Statistics

By
Ankush Garg(B. Tech, IIT Jodhpur)

MEASURES OF CENTRAL TENDENCY

An average value or a central value of a distribution is the value of variable which is representative of the entire distribution, this representative value are called the measures of central tendency.

1. Mean
2. Median
3. Mode

1 **

ARITHMETIC MEAN :

n observation

(i) For ungrouped dist. : If $x_1, x_2, \ldots\ldots x_n$ are n values of variate x_i then their A.M. $\bar{x}$ is defined as

$$\left[\bar{x} = \frac{x_1 + x_2 + \ldots\ldots + x_n}{n}\right] = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\Rightarrow \quad \Sigma x_i = n\bar{x}$$

$$\boxed{\text{Mean} = \frac{\Sigma x_i}{n}}$$

(ii) For ungrouped and grouped freq. dist. : If $x_1, x_2, \ldots\ldots x_n$ are values of variate with corresponding frequencies $f_1, f_2, \ldots f_n$ then their A.M. is given by

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \ldots + f_nx_n}{f_1 + f_2 + \ldots + f_n} = \boxed{\frac{\sum_{i=1}^{n} f_i x_i}{N}}, \quad \text{where } \boxed{N = \sum_{i=1}^{n} f_i}$$

year	No. g stu (frequency)
16	5
17	10
15	5

$$\frac{16 + 16 + 16 + 16 + 16 \;\_\_\_\_}{20}$$

$$(5 \times 16) + (17 \times 10)$$

(iii) By short method : If the value of x_i are large, then the calculation of A.M. by using previous formula is quite tedious and time consuming. In such case we take deviation of variate from an arbitrary point a

Let $\boxed{d_i = x_i - a}$

$$\therefore \quad \boxed{\bar{x} = a + \frac{\sum f_i d_i}{N}},$$ where a is assumed mean

d_i	f_i
23	5
146	6
232	4

x_i	f_i
1723	5
1846	6
1932	4

(iv) By step deviation method : Sometime during the application of short method of finding the A.M. If each deviation d_i are divisible by a common number h(let)

Let $\boxed{u_i = \frac{d_i}{h}} = \frac{x_i - a}{h}$

$$\therefore \quad \boxed{\bar{x} = a + \left(\frac{\sum f_i u_i}{N}\right) h}$$

u_i	d_i	x_i	f_i
-1	-100	1400	4
0	0	1500	5
1	+100	1600	6
			15

Q)

If for some $x \in R$, the frequency distribution of the marks obtained by 20 students in a test is

Marks	2	3	5	7
Frequency	$(x+1)^2$	$2x-5$	x^2-3x	x

Then, the mean of the marks is **(2019 Main, 10 April I)**

(a) 3.0 (b) 2.8 (c) 2.5 (d) 3.2

$$\frac{16\times 2+1\times 3+7\times 3}{20}$$

$$\frac{56}{20} = 2.8$$

x_i	f_i	
2	$(x+1)^2$	16
3	$2x-5$	1
5	x^2-3x	0
7	x	3

$\boxed{\Sigma f_i = 20}$

$(x^2+2x+1)+(2x-5)+(x^2-3x)+x=20$

$2x^2+2x-4=20$

$x^2+x-2=10$

$x^2+x-12=0$

$\boxed{x=3}$, $\boxed{x=-4}$

Properties

If $\bar{x}$ is the mean of variate x_i then

(1) A.M. of $(x_i + \lambda) = \bar{x} + \lambda$

(2) A.M. of $(\lambda x_i) = \lambda \bar{x}$

(3) A.M. of $(ax_i + b) = a\bar{x} + b$ (where λ, a, b are constant)

$x_1, x_2, \text{———} - x_n$

$\rightarrow x_1 + d, x_2 + d, - - - x_n + d$

$$\boxed{\bar{X} = \frac{x_1 + x_2 + x_3 - x_n}{n}}$$

$$\bar{X} = \frac{x_1 + x_2 + - - x_n + nd}{n}$$

$$= \left(\frac{x_1 + x_2 + x - - x_n}{n}\right) + d$$

$2x_1 + 3, \quad 2x_2 + 3,$

MEDIAN :

The median of a series is the value of middle term of the series when the values are written in ascending order. Therefore median, divided an arranged series into two equal parts.

Formulae of median :

(i) For ungrouped distribution : Let n be the number of variate in a series then

$$\text{Median} = \begin{cases} \left(\dfrac{n+1}{2}\right)^{th} \text{ term, (when n is odd)} \\ \text{Mean of } \left(\dfrac{n}{2}\right)^{th} \text{ and } \left(\dfrac{n}{2}+1\right)^{th} \text{ terms, (when n is even)} \end{cases}$$

1, 2, 3, 4, 5

3, 5, 7, 11, 2, 13

7, 2, 3, 4, 6, 10, 1

2, 3, 5, 7, 11, 13

~~1, 2, 3~~, 4, 6, 7, 10

n=6

n=7

$$\frac{3^{rd} + 4^{th}}{2}$$

$$\frac{5+7}{2}, 6$$

(ii) **For ungrouped freq. dist.** : First we prepare the cumulative frequency (c.f.) column and Find value of N then

$$\text{Median} = \begin{cases} \left(\dfrac{N+1}{2}\right)^{th} \text{ term, (when N is odd)} \\ \text{Mean of } \left(\dfrac{N}{2}\right)^{th} \text{ and } \left(\dfrac{N}{2}+1\right)^{th} \text{ terms, (when N is even)} \end{cases}$$

x_i	f_i	c.f
2	4	4
3	5	9
4	6	15
	$\Sigma f_i = 15$	

Median class

$\boxed{\text{Median} = 3}$

$\frac{15}{2} = 7.5$

(iii) For grouped freq. dist : Prepare c.f. column and find value of $\frac{N}{2}$ then find the class which contain value of c.f. is equal or just greater to N/2, this is median class

$$\therefore \boxed{\text{Median} = \ell + \frac{\left(\frac{N}{2} - F\right)}{f} h}$$

$$\boxed{l + \frac{\left(\frac{N}{2} - F\right)}{f} \times h}$$

where ℓ — lower limit of median class (20)

f — freq. of median class

F — c.f. of the class preceeding median class

h — Class interval of median class (10)

$$20 + \left(\frac{10 - 2}{8}\right) \times 10$$

$$20 + 10 = 30$$

x_i	f_i	cf
10 – 20	2	2
20 – 30	8	10
30 – 40	6	16
40 – 50	4	20
	$\Sigma f_i = 20$	

F

Median class

9) The mean and the median of the following ten numbers in increasing order 10, 22, 26, 29, 34, x, 42, 67, 70, y are 42 and 35 respectively, then $\frac{y}{x}$ is equal to

(2019 Main, 9 April II)

$\frac{126}{36}$ $\frac{42}{12}$ $\frac{7}{2}$

(a) $\frac{7}{3}$ (b) $\frac{7}{2}$

(c) $\frac{8}{3}$ (d) $\frac{9}{4}$

$\frac{y}{x}$

36

10, 22, 26, 29, 34, x, 42, 67, 70, y → 126

$\boxed{\bar{x} = 42}$

$\boxed{\text{Median} = 35}$

207 + 87

$35 = \frac{(5^{th}) + (6^{th})}{2}$

$35 = \frac{34 + x}{2}$

$\boxed{x = 36}$

$42 = \frac{294 + y}{10}$

$\boxed{y = 126}$