Statistics & Probability:-

Quartile:-

Sometimes median not represent the data properly. so we used quartile

Quartile is same as median but here the data is divided into four equal parts.

As median divide whole data into two equal parts so have one median value. Similarly Quartile divided the data into four equal parts so we have three Quartile to divide the data.

Quartile.

- First Quartile, Q_1 or Lower Quartile
- Median
- Third Quartile, Q_3 or Upper Quartile

~~Rules~~ Decile:-

Decile divides the given data into 10 equal parts. So here we will need 9 divisor or decile to divide the data.

$D_1, D_2, D_3, D_4, D_5, \dots D_9$.

$\frac{n}{10}, \frac{2n}{10}, \frac{3n}{10}, \text{---} \frac{9n}{10}$.

3. Percentile:-

It divides the given data into 100 equal parts Usually used when we have very large data.

$$D_J = \frac{Jn}{100}$$

Note:-

10 percentile — 1 decile
25 percentile → 1st Quartile.

Rules for Quartile, Desile & Percentile:-

- For Raw Data:-

The rule for raw data of Quartile, decile and percentile are same but just formula is different i.e.

For Quartile divide the number of observation by 4 i.e $n/4$ for decile $n/10$ & for percentile it is $n/100$.

Now if $n/4, n/10, n/100$ is integer than 1st Quartile will be ~~$[n/4 + 1]$~~ average of $n/4$ & $(n/4 + 1)$.

Similarly 1st decile will be if $n/10$ is Integer is $\left(n/10, \frac{n}{10}+1\right) \rightarrow$ Take its average.

& same for percentile.

→ If the result is not integer.

than for 1st Quartile $[n/4]+1$ will be 1st Quartile.

→ But the data should arrange in accending order.

(i) For decile

$[n/10+1]$ will be first decile.

so on for percentile.

For Ungroup data:-

→ First find Commulative frequency.

Now if we get from $n/4$, $n/10$ or $n/100$ non integer so we used greatest Integer rule & add 1 with that.

For e.g. if $n=50$.

Non Integer

So for 1st Quartile $= \frac{50}{4} = 12.5$

$[12.5]+1 = 13 \rightarrow$

Now check 13 in commulative freq. Table if present than infront of this value that become our Quartile.

Scanned by CamScanner

But if we use $n/4$ & we get Integer for e.g.

$$\frac{52}{4} = 13 \rightarrow \text{Integer.}$$

Now we will check in c. frequency table if present than average of that value & preceeding will be our first Quartile.

(¥) This rule is used if we get Integer without using greatest value/integer so we will consider average of that value infront of c. freq & preceding value.

Case #2 :-

If from $n/4$, $n/10$ or $n/100$ we get Integer so we will check in c. frequency column if its not present we will take range & consider the preceeding one & take its value as 1st Quartile, decile or 1st percentile

→ Similarly if $n/4$, $n/10$, or $n/100$ is not integer & we will use greatest Integer + 1 & than check in c. freq column. If In range of two number so consider the preceding one

For Grouped freq. data.

First step is to introduce class boundry.

→ First introduce Com. frequency.

→ Then find $n/4$, $n/10$ or $n/100$ & locate in C. freq Column to find median class.

Now after finding median Class. Use this formula.

For first Quartile.

$$\boxed{Q_1 = l + \frac{h}{f}\left(\frac{n}{4} - c\right)}$$

Here l = lower boundry of median or quartile Class. h = Class width, f = freq of quartile class. c is C. freq. of prevous Class.

For 3rd Quartile:-

Use Same process but consider $\frac{3n}{4}$ instead of $n/4$.

Pb#3.11:- Find Quartiles.

45, 32, 37, 46, 39, 41, 48, 36.

Solution:- First arrange the given data in assending order.

32, 36, 36, 37, 39, 41, 45, 46, 48.

1st Quartile:- Since $n = 9$ So,

$\frac{n}{4} = \frac{9}{4} = 2.25 \rightarrow$ Not Integer

$\lceil \frac{n}{4} \rceil + 1 = 2 + 1 = 3.$

So the 1st quartile is 3rd digit. which is.

