18 August 2023

Topic : Central Tendency

What is Population? ⇒ Performance ← marks of all students of class 11th in a school.

Total Students ⇒ 160 ⇒ population

20 20 20 20 = 80 ⇐ sample → representation of population.

sample ← population

sample is a subset of population.

Measures of Central Tendency

⇒ There are 3 types : ① Mean (or Average)

② Median

③ Mode.

also of 3 types

~~a)~~ Arithmetic mean (A·M)

~~b)~~ Geometric mean (G·M)

c) Harmonic mean (HM).

Formulas:

Data:

observation	x_1	x_2	x_3	$\cdots$	x_n
frequency f_i :	f_1	f_2	f_3	$\cdots$	f_n

① ARITHMETIC MEAN (A.M.)

a) Simple AM

Let $x_1, x_2, \dots, x_n$ be n observations repeated once.

(that is a data without frequency 'f')

Then, $$\boxed{\text{Simple AM } (\bar{x}) = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

b) weighted AM

b) Weighted AM

Let $x_1, x_2, \dots, x_n$ be 'n' observations with corresponding frequencies $f_1, f_2, \dots, f_n$.

Then, total frequency, $N = (f_1 + f_2 + \dots + f_n)$

or, $N = \sum_{i=1}^{n} f_i$

$$\therefore \text{Weighted AM}, \bar{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_n f_n}{f_1 + f_2 + f_3 + \dots + f_4}$$

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

or, $\boxed{\text{weighted mean}, \bar{x} = \frac{1}{N} \sum_{i=1}^{n} x_i f_i}$

Properties of Arithmetic Mean

① If all observations of a set of data are 'same' or 'equal' then arithmetic mean is the same value.

Proof

Let the 'n' observations be $x_1, x_2 \dots x_n$.

Since the observations have same value

then $x_i = c$ for all values of $i = 1(1)n$.

ie, $c, c, c, \dots, c$ (n times)

$$\therefore AM, \bar{x} = \frac{1}{n}(c + c + \dots + c) = \frac{1}{n}\sum_{i=1}^{n} c = \frac{n \cdot c}{n} = c$$

ie, $\boxed{\bar{x} = c}$

(2) If all the observations are repeated same number of times (in other words, frequencies are equal for all 'n' observations), then the $\boxed{\text{weighted AM} = \text{simple AM}}$.

Proof: Let the set of 'n' observations be $x_1, x_2, \ldots, x_n$

if observations are repeated same number of times then,

$f_i = k$ for all $i = 1(1)n$

$\therefore$ $\boxed{\text{total frequency}, N = \sum_{i=1}^{n} f_i = \sum_{i=1}^{n} k = nk}$

$$\therefore \boxed{\text{weighted AM}} = \frac{1}{\sum_{i=1}^{n} f_i} \sum_{i=1}^{n} x_i f_i$$

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

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

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

$$\boxed{= \text{simple AM}}$$

(3) The sum of the deviations of observations from its mean is zero.

ie $\boxed{\sum_{i=1}^{n} (x_i - \bar{x}) = 0}$.

Proof: Case I: without frequency.

Let the 'n' number of observations be $x_1, x_2, x_3, \ldots, x_n$

AM, $\boxed{\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i} \longrightarrow \boxed{\sum_{i=1}^{n} x_i = n\bar{x}}$

AM, $\boxed{\bar{x} = \frac{1}{n}\sum_{i=1} x_i} \longrightarrow \boxed{\sum_{i=1} x_i}$

Now. $\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x}$

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

$\Rightarrow \sum_{i=1}^{n}(x_i - \bar{x}) = 0$

Case II : with frequency:

Observations $\Rightarrow x_1, x_2, x_3, \ldots, x_n$ with corresponding frequencies.

$f_1, f_2, f_3, \ldots, f_n$

$\therefore$ Total frequency, $N = \sum_{i=1}^{n} f_i$

weighted AM $\boxed{\bar{x} = \frac{1}{N}\sum_{i=1}^{n} x_i f_i} \rightarrow \therefore N\bar{x} = \sum_{i=1}^{n} x_i f_i$

Now, $\sum_{i=1}^{n}(x_i - \bar{x}) f_i = \sum_{i=1}^{n} x_i f_i - \sum_{i=1}^{n} \bar{x} f_i$

$= N\bar{x} - \bar{x}\left(\sum_{i=1}^{n} f_i\right)$

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

$= 0$ ✓

(4) Arithmetic mean $(\bar{x})$ is dependent on change in both 'origin' and 'scale'.

