Measure of central Tendency

Arithmetic Mean

$x_1, x_2, \ldots, x_n \leftarrow$ set of data / obs.

and there are a total of N - obs.

Then

$$\bar{X} \text{ or } \bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{N}$$

Shortcut Method :

$$\bar{X} = A + \frac{\sum d}{N} \quad ;$$

$\bar{X} \leftarrow$ A.M.

$A \leftarrow$ assumed mean

$d \leftarrow$ deviation of items from assumed

$N \leftarrow$ total no. of obs.

Types of Data

- Grouped Data
- Ungrouped Data

Q. The following table contains the half-yearly bonus paid to 10 workers in a factory

S.No	1	2	3	4	5	6	7	8	9	10
Half Yearly Bonus	150	200	300	650	250	180	400	500	550	220

Find out the arithmetic mean.

$$\bar{X} = \frac{150+200+300+650+250+180+400+500+550+220}{10}$$

$$= 340$$

Shortcut :

Let, assumed mean be $A = 300$

S.No.	Half Yearly bonus (x)	$d = x - 300$
1	150	-150
2	200	-100
3	300	0
4	650	350
5	250	-50
6	180	-120
7	400	100
8	500	200
9	550	250
10	220	-80
		$\Sigma d = 400$

Now,

$$\bar{x} = A + \frac{\Sigma d}{N}$$

$$= 300 + \frac{400}{10}$$

$$= \frac{3400}{10}$$

$$= 340$$

Discrete Series

Direct :

$$\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \cdots + f_n x_n}{f_1 + f_2 + \cdots + f_n} = \frac{\sum_i f_i x_i}{\sum_i f_i} \quad \text{or} \quad \frac{\sum f x}{\sum f}$$

where f_i's are frequencies corresponding to the values x_i's.

Shortcut :

$$\bar{x} = A + \frac{\sum f d}{N} \quad ; \quad N = \sum f$$

step deviation method :

$$\bar{x} = A + \frac{\sum f d'}{N} \times h \quad ; \quad d' = \frac{x - A}{h}$$

h = step deviation

Q. Calculate the mean of the following frequency distribution of marks in a test in statistics

Marks :	10	20	30	40	50	60	70	80
No. of students :	3	6	10	12	9	6	2	2

Direct method

marks (x)	No. of students (f)	fx
10	3	30
20	6	120
30	10	300
40	12	480
50	9	450
60	6	360
70	2	140
80	2	160
	$N = \Sigma f = 50$	$\Sigma fx = 2040$

$$\bar{X} = \frac{\Sigma f x}{\Sigma f} = \frac{2040}{50} = 40.8$$

Shortcut method

let, assumed mean, $A = 40$

Marks (x)	No. of students (f)	$d = x - 40$	fd
10	3	-30	-90
20	6	-20	-120
30	10	-10	-100
40	12	0	0
50	9	10	90
60	6	20	120
70	2	30	60
80	2	40	80

$N = \Sigma f = 50$ $\quad\quad \Sigma fd = 40$

$$\therefore \bar{X} = A + \frac{\Sigma fd}{N} = 40 + \frac{40}{50} = 40.8$$

Step deviation method

$A = 40$, $h = 10$

Marks (x)	No. of students (f)	$d' = \frac{x-40}{10}$	fd'
10	3	-3	-9
20	6	-2	-12
30	10	-1	-10
40	12	0	0
50	9	1	9
60	6	2	12
70	2	3	6
80	2	4	8
	$N = \Sigma f = 50$		$\Sigma fd' = 4$

$$\bar{X} = A + \frac{\Sigma fd'}{N} \times h = 40 + \frac{4}{50} \times 10 = 40.8$$

A point is randomly selected with uniform probability in the xy plane with in the rectangle with corners at (0,0), (1,0), (1,2) and (0,2). If p is the length of the p.v. of the point then expected value of p^2 is

$$f(x) = \frac{1}{B-A} \quad , \quad A \le x \le B$$

Uniform probability | Rectangular prob.

$U(a, b)$

parameters

$a \leftarrow min^m \quad , \quad b \leftarrow max^m$

Continuous Series

Direct: $\bar{X} = \frac{\Sigma fm}{N}$; m = midpoint of various classes

Shortcut method: $\bar{x} = A + \frac{\Sigma fd}{N}$

Step deviation method: $\bar{X} = A + \left(\frac{\Sigma fd'}{N} \times h\right)$

Q. Find the arithmetic mean

Marks	No. of students
0-10	5
10-20	10
20-30	40
30-40	20
40-50	25

Direct method

Marks	Mid-values (m)	No. of students (f)	fm
0 - 10	5	5	25
10 - 20	15	10	150
20 - 30	25	40	1000
30 - 40	35	20	700
40 - 50	45	25	1125
		$N = \Sigma f = 100$	$\Sigma fm = 3000$

$$\bar{X} = \frac{\Sigma fm}{N} = \frac{3000}{100} = 30$$

Shortcut method

Marks	Mid-point (x)	No. of students (f)	Deviation from assumed mean ($d = x - 25$)	fd
0-10	5	5	-20	-100
10-20	15	10	-10	-100
20-30	25	40	0	0
30-40	35	20	10	200
40-50	45	25	20	500
		$N = \Sigma f = 100$		$\Sigma fd = 500$