$Q_1 = 36$ marks

$Q_2 =$ Median

Median = ?

$\frac{n}{2} = \frac{9}{2} = 4.5 \rightarrow$ not Integer.

$[4.5] + 1 = 5^{th}$ student.

Q_2 = Median = 39 Marks.

$Q_3 = ?$

non Integer

$\frac{3n}{4} = \frac{3(9)}{4} = \frac{27}{4} = 6.75$

So; $\lceil \frac{3n}{4} \rceil + 1 = \lceil 6.75 \rceil + 1 = 6 + 1 = 7.$

So the 3rd quartile is on 7th position in data i.e $\boxed{45 \text{ Marks} = Q_3}$

Exp # 3.12 Consider the following distribution.

No. of Assistants	0	1	2	3	4	5	6	7	8	9
f	3	4	6	7	10	6	5	5	3	1

Calculate Quartile & 7th decile.

Solution:- First we have to find c. freq.

No. of Assistants	0	1	2	3	4	5	6	7	8	9
f	3	4	6	7	10	6	5	5	3	1
c. f	3	7	13	20	30	36	41	46	49	50

Q_1 = Quartile first $= \frac{n}{4} = \frac{50}{4} = 12.5$ Not Integer

So;

$\left[\frac{n}{4}\right] + 1 = 12 + 1 = 13.$

So by looking c.freq table In front of 13 we have $\boxed{Q_1 = 2}$

Now for Q_2 = Median $= \frac{n}{2} = \frac{50}{2} = 25$ Integer

25 is b/w 20 & 30 so Consider 30 &

$\boxed{Q_2 = 4}$

~~For $Q_3 = \frac{3n}{4} = \frac{3}{4}(50) = \frac{150}{4} = 37.5$~~ not Integer

~~$\left[\frac{3n}{4}\right] + 1 = 37 + 1 = 38.$~~

$Q_3 = ?$

$\frac{3n}{4} = \frac{3(50)}{4} = \frac{150}{4} = 37.5$ → not Integer.

$\left[\frac{3n}{4}\right] + 1 = 37 + 1 = 38.$

We now in C. freq whether 38 there or not
as not, so consider preceeding one

$\boxed{Q_3 = 6}$ Ans

7th decile:-

$\frac{7}{10} n = \frac{7}{10}(50) = 35$ → Integer.

See in C. freq table & consider the
preceeding range.

i.e $\boxed{Q_3 = 5}$ Ans.

Mode:- An observation which repeat mostly in
data.

In ungrouped data That will be mode which
will have highest frequency.

(*) If not value repeated mostly than we
have two theories
(i) There is no mode (ii) Every no. is mode

We will consider that every observation is mode.

For grouped data:-

① Introduce class boundry.

② Whose frequency is more that class or boundry called modal class.

Then use this formule.

$$\text{Mode} = l + h \frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)}$$

where l = lower class boundry of M.C

h = Class width.

f_m = frequency of modal class.

f_1 = Frequency of prevous class.

f_2 = frequency of preceeding class

Imperical Relation:-

It is relation between mean, mode & median. i.e

$$3(\text{Mean} - \text{Median}) = \text{Mean} - \text{Mode}$$

If the date is skewed or J shaped than Imperical formula is not used.

(4) The mode can also be calculated by $\text{Mode} = l + \frac{f_2}{f_1 + f_2}(h)$.

Ex 3.14 Calculate the mode for the distribution of examination marks given in exp 3.13

Marks	30-39	40-49	50-59	60-69	70-79	80-89	90-99
No. of students	8	87	190	304	211	85	20

Class boundry	No. of students (f)
29.5 - 39.5	8
39.5 - 49.5	87
49.5 - 59.5	190
59.5 - 69.5	304
69.5 - 79.5	211
79.5 - 89.5	85
89.5 - 99.5	20

We have modal class i.e 59.5-69.5

As we know that;

$$\text{Mode} = l + h \times \frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)}$$

$$= \cancel{29.5}\ 59.5 + 10 \times \frac{(304 - 190)}{(304-190) + (304 - 211)}$$

$$\text{Mode} = 59.5 + 5.8 \Rightarrow 65.3$$

$\boxed{M = 65 \text{ Marks}}$ Ans

Measure of Dispersion:-

As measure of Central tendency represent a single number which is not enough to represent the data.

