

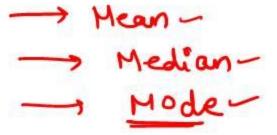
Statistics

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.



Averag

ARITHMETIC MEAN :

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

$$\sum_{x_i=n}^n \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

$$\sum_{x_i=n}^n x_i$$

$$\sum_{x_i=n}^n x_i$$

$$\sum_{x_i=n}^n x_i$$

$$\sum_{x_i=n}^n x_i$$

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

$$\overline{x} = \frac{f_1 x_1 + f_2 x_2 + \ldots + f_n x_n}{f_1 + f_2 + \ldots + f_n} = \frac{\sum_{i=1}^n f_i x_i}{N}, \quad \text{where } N = \sum_{i=1}^n f_i$$

* 3). Find Hean of the distribution

$$\overline{X} = (48 \times 2) + (28 \times 7) + (30 \times 8) + (40 \times 3)$$



(iii) By short method: If the value of x 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.

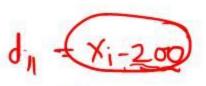
Let

$$d_i = x_i - a_i$$

.

$$\bar{x} = a + \left(\frac{\sum f_i d_i}{N}\right)$$

where a is assumed mean



>d1 30 20

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

Let

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

:.

$$= a + \left(\frac{\sum f_i u_i}{N}\right)$$

11, - 3 2 5



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

Marks **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

$$\bar{x} = \frac{32 + 3 + 2}{20}$$

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

$$x^{2} + 2x + 1 + 2x - 5 + x^{2} - 3x + x = 20$$

$$2x^{2} + 2x - 24 = 0$$

$$x^{2} + x - 12 = 0$$

$$(x+4)(x-3) = 0$$

$$x = 3$$



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

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

A.M. of
$$(\lambda x) = \lambda \overline{x}$$

A.M. of
$$(ax_1 + b) = a\overline{x} + b$$

(where λ , a, b are constant)

$$x_1, x_2, --x_n \longrightarrow X$$

 $x_1, x_2, --x_n \longrightarrow X$

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

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

$$\chi_1, \chi_2 - - - \chi_{10}$$

$$\chi_1, \chi_2 - - - \chi_{10}$$

$$\chi_1, \chi_2 - - - \chi_{10}$$

$$\chi_2 + \chi_2 - - - \chi_{10}$$

$$\chi_2 + \chi_2 - - - \chi_{10}$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_2 + \chi_4 + \chi_5$$

