Statistics

1. Measures of central tendency :

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

(A) Mathematical average
(i) Arithmetic mean or mean
(ii) Geometrical mean
(iii) Harmonic mean

(B) Positional average
(i) Median
(ii) Mode

2. Mean (Arithmetic mean)

If $x_1, x_2, x_3, ...x_n$ are n values of variate x_i then their A.M. $\bar{x}$ is defined as

$$\bar{x} = \frac{x_1 + x_2 + x_3 + ... + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

If $x_1, x_2, x_3, ... x_n$ are values of veriate with frequencies $f_1, f_2, f_3, ...f_n$ then their A.M. is given by

$$\bar{x} = \frac{f_1x_1 \times f_2x_2 + f_3x_3 + ...f_nf_n}{f_1 + f_2 + f_3 + ... + f_n} = \frac{\sum_{i=1}^{n} f_i x_i}{N}, \text{ where } N = \sum_{i=1}^{n} f_i$$

(i) Properties of arithmetic mean :

(a) Sum of deviation of variate from their A.M. is always zero that is $\Sigma(x_i - \bar{x}) = 0$.

(b) Sum of square of deviation of variate from their A.M. is minimum that is $\Sigma(x_i - \bar{x})^2$ is minimum

(c) If $\bar{x}$ is mean of variate x_i then

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

A.M. of $\lambda_i . x_i = \lambda . \bar{x}$

A.M. of $(ax_i + b) = a\bar{x} + b$

(ii) Merits of arithmetic mean :

(a) It is rigidly defined.
(b) It is based on all the observation taken.
(c) It is calculated with reasonable ease.
(d) It is least affected by fluctuations in sampling.
(e) It is based on each observation and so it is a better representative of the data.
(f) It is relatively reliable
(g) Mathematical analysis of mean is possible.

(iii) Demerits of Arithmetic Mean :

(a) It is severely affected by the extreme values.

(b) It cannot be represented in the actual data since the mean does not coincide with any of the observed value.

(c) It cannot be computed unless all the items are known.

Example # 1 : Find mean of data 2, 4, 5, 6, 8, 17.

Solution : Mean = $\frac{2+4+5+6+8+17}{6}$ = 7

Example # 2 : Find the mean of the following distribution :

x :	4	6	9	10	15
f :	5	10	10	7	8

Solution : Calculation of Arithmetic Mean

x_i	f_i	f_ix_i
4	5	20
6	10	60
9	10	90
10	7	70
15	8	120
	$N = \sum f_i = 40$	$\sum f_ix_i = 360$

∴ Mean = $\bar{X} = \frac{\sum f_i x_i}{\sum f_i} = \frac{360}{40} = 9$

Example # 3 : Find the mean wage from the following data :

Wage (in Rs) :	800	820	860	900	920	980	1000
No. of workers :	7	14	19	25	20	10	5

Solution : Let the assumed mean be A = 900 and h = 20.

Calculation of Mean

Wage (in Rs) x_i	No. of workers f_i	$d_i = x_i - A = x_i - 900$	$u_i = \frac{x_i - 900}{20}$	f_iu_i
800	7	–100	–5	–35
820	14	–80	–4	–56
860	19	–40	–2	–38
900	25	0	0	0
920	20	20	1	20
980	10	80	4	40
1000	5	100	5	25
	$N = \sum f_i = 100$			$\sum f_i u_i = -44$

We have,

$N = 100$, $\sum f_iu_i = -44$, $A = 900$ and $h = 20$

∴ Mean = $\bar{X} = A + h\left(\frac{1}{N}\sum f_iu_i\right)$

⇒ $\bar{X} = 900 + 20 \times \frac{-44}{100} = 900 - 8.8 = 891.2$

Hence, mean wage = Rs. 891.2

3. Geometric mean :

If x_1, x_2, x_3 ,.......x_3 are n positive values of variate then their geometric mean G is given by G = $(x_1\, x_2\, x_3\, ...\, x_n)^{1/n}$

$\Rightarrow \quad G = \text{antilog}\left[\frac{1}{n}\sum_{i=1}^{n}\log \; x_i\right]$

4. Median :

The median of a series is values of middle term of series when the values are written is ascending order or descending order. Therefore median, divide on arranged series in two equal parts

(i) For ungrouped distribution :

If n be number of variates in a series then

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

(ii) For ungrouped frequency distribution :

First we calculate cumulative frequency (sum of all frequencies). Let it be N then

$$\text{Median} \begin{cases} \left(\frac{N+1}{2}\right)^{th} \text{term (when n odd)} \\ \text{Mean of } \left(\frac{N}{2}\right) \& \left(\frac{N+2}{2}\right) \text{ (when n is even)} \end{cases}$$

5. Merits and demerits of median :

The following are some merits and demerits of median :

(i) Merits :

(a) It is easy to compute and understand.

(b) It is well defined an ideal average should be

(c) It can also be computed in case of frequency distribution with open ended classes.

(d) It is not affected by extreme values.

(e) It can be determined graphically.

(f) It is proper average for qualitative data where items are not measured but are scored.