For e.g Median or Mean of student who got in one paper 99 & in other 1 so mean is 50 similarly if other get 60 in one Subject & 40 in other so its mean mean is also 50. As mean is 50 i.e same but represent different data.

→ So we also need to represent the variability i.e dispersion.

Range:-

By range we can know how much the data is; or how large date is.

Range = Highest Value - Smallest Value

$$R = Max - Min$$

$$\boxed{R = x_m - x_0}$$

Co-efficient of Dispersion:-

Co-efficient is physical quantity which is used for comparison.

If we have two groups in which one has we measure weight and other has mass In these two groups. In order to find variation we need are a constant which is used to compare that datas. Such comparison for variation is called Co-efficient of dispersion.

$$\boxed{C.O.D = \frac{x_m - x_o}{x_m + x_o}}$$

We used this formule to find variation than used this formula & find its variation than by coefficient compare it

Example #4.1 The marks obtained by 9 students are given below.

45, 32, 37, 46, 39, 36, 41, 48, 36.

Find range & dispersion. (coefficient of dispersion)

Solution:- Range $= X_m - x_o$.

$= 48 - 32 = 16$.

Co-efficient of dispersion $= \frac{x - x_o}{x_m + x_o} = \frac{48-32}{48+32}$

$= \frac{16}{80} = 0.2$ Ans

Some times in data we have out layer

i.e for e.g consider the data 1, 2, 3, 4, 5, 99

The range = 99 - 1 = 98.

Now if we change data from mid

There is no effect on range.

Not this range properly represent the data.

So; we move to another topic i.e

Semi quartile deviation / Semi Interquartile :-

$$Q.D = \frac{Q_3 - Q_1}{2}$$

As Semi so divide by 2

→ First step is to order the data accordingly.

→ Then find Q_3 & Q_1 so the out layer will move to end. So we will find the number which will properly represent the data.

Co-efficient of Q.D :-

$$\text{Co-efficient of } Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

Scanned by CamScanner

Example # 4.2

Find the Quartile deviation & Co-efficient of Quartile deviation for the data in Exp 3.11 & 3.13.

Sol:- As $Q.D = \frac{Q_3 - Q_1}{2}$

From the data of Exp 3.11

$Q_3 = 45$ & $Q_1 = 36$

So; $Q.D = \frac{45-36}{2} = \frac{9}{2} = 4.5$

Co-efficient of $Q.D = \frac{45-36}{45+36} = \frac{9}{81} = 0.11$

(b) For Example 3.13

$Q_1 = 56$ & $Q_3 = 74$

$Q.D = \frac{Q_3 - Q_1}{2} = \frac{74-56}{2} = 9$ marks.

Coeff of $Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{74-56}{74+56} = \frac{18}{130}$

$= 0.138$ Ans

Absolute Mean deviation

or

Average deviation.

$$M.D = \frac{\sum |x_i - \mu|}{N}$$ For population.

or ~~Sample~~

where N = Total no. of observation.

For Sample

$$M.D = \frac{\sum |x_i - \bar{x}|}{n}.$$

Mean deviation tell us about the mean how variation is in data? OR how the data is distributed about the mean.

For Grouped data:-

$$M.D = \frac{\sum f |x_i - \bar{x}|}{n}$$ where $n = \sum f$.

Co-efficient of Mean deviation :-

$$\frac{M.D}{Mean} \text{ or } \frac{M.D}{Median}.$$

Example # 4.3

Calculate the mean deviation from (i) the mean (ii) The median of following. Also calculate co-efficient of mean deviation.

32, 36, 36, 37, 39, 41, 45, 46, 48.

Solution:- To find mean.

$$\bar{x} = \frac{\Sigma x}{n} = \frac{32+36+36+37+39+41+45+46+48}{9}$$

$$\boxed{\bar{x} = 40}$$

To find median:-

As $n = 9$.

so; $\frac{n}{2} = \frac{9}{2} = 4.5$ is not integer.