$$\chi_2 + \chi_4 + \chi_5$$

$$\chi_1 + \chi_4 + \chi_5$$

$$\chi_2 + \chi_5 + \chi_6$$

$$\chi_1 + \chi_5 + \chi_6$$

$$\chi_1 + \chi_5 + \chi_6$$

$$\chi_1 + \chi_5 + \chi_6$$

$$\chi_2 + \chi_5 + \chi_6$$

$$\chi_1 + \chi_2 + \chi_6$$

$$\chi_1 + \chi_6$$

$$\chi_2 + \chi_6$$

$$\chi_1 + \chi_6$$

$$\chi_1 + \chi_6$$

$$\chi_1 + \chi_6$$

$$\chi_2 + \chi_6$$

$$\chi_1 + \chi_6$$

$$\chi_2 + \chi_6$$

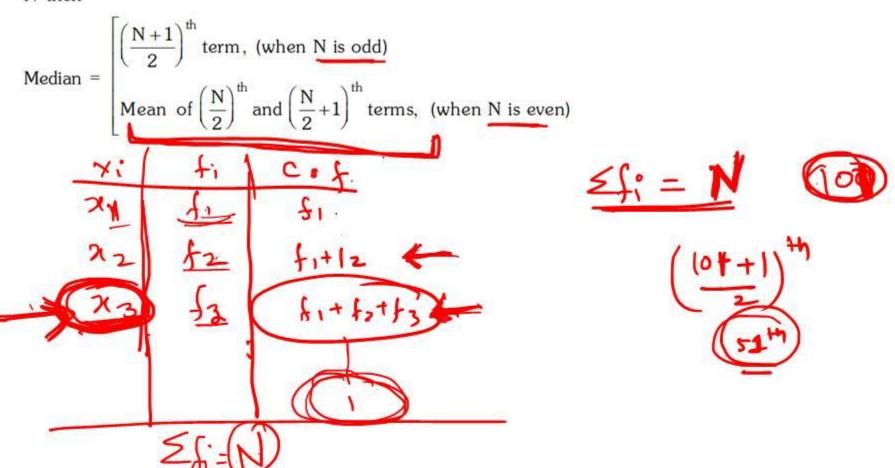
$$\chi_1 + \chi_6$$

$$\chi_1 + \chi_6$$

$$\chi_2 + \chi_$$



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



For grouped freq. dist: Prepare c.f. column and find value of the find the class which contain

value of c.f. is equal or just greater to N/2, this is median class

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

where

ℓ — lower limit of median class ∠—

f - freq. of median class

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

h - Class interval of median class

				East 100 miles	
	ni	f:	C.t		
(D)	10-20	10			
	20-30	29		Median	class
4	9 40		-		
			_=		
6		Eti= N			





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)

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

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

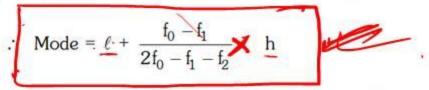
(c) $\frac{8}{3}$

(d) $\frac{9}{4}$

In a frequency distribution the mode is the value of that variate which have the maximum frequency Method for determining mode:

For ungrouped dist.: The value of that variate which is repeated maximum number of times for ungrouped freq. dist.: The value of that variate which have maximum frequency.

(iii) For grouped freq. dist. : First we find the class which have maximum frequency, this is model calss



where

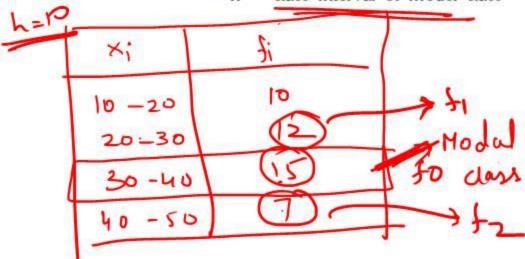
- lower limit of model class

 f_0 - freq. of the model class

 f_1 - freq. of the class preceeding model class .

f₂ - freq. of the class succeeding model class

h - class interval of model class





RELATION BETWEEN MEAN, MEDIAN AND MODE



MEASURES OF DISPERSION :



OM

The dispersion of a statistical distribution is the measure of deviation of its values about the their average (central) value.

Mean deviation (M.D.): The mean deviation of a distribution is, the mean of absolute value of deviations of variate from their statistical average (Mean, Median, Mode).

If A is any statistical average of a distribution then mean deviation about A is defined as

Mean deviation =
$$\frac{\sum_{i=1}^{n} |x_i - A|}{n}$$
 (for ungrouped dist.)

Mean deviation =
$$\frac{\sum_{i=1}^{n} f_i |x_i - A|}{N}$$
 (for freq. dist.)

$$\begin{cases} 1 & \text{find } A \\ \text{find } A \\$$



If the sum of the deviations of 50 observations from

30 is 50, then the mean of these observations is

(2019 Main, 12 Jan I)

(a) 50

(b) 30

(c) 51

(d) 31

$$(x_{1}-30) + (x_{2}-30) - + (x_{50}-30) = 50$$

$$x_{1}+x_{2}-x_{50}-(30x_{50}) = 50$$

$$x_{1}+x_{2}-x_{50} = 50+30x_{50}$$

$$= 31x_{50}$$

$$x_{1}+x_{2}-x_{50}$$

$$= 31x_{50}$$

$$= 31$$



(iii) Variance and standard deviation: The variance of a distribution is, the mean of squares of deviation of variate from their mean. It is denoted by σ^2 or var(x).

The positive square root of the variance are called the standard deviation. It is denoted by σ or S.D.