(ii) Demerits :

(a) For computing median data needs to be arranged in ascending or descending order.

(b) It is not based on all the observations of the data.

(c) It cannot be given further algebraic treatment.

(d) It is affected by fluctuations of sampling.

(e) It is not accurate when the data is not large.

(f) In some cases median is determined approximately as the mid-point of two observations whereas for mean this does not happen.

Example # 4 : Find the median of observations 4, 6, 9, 4, 2, 8, 10

Solution : Values in ascending order are 2, 4, 4, 6, 8, 9, 10 here n = 7 so $\frac{n+1}{2}$ = 4

so median = 4_{th} observaiton = 6

Example # 5 : Obtain the median for the following frequency distribution :

x	11	13	15	18	21	23	30	40	50
f	8	10	11	16	20	25	15	9	6

Solution :

x	f	cf
11	8	8
13	10	18
15	11	29
18	16	45
21	20	65
23	25	90
30	15	105
40	9	114
50	6	120
	N = 120	

Here, N = 120 ⇒ $\frac{N}{2}$ = 60

We find that the cummulative frequency just greater than $\frac{N}{2}$ i.e., 60 is 65 and the value of x corresponding to 65 is 21. Therefore, Median = 21.

6. Harmonic Mean :

If x_1, x_2, x_3 ,.......x_3 are n non-zero values of variate then their harmonic mean H is defined as

$$H = \frac{n}{\frac{1}{x_i}+\frac{1}{x_2}+....+\frac{1}{x_n}} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}$$

7. Mode :

If a frequency distribution the mode is the value of that variate which have the maximum frequency. Mode for

(i) For ungrouped distribution :

The value of variate which has maximum frequency.

(ii) For ungrouped frequency distribution :

The value of that variate which have maximum frequency.

Relationship between mean, median and mode.

- In symmetric distribution, mean = mode = median
- In skew (moderately asymmetrical) distribution,
  median divides mean and mode internally in 1 : 2 ratio.

⇒ median = $\frac{2(\text{Mean})+(\text{Mode})}{3}$

8. Merits and demerits of mode :

The following are some merits and demerits of mode :

(i) Merits :

(a) It is readily comprehensible and easy to compute. In some case it can be computed merely by inspection.

(b) It is not affected by extreme values. It can be obtained even if the extreme values are not known.

(c) Mode can be determined in distributions with open classes.

(d) Mode can be located on graph also.

(ii) Demerits :

(a) It is ill-defined. It is not always possible to find a clearly defined mode. In some cases, we may come across distributions with two modes. Such distributions are called bimodal. If a distribution has more than two modes, it is said to be multimodal.

(b) It is not based upon all the observation.

(c) Mode can be calculated by various formulae as such the value may differ from one to other. Therefore, it is not rigidly defined.

(d) It is affected to a greater extent by fluctuations of sampling.

Example # 6 : Find mode of data 2, 4, 6, 8, 8, 12, 17, 6, 8, 9.

Solution : 8 occurs maximum number of times so mode = 8

9. Measure of dispersion :

It is measure of deviation of its value about their central values. It gives an idea of scatterdness of different values from the central values.

Types :

(i) Range :

Difference between greatest values & least values of variates of a distribution are called the range of distribution.

Also coefficient of range = $\frac{\text{difference of extreme values}}{\text{sum of extreme values}} = \frac{L-S}{L+S}$

where L = largest value and S = smallest value

(ii) Mean deviation :

Mean deviation of a distribution is, the mean of absolute value of deviation of variate from their statistical average (median, mean or mode).

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

Mean deviation = $\frac{\sum_{i=1}^{n} |x_i - A|}{n}$

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

Example # 7 : Calculate mean deviation about median for the following data 3, 9, 5, 3, 12, 10, 18, 4, 7, 19, 21.

Solution : Data in ascending order is 3, 3, 4, 5, 7, 9, 10, 12, 18, 19, 21

Median = $\frac{n+1}{2}$ th value = 6_{th} value = 9

Mean deviation about median = $\frac{\sum_{i=1}^{11} |x_i - \text{median}|}{11} = \frac{58}{11}$

Example # 8 : Find mean deviation from mean

x	5	7	9	10	12	15
f	8	6	2	2	2	6

Solution :

x	f	fx	$x-\bar{x}$	$\lvert x-\bar{x}\rvert$	$f\lvert x-\bar{x}\rvert$
5	8	40	–4	4	32
7	6	42	–2	2	12
9	2	18	0	0	0
10	2	20	1	1	2
12	2	24	3	3	6
15	6	90	6	6	36
	N = 26	$\sum fx = 234$			$\sum f\lvert x-\bar{x}\rvert = 88$

$\bar{x} = \frac{\sum fx}{\sum f} = 9$ Now, M.D. $(\bar{x}) = \frac{\sum f\lvert x-\bar{x}\rvert}{\sum f} = \frac{88}{26} = 3.38$

(iii) **Variance** :

It 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.

