Measures of Location or measure of Central tendency

definition: (Central tendency)

The tendency of a set of quantative data is called central tendency.

A measure of central tendency (also referred to as measures of centre or central location) is a summary that attempts to describes a whole set of data with a single value that represents the middle or centre of its distribution.

For example, we often talk of average income, average weight, average age of employees etc. Thus an average is a single value which is considered as the most representative value ~~fore the~~ respective set of data.

There are three main measures of central tendency:

i) Mean
ii) Median
iii) Mode

Again there are three types of mean
iv) Arithmatic mean
v) Geometric mean
vi) Harmonic mean.

Arithmatic mean:

Arithmatic mean of a set of observation is the sum of all observations divided by the number of observations.

Arithmatic mean $\bar{x}$ of n ungrouped observations $x_1, x_2, \ldots, x_n$ is given by

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

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

Example: Find the arithmatic mean of 2, 5, 7, 9, 4 and 3

Solution:

Arithmatic mean, $\bar{x} = \dfrac{2+5+7+9+4+3}{6}$

$= \dfrac{30}{6} = 5$

This method is called direct method for finding arithmatic mean from the ungrouped data.

For frequency distribution (grouped data)

the arithmatic mean is $\bar{x} = \dfrac{f_1x_1+f_2x_2+\cdots+f_nx_n}{f_1+f_2+\cdots+f_n}$

$= \dfrac{\sum_{i=1}^{n} f_ix_i}{\sum_{i=1}^{n} f_i}$

here $\cancel{x = \sum_{i=1}^{n} f_i}$

f_i = frequency

x_i = mid-values of the interval.

Class interval	Mid point x_i	Frequency: f_i
5-10	7.5	5
10-15	12.5	7
15-20	17.5	11
20-25	22.5	8
25-30	27.5	4
30-35	32.5	2

Here $\sum f_i = 37$, $\sum fx = 672.50$

Arithmatic mean $\bar{x} = \dfrac{\sum f_ix_i}{\sum f_i}$

$= \dfrac{672.50}{37} = 18.18$

This method is called direct method for finding arithmatic mean

Shortcut method for ~~group~~ ungrouped data:

In this case, a new variable defined by

$$d_i = \frac{x_i - A}{c}$$

$$x_i - A = d_i c$$

$$x_i = A + d_i c$$

$$\bar{x} = A.$$

$$u_i = \frac{x_i - A}{h} \qquad \bar{u} = \frac{\frac{x_i - a}{h}}{n}$$

$$\Rightarrow x_i - A = h u_i$$

$$x_i = h u_i + A$$

$$\bar{x} = h \bar{u} + A$$

$$\therefore \cancel{\bar{x} = h \cdot \bar{x}}$$

Shortcut method for grouped data:

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

$$x_i = h u_i + A$$

$$\bar{X} = h \bar{u} + A$$

$$\bar{u} = \frac{\sum_{i=1}^{n} f_i u_i}{\sum_{i=1}^{n} f_i}$$

Example: Find the arithmatic mean by direct method as well as short method from the frequency distribution of ages with class interval of two years each from the following data of ages of 35 students in a certain locality.

Class interval of ages years	Number of students
11–13	3
13–15	4
15–17	5
17–19	10
19–21	6
21–23	4
23–25	3
total	35

Solution: calculation of arithmatic mean by both methods.

class interval of ages Years	Number of Students f_i	Mid-values of class interval x_i	$f_i x_i$	$u_i = \frac{x_i - 18}{2}$	$f_i u_i$
11-13	3	12	36	-3	-9
13-15	4	14	56	-2	-8
15-17	5	16	80	-1	-5
17-19	10	18	180	0	0
19-21	6	20	120	+1	6
21-23	4	22	88	2	8
23-25	3	24	72	3	9
total	35		632		

a) Direct method:

Arithmatic mean $\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{632}{35} = 18.06$ Years

b) Short cut method:

Aritmatic mean: $\bar{x} = c\bar{u} + A$

where $c = 2$, $A = 18$

$\bar{u} = \frac{\Sigma f_i u_i}{\Sigma f_i} = \frac{1}{35} = .03$

$\therefore \bar{x} = 2 \times .03 + 18 = 18.06$ years

Hence the mean of age students = 18.06 years is obtained from the both methods.

Arithmatic mean is dependent on origin and scale.

For ungrouped data:

Let $u_i = \frac{x_i - a}{c}$ [a = origin, c = scale]

$x_i - a = cu_i$

$\frac{x_i - a}{\cancel{n}} = \frac{cu_i}{\cancel{n}}$

$\Sigma(x_i - a) = c\Sigma u_i$

$\Sigma x_i - na = c\Sigma u_i$

$\frac{\Sigma x_i - na}{n} = c\frac{\Sigma u_i}{n}$ [divide by n]

$\Rightarrow \frac{\Sigma x_i}{n} - \frac{na}{n} = c\bar{u}$

$\Rightarrow \bar{x} - a = c\bar{u}$

$\bar{x} = c\bar{u} + a$

For grouped data:

Let new variable. $u_i = \frac{x_i - a}{c}$

$x_i - a = cu_i$

$x_i f_i - af_i = cu_i f_i$ [multiply by f_i]

$\Sigma x_i f_i - a\Sigma f_i = c\Sigma u_i f_i$

$\Rightarrow \cancel{\sum f_i}$

$$\frac{\sum x_i f_i - a\sum f_i}{N} = c\,\frac{\sum u_i f_i}{N} \qquad \left[N = \sum f_i\right]$$

$$\Rightarrow \frac{\sum x_i f_i}{N} - a\,\frac{\sum f_i}{N} = c\,\frac{\sum u_i f_i}{N}$$

$$\Rightarrow \bar{x} - a = c\bar{u}$$

$$\bar{x} = c\bar{u} + a$$

Theorem: The algebric sum of the deviations of the observations from their mean is zero.

Symbolically,

$$\sum (x_i - \bar{x}) = 0$$

Proof:

$$\sum (x_i - \bar{x}) = (x_1 - \bar{x}) + (x_2 - \bar{x}) + \cdots + (x_n - \bar{x})$$

$$= (x_1 + x_2 + x_3 + \cdots + x_n) - (\bar{x} + \bar{x} + \bar{x} + \cdots + \bar{x})$$

$$= \sum x_i - n\bar{x}$$

$$= n\bar{x} - n\bar{x}$$

$$= 0$$

$$\left[\bar{x} = \frac{\sum x_i}{n},\quad \sum x_i = n\bar{x}\right]$$

Example: Show that for the values 7, 9, 4, 8, 2, $\sum (x_i - \bar{x}) = 0$

Soln: Here $x_1 = 7,\ x_2 = 9,\ x_3 = 4,\ x_4 = 8,\ x_5 = 2$

$$\bar{x} = \frac{7+9+4+8+2}{5} = 6$$

Hence,

$$\sum (x_i - \bar{x}) = (x_1 - \bar{x}) + (x_2 - \bar{x}) + (x_3 - \bar{x}) + (x_4 - \bar{x}) + (x_5 - \bar{x})$$

$$= (7-6) + (9-6) + (4-6) + (8-6) + (2-6)$$

$$= 0$$