Formulae for variance :

(i) for ungrouped dist. :

$$\sigma_{\rm d}^2 = 0$$

$$\leq \left[\frac{x^2+(\overline{x})^2-2x_i\overline{x}}{x}\right]$$

$$=\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2$$

where
$$d_i = x_i - a$$

Find the variance of first n natural numbers



$$\frac{1}{x^{2}}, \frac{1}{x^{2}}, \frac{1}{x^{2}} = \frac{1}{x^{2}} \frac{1}{x^{2}} - \frac{1}{x^{2}} = \frac{1}{x^{2}} \frac{1}{x^{2}} - \frac{1}{x^{2}} = \frac{1}{x^{2}} \frac{1}{x^{2}} =$$

If the data x_1, x_2, \dots, x_{10} is such that the mean of first four of these is 11, the mean of the remaining six is 16 and the sum of squares of all of these is 2000, then the standard deviation of this data is (2019 Main, 12 April I)

x 1+ x2+ x3+ x4 = 11

X1+ x2+ -- >4 = 44 X5+ X6 - - X10 = 96

(a) $2\sqrt{2}$ (b) 2

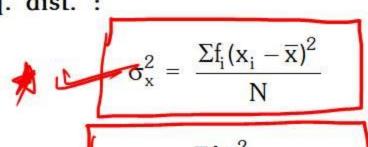
(c) 4

(d) $\sqrt{2}$

a = Travance \(\si^2\) - (\overline{\times})^* 200 - ()2 200 -196



(ii) For freq. dist. :



$$\sigma_{\rm x}^2 = \frac{\Sigma}{2}$$

$$\sigma_{x}^{2} = \frac{\Sigma f_{i} x_{i}^{2}}{N} - (\overline{x})^{2} = \frac{\Sigma f_{i} x_{i}^{2}}{N} - \left(\frac{\Sigma f_{i} x_{i}}{N}\right)^{2}$$

$$\sigma_{d}^{2} = \frac{\Sigma f_{i} d_{i}^{2}}{N} - \left(\frac{\Sigma f_{i} d_{i}}{N}\right)^{2}$$

$$\sigma_{u}^{2} = h^{2} \left[\frac{\Sigma f_{i} u_{i}^{2}}{N} - \left(\frac{\Sigma f_{i} u_{i}}{N} \right)^{2} \right] \qquad \text{where } u_{i}^{2} = \frac{d_{i}}{h}$$

where
$$u_i = \frac{d_i}{h}$$



(iii) Coefficient of S.D. =
$$\frac{\sigma}{\chi}$$

Coefficient of variation =
$$\frac{\sigma}{\overline{x}} \times 100$$

Note :-
$$\sigma^2 = \sigma_x^2 = \sigma_d^2 = h^2 \sigma_u^2$$

$$A^{3} = 0$$

$$A^{2} \neq 0$$

$$B - I = 0$$

$$A = B - BA$$

$$A = B \cdot A - BA^{2}$$

$$B - I = 0$$

$$B = I$$

$$A^{2} = A^{2} = A^{2}$$

$$B - A^{2} = A^{2} = A^{2}$$



The mean of five observations is 5 and their variance is 9.20. If three of the given five observations are 1, 3 and 8, then a ratio of other two observations is (2019 Main, 10 Jan I)

- 4:9
- (b) 6:7
- (c) 10:3
- (d) 5:8

$$\frac{1,3,8,7,3}{5} = 5$$

$$\frac{1+3+8+x+y}{5} = 5$$

$$\frac{3\cdot 2}{5} = \frac{5\times 1^2}{n} - (x)^2$$

$$\frac{3\cdot 2}{5} = \frac{1+9+6+x^2+5}{5} - 25$$

$$\frac{3+y=13}{5}$$

$$\frac{3+y=13}{5}$$

$$\frac{3+y=13}{5}$$

$$\frac{3+y=13}{5}$$

$$\frac{3+y=13}{5}$$

$$\frac{3+y=13}{5}$$

$$\frac{3+y=13}{5}$$

$$\frac{3+y=13}{5}$$

$$\frac{3+y=13}{5}$$

$$\frac{3+y=1}{5}$$

$$\frac{3+y=1}{5}$$