11 12 13

Lower half upper half

7 8 9 11 12 13

$Q_1 = 8$ $Q_3 = 12$

Odd number of values :-

⇒ First Arrange the data.

⇒ Find Median and delete it.

⇒ Split remaining data into two parts.

⇒ Median of upper half is Q_3 and Median of lower half is Q_1.

e.g

7 8 11 13 $\boxed{15}$ 18 19 20 22

Median

~~upper~~ half upper half

7 8 11 13 18 19 20 22

$Q_1 = \frac{8+11}{2} = 9.5$

$Q_3 = \frac{19+22}{2}$

$Q_3 = 19.5$

Deciles :-

The nine values which divides an arranged data set into ten equal parts are called Deciles.

$D_1, D_2 \ldots D_9$.

e.g

$D_8 = \frac{8(n+1)}{10}^{th}$ value (odd)

$D_8 = \frac{1}{2}\left[\frac{8n}{10} + \left(\frac{8n}{10}\right)+1\right]^{th}$ value (even)

Grouped data:-

$D_8 = l + \frac{h}{f}\left(\frac{8n}{10} - c\right)$

Percentiles :-

The ninety nine values which divides an arranged data set into hundred equal parts are called percentiles.

$P_1, P_2 \ldots P_{99}$

e.g

$$P_{60} = \frac{60(n+1)}{100}^{th} \text{ value} \quad \text{(odd)}$$

$$P_{60} = \frac{1}{2}\left[\frac{60n}{100} + \left(\frac{60n}{100}\right)+1\right]^{th} \text{ value} \quad \text{(even)}$$

Grouped data

$$P_{60} = l + \frac{h}{f}\left(\frac{60\,\Sigma f}{100} - c\right)$$

Mode :-

Mode is "most frequent" value of a data set. A data set may have more than one mode or no mode at all when each observation occurs the same number of times.

$$Mo = l + \frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)} \times h$$

where l = lower class boundary of the modal class

f_m = frequency of modal class

f_1 = frequency of class preceding the modal class.

f_2 = frequency associated with the class following the modal class.

h = width of class interval.

⇒ A distribution having one mode is called a unimodal distribution.

⇒ A distribution having two modes called bimodal distribution.

⇒ A distribution having more than two modes is called multimodal.

Example

2, 3, 5, 7, 5, 8, 10, 5

Mode is 5. (Most frequent value)

Group data: (Example 3.14)
Pg 73