So median will be at $[4.5] + 1 = 5^{th}$ one

Which is 39.

x_i	$x_i - \bar{x}$	$\lvert x_i - \bar{x} \rvert$	$\lvert x_i - Med \rvert$
32	-8	8	7
36	-4	4	3
36	-4	4	3
37	-3	3	2
39	-1	1	0
41	1	1	0
45	5	5	4
46	6	6	5
48	8	8	7
		$\Sigma \lvert x_i - \bar{x} \rvert = 40$	$\Sigma \lvert \bar{x} - Med \rvert = 39$

M.D from mean $= \dfrac{\sum |x_i - \bar{x}|}{n}$

$= \dfrac{40}{9} = 4.4$

M.D from median $= \dfrac{\sum |x_i - \text{Median}|}{n}$

$= \dfrac{39}{9} = 4.3$.

Co-efficient of M.D :-

$= \dfrac{M.D}{\bar{x}}$ or $\dfrac{M.D}{\text{Median}}$

$= \dfrac{4.4}{40} \Rightarrow 0.11.$ $\Big\}$ $\dfrac{4.4}{39} = 0.1128.$

Example # 4.4 :-

Find the mean deviation of the following frequency distribution.

weights	65-84	85-104	105-124	125-144	145-164	165-184	185-204
f	9	10	17	10	5	4	5.

Solution:- First find mean for the following freq. distribution data.

Weight	x_i	f_i	fx_i	$\lvert x - x_i \rvert$	$f\lvert x_i - \bar{x} \rvert$
65 – 84	74.5	9	670.5	+48	432
85 – 104	94.5	10	945	28	280
105 – 124	114.5	17	1946.5	8	136
125 – 144	134.5	10	1345	12	120
145 – 164	154.5	5	772.5	32	160
165 – 184	174.5	4	698	52	208
185 – 204	194.5	5	972.5	72	360
		$\Sigma f = 60$	$\Sigma f x_i = 7350$		$\Sigma f \lvert x_i - \bar{x} \rvert = 169$

to find mean $= \dfrac{\Sigma f x_i}{\Sigma f} = \dfrac{7350}{60}$

$$\boxed{\bar{x} = 122.5\,g}$$

But as we know that;

$$M.D = \frac{\Sigma f \lvert x_i - \bar{x} \rvert}{n} \qquad \because n = \Sigma f$$

By putting values

$$M.D = \frac{1696}{60}$$

$$\boxed{M.D = 25.27 \text{ grams}}$$ Ans

Variance & standard deviation.

For population:-

Variance is;

$$\delta^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}.$$

For Sample:-

$$\delta^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})}{n}.$$

Standard deviation:-

$$S.D = \sqrt{\delta^2} = \sqrt{\text{Variance}}.$$

For ^un^ Grouped data:-

$$\delta^2 = \frac{\sum_{i=1}^{n} f(x_i - \bar{x})^2}{\Sigma f}.$$

(*) For good estimator, the data should be unbaised. i.e divide by $n-1$. instead of n. i.e $\Sigma f - 1$.

Short method:-

This is short method to find variance & standard deviation.

As

- Variance $= \delta^2 = \dfrac{\Sigma (x_i - \mu)^2}{N}$

$$= \frac{\Sigma (x_i^2 - 2\mu x_i + \mu^2)}{N}$$

$$= \frac{\Sigma x_i^2}{N} - 2\mu \frac{\Sigma x_i}{N} + \cancel{\Sigma} \frac{\mu^2}{N}$$

$$= \frac{\Sigma x_i^2}{N} - 2\mu\,\mu + \cancel{N}\frac{\mu^2}{\cancel{N}}$$

$$= \frac{\Sigma x_i^2}{N} - 2\mu^2 + \mu^2$$

$$= \frac{\Sigma x_i^2}{N} - \mu^2$$

$$= \Sigma \frac{x_i^2}{N} - \left(\frac{\Sigma x_i}{N}\right)^2$$

Scanned by CamScanner

Example #4.5

A population of $N = 10$ has the observation 7, 8, 10, 13, 14, 19, 20, 25, 26 & 28
Find its variance & S.D.