Here, assumed mean, $A = 25$

$$\therefore \bar{x} = A + \frac{\Sigma fd}{N} = 25 + \frac{500}{100} = 25 + 5 = 30$$

Step deviation method

Marks	Mid-point (x)	No. of students (f)	$d' = \frac{x-25}{10}$	fd'
0-10	5	5	-2	-10
10-20	15	10	-1	-10
20-30	25	40	0	0
30-40	35	20	1	20
40-50	45	25	2	50
		$N = \Sigma f = 100$		$\Sigma fd = 50$

Here, assumed mean, $A = 25$

step deviation, $h = 10$

$$\therefore \bar{X} = A + \frac{\Sigma fd'}{N} \times h = 25 + \frac{50}{100} \times 10 = 30$$

Properties

① Total of the deviations of items from the mean is equal to zero

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

② If a series of observation consists of two or more component series, the mean of the whole series is given by

$$\bar{X}_{1,2} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2}$$

where

$\bar{X}_{1,2}$ = combined mean

$\bar{X}_1$ = mean of 1st series

$\bar{X}_2$ = mean of 2nd series

N_1 = no. of obs in 1st group | N_2 = no. of obs in 2nd group

$$\bar{X}_{1,2,\ldots,n} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2 + \ldots + N_n\bar{X}_n}{N_1 + N_2 + \ldots + N_2}$$

Q. A cooperative bank has two branches employing 50 and 70 workers respectively. The average salaries paid by two respective branches are ₹ 360 and ₹ 390 per month. Calculate the mean of the salaries of all the employees

Given, $N_1 = 50$, $N_2 = 70$

$\bar{X}_1 = 360$, $\bar{X}_2 = 390$

$$\therefore \bar{X}_{1,2} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2} = \frac{(50 \times 360) + (70 \times 390)}{50 + 70} = 377.5$$

9. The arithmetic mean of a series of 40 items was calculated by a student as ₹ 265, while calculating it an item ₹ 115 was misread as ₹ 150. Find the correct arithmetic mean.

we know that,

$$\bar{X} = \frac{\Sigma X}{N}$$

$$\Rightarrow \Sigma X = \bar{X} N \quad \text{——} \; (1)$$

Given, $\bar{X} = 265$

$N = 40$

Wrong calculation : $\Sigma X(w) = 265 \times 40 = 10600$

Correct value, $C = 115$

Wrong value, $W = 150$

$$\text{Correct } \bar{X} = \frac{\Sigma X(w) + C - W}{N}$$

$$= \frac{10600 + 115 - 150}{40} = 264.12$$

Median

Let, $x_1, x_2, \ldots, x_N$ be the N values of a variable written in ascending (or descending) order of magnitude. The median denoted M or M_d is given by

Median = value of the middle item

① when N is odd, then median = $\left(\frac{N+1}{2}\right)^{th}$ value

② when N is even, then median = $\dfrac{\left(\frac{N}{2}\right)^{th} + \left(\frac{N}{2}+1\right)^{th}}{2}$ value

Discrete Series

$$\text{Median} = \left(\frac{N+1}{2}\right)^{th} \text{value} \quad ; \quad N = \Sigma f$$

* Values must be arranged in ascending or descending order and construct a table consisting freq. & c.f.

cumulative frequency

Continuous Series

Steps

① compute cumulative frequency (c.f)

② Median item is located by finding out of $\left(\frac{N}{2}\right)^{th}$ item

③ Locate the median group in c.f. column where the size of $\left(\frac{N}{2}\right)^{th}$ item falls

④ Apply the formula

$$\boxed{\text{Median} = L + \frac{\frac{N}{2} - cf}{f} \times h}$$ where

L = lower limit of median class

N = sum of frequencies

f = frequency of median class

cf = cumulative frequency of the class preceeding the median class

h = width of class interval

Q. The consumption of printing paper reams (in units) for the first 11 months of a computer operator is given by

20, 25, 30, 15, 17, 35, 26, 18, 40, 45, 50

Soln: Arranging the given data in ascending order,

15, 17, 18, 20, 25, 26, 30, 35, 40, 45, 50

No. of terms = 11 i.e. odd

$$\text{Median} = \left(\frac{N+1}{2}\right)^{th} \text{ obs. value}$$

$$= \left(\frac{11+1}{2}\right)^{th} \text{ obs. value.}$$

$$= 6^{th} \text{ obs. value}$$

$$= 26$$

Q. Calculate the median of the following data that relates to monthly salaries of employees (in thousand rupees) : 110, 115, 108, 112, 120

116, 140, 135, 128, 132.