Let us change each observation of x_i with origin 'a' and scale 'b' such that $\boxed{y_i = a + bx_i}$, Let us change each obs... and scale 'b' such that $y_i = a+bx_i$, then we have show that $\bar{y} = a+b\bar{x}$

we have $y_i = a+bx_i$

let us take summation on both side, then,

$$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} (a+bx_i)$$

let us divide both sides by 'n' we get,

$$\frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n} (a+bx_i)$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} a + \frac{b}{n}\sum_{i=1}^{n} x_i$$

$$\bar{y} = \frac{1}{\cancel{n}} \times \cancel{n}a + b\bar{x}$$

$$\boxed{\bar{y} = a+b\bar{x}}$$ (Proved)

(*) In case of weighted AM: $y_i = a+bx_i$

$$\frac{1}{N}\sum_{i=1}^{n} y_i f_i = \frac{1}{N}\sum_{i=1}^{n} (a+bx_i) f_i$$

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

then, $\bar{y} = \frac{1}{N}\sum_{i=1}^{n} a f_i + \frac{b}{N}\sum x_i f_i$

$= a + b\bar{x}$

$$\bar{y} = \frac{1}{\cancel{N}} a\cancel{N} + b\bar{x}$$

$$\boxed{\bar{y} = a + b\bar{x}}$$

(5) Sum of the square of the deviation is least around mean.

i.e, $\Sigma(x_i - A)^2$ is minimum at $A = \bar{x}$.

Proof:

Let $(x_i - A) = (x_i - \bar{x} + \bar{x} - A)$

Squaring both sides and taking summation we get,

$$\sum_{i=1}^{n}(x_i - A)^2 = \sum_{i=1}^{n}\left\{\underbrace{(x_i - \bar{x})}_{a} + \underbrace{(\bar{x} - A)}_{b}\right\}^2$$

$$\sum_{i=1}^{n} x_i - A = \sum_{i=1}^{n}\left\{(x_i - \bar{x})^2 + 2(x_i - \bar{x})(\bar{x} - A) + (\bar{x} - A)^2\right\}$$

$$= \sum_{i=1}^{n}(x_i - \bar{x})^2 + \underbrace{2(\bar{x} - A)\sum_{i=1}^{n}(x_i - \bar{x})}_{=0} + \sum_{i=1}^{n}(\bar{x} - A)^2$$

Since sum of deviation of obs from mean is 0

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

$$\therefore \sum_{i=1}^{n}(x_i - A)^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - A)^2$$

$\therefore \sum_{i=1} (x_i - ...$

comparing both sides
$$\sum_{i=1}^{n} (x_i - A)^2 \geq \sum_{i=1}^{n} (x_i - \bar{x})^2$$

from ① if $n(\bar{x} - A)^2 = 0$, then eq ① will be minimum

or, $n(\bar{x} - A) = 0$

or, $n\bar{x} - nA$

or, $n\bar{x} = nA$

or, $\boxed{\bar{x} = A}$

proved

What is combined mean of Grand mean $(\bar{X})$?

Let there be two sets of observations

set 1 : $x_{11}, x_{21}, x_{31} \cdots x_{n_1}$ (n_1 no. of observation)

mean of set 1 is $\bar{x}_1$

set 2 : $x_{12}, x_{22}, x_{32} \cdots x_{n_2}$ (n_2 no of obs)

mean of set 2 is $\bar{x}_2$.

$\therefore$ Grand mean, $$\boxed{\bar{X} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{(n_1 + n_2)}}$$

or, $\bar{X} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{N}$ where $N = (n_1 + n_2)$

m_1 $\bar{X} = \dfrac{n_1 x_1 + n_2 x_2}{N}$ where $N = (n_1 + n_2)$

m_1 $\boxed{N\bar{X} = n_1\bar{x}_1 + n_2\bar{x}_2}$