Sol:- As we know that

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N} \longrightarrow ①$$

to find $\mu = ?$

$$\mu = \frac{\sum x_i}{N} = \frac{170}{10} = 17$$

S.No	x_i	$x_i - \mu$	$(x_i - \mu)^2$	x_i^2
1	7	-10	100	49
2	8	-9	81	64
3	10	-7	49	100
4	13	-4	16	169
5	14	-3	9	196
6	19	2	4	361
7	20	3	9	400
8	25	8	64	625
9	26	9	81	676
10	28	11	121	784
	$\sum x_i = 170$		$\sum (x_i - \mu)^2 = 534$	$\sum x_i^2 = 3424$

eq ① $\Rightarrow$ $\sigma^2 = \frac{534}{10} \Rightarrow \boxed{53.4 = \text{Variance}}$

Deviation $= \sqrt{\sigma^2} \Rightarrow \sqrt{53.4}$

$$\boxed{S.D = 7.31}$$

By using short method.

As;

$$\sigma^2 = \frac{\Sigma x_i^2}{N} - \left(\frac{\Sigma x_i}{N}\right)^2$$

$$= \frac{3424}{10} - \left(\frac{170}{10}\right)^2$$

$$= \cancel{3424}\ 342.4 - 289.$$

$$\boxed{\sigma^2 = 53.4}$$

$$S.D = \sqrt{\sigma^2} = \text{or} \sqrt{\frac{\Sigma x_i^2}{N} - \left(\frac{\Sigma x_i}{N}\right)^2}$$

$$= \sqrt{53.4} = 7.31 \quad \underline{\text{Ans}}.$$

Example:-

Calculate the variance & S.D of the following data

45, 32, 37, 46, 39, 36, 41, 48, 36.

Sol:- As $\sigma^2 = \frac{\Sigma |x_i - \bar{x}|}{n}$

$\bar{x} = ?$

36
45
32
37
46
39
36
41
48

First find $\bar{x} = ?$

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{360}{9} = 40$$

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	x_i^2
45	5	25	2025
32	-8	64	1024
37	-3	9	1369
46	6	36	2116
39	-1	1	1521
36	-4	16	1296
41	1	1	1681
48	8	64	2304
36	-4	16	1296
$\Sigma x_i = 360$		$\Sigma (x_i - \bar{x})^2 = 232$	$\Sigma x_i^2 = 14632$

$$\delta^2 = \frac{\Sigma (x_i - \bar{x})^2}{n}$$

$$= \frac{232}{9} = 25.78$$

$$\text{S.D } \cancel{\text{Variance}} = \sqrt{\delta^2} \Rightarrow \sqrt{25.78}$$

$$\boxed{\text{S.D} = 5.08}$$

→ **By short method**

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

$$= \frac{14632}{9} - \left(\frac{360}{9}\right)^2$$

$$\boxed{\delta^2 = 25.78}$$

$$\text{S.D} = \sqrt{25.78} \Rightarrow \boxed{5.08 \text{ marks}}$$ Ans

Scanned by CamScanner

Example # 4.7 Calculate the variance & S.D from the data as given below.

x_i	f_i	$f_i x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i(x_i-\bar{x})^2$
74.5	9	670.5	-48	2304	20736
94.5	10	945	-28	784	7840
114.5	17	1946.5	-8	64	1088
134.5	10	1345	12	144	1440
154.5	5	772.5	32	1024	5120
174.5	4	698	52	2704	10816
194.5	5	972.5	72	5184	25920
	$\Sigma f = 60$	$\Sigma f x_i = 7350$		$\Sigma(x_i-\bar{x})^2 = 12208$	$\Sigma f(x_i-\bar{x})^2 = 72960$

Sol:- As we know that variance for grouped freq. data.