Ans : Median salary : ₹ 118,000

Q. Obtain the median size of shoes sold from the following data.

Size	5	5½	6	6½	7	7½	8	8½	9	9½	10	10½	11	11½
No. of pairs	30	40	50	150	300	600	950	820	750	440	250	150	40	39

Soln :

Size (x)	No. of pairs (f)	cumulative frequency (c.f)
5	30	30
$5\frac{1}{2}$	40	70
6	50	120
$6\frac{1}{2}$	150	270
7	300	570
$7\frac{1}{2}$	600	1170
8	950	2120
$8\frac{1}{2}$	820	2940
9	750	3690
$9\frac{1}{2}$	440	4130
10	250	4380
$10\frac{1}{2}$	150	4530
11	40	4570
$11\frac{1}{2}$	39	4609

$N = \Sigma f = 4609$

$$\text{Median} = \left(\frac{N+1}{2}\right)^{th} \text{value} = \left(\frac{4609+1}{2}\right)^{th} \text{value} = 2305^{th} \text{ value}$$

∴ Median corresponds to 2305th value in the series

⇒ Median shall be the value corresponding to 2940th cumulative frequency ecause it is the first value after 2305th value

⇒ Median = $8\frac{1}{2}$

Hence, median size of the shoe is $8\frac{1}{2}$

Q. An insurance company obtained the following data for accident claims from a particular region. Obtain the median from this data

Amount of claim (in 1000)	Frequency
1-3	6
3-5	53
5-7	85
7-9	56
9-11	21
11-13	16
13-15	4
15-17	4

Sol^n

Amount	Frequency (f)	c.f.
1-3	6	6
3-5	53	59
5-7	85	144
7-9	56	200
9-11	21	221
11-13	16	237
13-15	4	241
15-17	4	245
	$N = \Sigma f = 245$	

Here, $N = \Sigma f = 245$ i.e. odd

$$\frac{N}{2} = \frac{245}{2} = 122.5$$

Median class is 5-7

l = lower limit of median class

$= 5$

f = frequency of median class

$= 85$

cf = cumulatative freq. of the class proceeding median class

$= 59$

h = width of the class interval

$= 2$

$$\therefore \text{Median} = L + \frac{\frac{N}{2} - cf}{f} \times h$$

$$= 5 + \frac{122.5 - 59}{85} \times 2$$

$$= 6.5$$

Q. In the frequency of 100 families given below, the number of families corresponding to expenditure group 10-20 and 40-50 are missing from the table. However, the median is known to be 30. Find the missing frequencies.

expenditure	0-10	10-20	20-30	30-40	40-50	50-60
No. of families	10	?	25	30	?	10

Let, the missing freq of 10-20 be f_1

and 40-50 be f_2

Now,

$N = 100$ (given)

$\Rightarrow 10 + f_1 + 25 + 30 + f_2 + 10 = 100$

$\Rightarrow f_1 + f_2 = 25$ —— 1

exp.	No. of families (f)	c.f.
0-10	10	10
10-20	f_1	$10+f_1$
20-30	25	$35+f_1$
30-40	30	$65+f_1$
40-50	f_2	$65+f_1+f_2$
50-60	10	$75+f_1+f_2$

$$N = 100$$

$$\frac{N}{2} = 50$$

Median class is 30-40

Given, median = 30

Now,

$$\text{Median} = l + \frac{\frac{N}{2} - cf}{f} \times h$$

$$\Rightarrow 30 = 30 + \frac{50 - (35 + f_1)}{30} \times 10$$

$$\Rightarrow 0 = \frac{15 - f_1}{3}$$

$$\Rightarrow f_1 = 15$$

From ①,

$$15 + b_2 = 25$$

$$\Rightarrow b_2 = 10$$

$\therefore$ Missing frequencies are 15, 10

Discrete Series

Mode = data with the highest frequency i.e. data that repeats the most.

Continuous Series

$$\text{Mode } (M_0) = L + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

where L = Lower limit of the modal class

f_1 = frequency of the modal class

f_2 = frequency of the class succeeding the modal class

f_0 = frequency of the class preceeding the modal class

h = width of the modal class.

* Mode = 3 Median − 2 Mean

Q. Find the mode of the following marks obtained by 15 students given by

4, 6, 5, 7, 9, 8, 10, 4, 7, 6, 5, 8, 7, 7, 9

Soln:

Marks	Frequency
4	2
5	2
6	2
7	4
8	2
9	2
10	1

$\therefore$ Mode = 7

Q. Calculate the mode of the following series

Marks	No. of students
200-220	7
220-240	15
240-260	20
260-280	20
280-300	6
300-320	4
320-340	2

Here,

modal class = 240-260

$L = 240$

$f_1 = 20$

$f_0 = 15$

$f_2 = 20$

$h = 20$

we know that,

$$M_o = L + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

$$= 240 + \frac{20-15}{40-15-20} \times 20$$

$$= 240 + 20$$

$$= 260$$

Q. In an asymmetrical distribution mean is 16 and median is 20. Calculate mode.