$$\text{Variance} = \delta^2 = \frac{\Sigma f(x_i-\bar{x})^2}{\Sigma f}$$

$$\bar{x} = \frac{\Sigma f x_i}{\Sigma f}$$

$$\bar{x} = \frac{7350}{60} \Rightarrow \boxed{\bar{x} = 122.5}$$

$$\delta^2 = \frac{72960}{60} \Rightarrow \boxed{\delta^2 = 1216}$$

Short method:-

$$\delta^2 = \frac{\Sigma f x_i^2}{n} - \left(\frac{\Sigma f x_i}{n}\right)^2$$

$$= \frac{973335}{60} - \left(\frac{7350}{6}\right) \Rightarrow \boxed{\delta^2 = 1216}$$

$$S.D = \sqrt{1216} \Rightarrow \boxed{SD = 34.87}$$ Ans

Change of Scale & Origin:-

If we have a large data or such number which are very large so we Introduce a new number say u which

is; $u_i = \dfrac{x_i - a}{h}$.

a - is observation whose freq. high

h is class width

For mean

$\bar{u} = \dfrac{\bar{x} - a}{h}$.

$\bar{x} = a + h\bar{u}$.

$$Var(u) = \sigma^2 = \frac{1}{h^2}\,(var(x))$$

or

$$var(x) = h^2 * var(u).$$

It is better to use this formula or This method.

Example #4.8:- Find the standard deviation by short method from the data.

x_i	f_i	u_i	fu_i	fu_i^2
74.5	9	-2	-18	36
94.5	10	-1	-10	10
114.5	17	0	0	0
134.5	10	1	10	10
154.5	5	2	10	20
174.5	4	3	12	36
1945	5	4	20	80
$\Sigma f = 60$			$\Sigma fu = 24$	$\Sigma fu_i^2 = 192$

As by short method standard deviation is;

$$S.D = \sqrt{\frac{\Sigma f x_i^2}{\Sigma f} - \left(\frac{\Sigma f x_i}{\Sigma f}\right)^2}$$

But by using change of origin & scale than.

$$S.D = \sqrt{h^2 \, var(u)}$$

$$= h \sqrt{var(u)} \longrightarrow (1)$$

To find $var(u) = \frac{\Sigma f_i u_i^2}{\Sigma f} - \left(\frac{\Sigma f u}{\Sigma f}\right)^2 \longrightarrow (ii)$

$$u = \frac{x_i - a}{h}$$

$$u = \frac{74.5 - 114.5}{20} \Rightarrow -2.$$

eq (ii) $\Rightarrow$ $var(u) = \frac{192}{60} - \left(\frac{24}{60}\right)^2 \Rightarrow var(u) = 3.2 - 0.16.$

$var(u) = 3.04.$

eq (1) $\Rightarrow$ $S.D = h\sqrt{var(u)} = 20\sqrt{3.04}$

$\boxed{S.D = 34.87}$ Ans

Chabishi rule:-

It tells us if you give S.D & mean than I can tell u an interval which will show about max. data range.

that interval is;

$$(\bar{x} - k\,S.D\ ,\ \bar{x} + k\,S.D)$$

k should be > 1.

This interval will contain at least

$(1 - 1/k^2)$ fraction of data.

For eg if k = 2.

$$(1 - 1/4) = 3/4 = 75\%.$$

at least represent 75% of the data

For e.g

For previous example:

$\bar{x} = 122.5.$

$S.D = 34.87.$

If let k = 2.

∵ As k increase the degree of accuracy increases.

$$(122.5 - 2(34.87)\ ,\ 122.5 + 2(34.87))$$

which is at least 75% of data representation.

$(52.82\ ,\ 192.1)$ Interval

lower class limit

Higher class limit

Co-efficient of Variation:-

$$C.V = \frac{\delta}{\mu} * 100\%$$ For population

$$C.V = \frac{\delta}{\bar{x}} * 100\%$$

The more C.V value for data the more there is variation in that data & vice versa.

→ The data which have less coefficient of variation than the data is more consistant.

Example #4.9 Using the co-efficient of variation determine whether or not there is gap variation among the prices of certain similar commidities given, than among the life in hour under.

Price In Rupees : 8, 13, 18, 23, 30.

Life in hours 130, 150, 180, 250, 345.

Sol:- For Co-efficient of variation;

We have required standard deviation and mean.

Since we know that;

$$C.V = \frac{S.D}{\bar{x}} \times 100 \longrightarrow (A)$$

To find $\bar{x} = ?$

$$\bar{X} = \frac{\sum^{n} x_i}{n} \Rightarrow \frac{92}{5} = 18.4$$

For other data.

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

$$\bar{x} = \frac{1055}{5}$$

$= 211$ hours.

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

For Price in Rs/-.

x_i	x_i^2
8	64
13	169
18	324
23	529
30	900
$\sum x_i = 92$	$\sum x_i^2 = 1986$

For life In hours

x_i	x_i^2
130	16900
150	22500
180	32400
250	62500
345	119025
1055	253325

Price In Rs-/.

$$S.D = \sqrt{\frac{1986}{5} - \left(\frac{92}{5}\right)^2}$$

$S.D = 7.66$ Rs/-

& (A). $C.V = \frac{7.66}{18.4} \times 100$

$\boxed{C.V = 41.6\ \%}$

For life In hours.

$$S.D = \sqrt{\frac{253325}{5} - \left(\frac{1055}{5}\right)}$$

$S.D = 78.38$ hours

& $C.V = \frac{78.38}{211} \times 100$

$\boxed{C.V = 37.15\%}$

From the two data having C.V. So Price in Rs/- has greater variation as compared to life in hours.

Example # 4.10 Goals scored by two teams are as followed. Find C.V to tell which team is more consistant.

No. of Goals Scored (x_i)	Number of Matches A	Number of Matches B
0	27	17
1	9	9
2	8	6
3	5	5
4	4	3

Sol:- First we have to find co-efficient of variation for both teams.

As $C.O.V = \frac{S\,(S.D)}{\bar{X}}$. $\longrightarrow$ ①.

Find mean & standard deviation.

No. of goals (x_i)	Team A f_i	$f_i x_i$	$f_i x_i^2$	Team B f_i	$f_i x_i$	$f_i x_i^2$
0	27	0	0	17	0	0
1	9	9	9	9	9	9
2	8	16	32	6	12	24
3	5	15	45	5	15	45
4	4	16	64	3	12	48
	$\Sigma f_i = 53$	$\Sigma f x_i = 56$	$\Sigma f x_i^2 = 150$	$\Sigma f = 40$	$\Sigma f_i x_i = 48$	$\Sigma f_i x_i^2$

A3 $\bar{x} = \frac{\Sigma f x_i}{\Sigma f}$.

& $S.D = \sqrt{\text{Variance}}$

& $\text{Variance} = \frac{\Sigma f(x_i - \bar{x})}{\Sigma f}$.

or

$$\frac{\Sigma f x_i^2}{n} - \left(\frac{\Sigma f x_i}{n}\right)^2 \rightarrow ①.$$

Using eq ①.

$$S.D = \sqrt{\text{Variance}} = \sqrt{\frac{\Sigma f x_i^2}{\Sigma f} - \left(\frac{\Sigma f x_i}{\Sigma f}\right)^2}$$

<u>For team A:-</u>

$$S.D = \sqrt{\frac{150}{53} - \left(\frac{\cancel{1}56}{53}\right)^2} \Rightarrow S.D = 1.308.$$

$$\bar{x} = \frac{\Sigma f x_i}{\Sigma f} = \frac{56}{53} = 1.06.$$

$$C.V = \frac{S.D \times 100}{\bar{x}} \Longrightarrow \frac{1.308}{1.06} \times 100 \Rightarrow 123.4\%.$$

<u>For team B:-</u> $\bar{x} = \frac{\Sigma f x_i}{\Sigma f} = \frac{48}{40} = 1.20$.

$$S.D = \sqrt{\frac{\Sigma f x_i^2}{\Sigma f_i} - \left(\frac{\Sigma f x_i}{\Sigma f_i}\right)^2}$$

$$= \sqrt{\frac{126}{40} - \left(\frac{48}{40}\right)^2} \Rightarrow S.D = 1.308$$

$$C.V = \frac{S.D \times 100}{\bar{x}} = \frac{1.308}{1.20} \times 100.$$

$$\boxed{C.V = 109\%}$$

As the C.V of team B is small as compare to team A so it means team B is more consistant than team A