so standard deviation = + $\sqrt{\text{variance}}$ formula

$\sigma_x^2 = \frac{\Sigma (x_i - \bar{x})^2}{n}$ ⇒ $\sigma_x^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - (\bar{x})^2$

$\sigma_d^2 = \frac{\Sigma d_i^2}{n} - \left(\frac{\Sigma d_i}{n}\right)^2$, where $d_i = x_i - a$, where a = assumed mean

coefficient of S.D. = $\left(\frac{\sigma}{\bar{x}}\right)$ ⇒ coefficient of variation = $\left(\frac{\sigma}{\bar{x}}\right) \times 100$ (in percentage)

(a) **Properties of variance :**

• $\text{var}(x_i + \lambda) = \text{var}(x_i)$ • $\text{var}(\lambda . x_i) = \lambda^2(\text{var } x_i)$ • $\text{var}(a\, x_i + b) = a^2(\text{var } x_i)$

where λ, a, b are constant.

Example # 9 : Find the mean and variance of first n natural numbers.

Solution : $\bar{x} = \frac{\sum x}{n} = \frac{1+2+3+.\ .\ .\ .\ .+\ n}{n} = \frac{n+1}{2}$

Variance = $\frac{\sum x^2}{n} - (\bar{x})^2 = \frac{1^2+2^2+3^2+.\ .\ .\ .\ .+\ n^2}{n} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}$

Example # 10 : Find the variance and standard deviation of the following frequency distribution :

Variable (x_i)	2	4	6	8	10	12	14	16
Frequency (f_i)	4	4	5	15	8	5	4	5

Solution : Calculation of variance and standard deviation

Variable x_i	Frequency f_i	$f_i x_i$	$x_i - \overline{X}$ $= x_i - 9$	$(x_i - \overline{X})^2$	$f_i(x_i - \overline{X})^2$
2	4	8	–7	49	196
4	4	16	–5	25	100
6	5	30	–3	9	45
8	15	120	–1	1	15
10	8	80	1	1	8
12	5	60	3	9	45
14	4	56	5	25	100
16	5	80	7	49	245
	$N = \sum f_i = 50$	$\sum f_i x_i = 450$			$\sum f_i(x_i - \overline{X})^2 = 754$

Here N = 50, $\sum f_i x_i = 450$

∴ $\overline{X} = \frac{1}{N}(\sum f_i x_i) = \frac{450}{50} = 9$

We have $\sum f_i(x_i - \overline{X})^2 = 754$

∴ $\text{Var}(X) = \frac{1}{N}\left[\sum f_i (x_i - \overline{X})^2\right] = \frac{754}{50} = 15.08$

$\text{S.D.} = \sqrt{\text{Var}(X)} = \sqrt{15.08} = 3.88$

Example # 11 : Calculate the mean and standard deviation for the following data :

Wages upto (in Rs.)	15	30	45	60	75	90	105	120
No. of workers	12	30	65	107	157	202	222	230

Solution : We are given the cummulative frequency distribution. So first we will prepare the frequency distribution as given below :

Class Interval	Cummlative frequency	Mid–values	Frequency	$u_i = \frac{x_i - 67.5}{15}$	$f_i u_i$	$f_i u_i^2$
0 – 15	12	7.5	12	–4	–48	192
15 – 30	30	22.5	18	–3	–54	162
30 – 45	65	37.5	35	–2	–70	140
45 – 60	107	52.5	42	–1	–42	42
60 – 75	157	67.5	50	0	0	0
75 – 90	202	82.5	45	1	45	45
90 – 105	222	97.5	20	2	40	80
105 – 120	230	112.5	8	3	24	72
			$\sum f_i = 230$		$\sum f_i u_i = -105$	$\sum f_i u_i^2 = 733$

Here A = 67.5, h = 15, N = 230, $\sum f_i u_i = -105$ and $\sum f_i u_i^2 = 733$

∴ $\text{Mean} = A + h\left(\frac{1}{N}\sum f_i u_i\right) = 67.5 + 15\left(\frac{-105}{230}\right) = 67.5 - 6.85 = 60.65$

and $\text{Var}(X) = h^2\left[\frac{1}{N}\sum f_i u_i^2 - \left(\frac{1}{N}\sum f_i u_i\right)^2\right]$

⇒ $\text{Var}(X) = 225\left[\frac{733}{230} - \left(\frac{-105}{230}\right)^2\right] = 225[3.18 - 0.2025] = 669.9375$

∴ $\text{S.D.} = \sqrt{\text{Var}(X)} = \sqrt{669.9375} = 25.883$

10. For NCERT and Board purpose :

(i) Arithmetic mean :

Arithmetic mean of continuous grouped data :

Take mid points of given classes as x_i and use formula as given for discrete grouped data.

Example # 12 : Find the mean of the following frequency distribution :

Class-interval :	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50
No. of workers f :	7	10	15	8	10

Solution : Calculation of Mean

Class-interval	Mid-values (x_i)	Frequency f_i	$d_i = x_i - 25$	$u_i = \frac{x_i - 25}{10}$	f_iu_i
0 – 10	5	7	–20	–2	–14
10 – 20	15	10	–10	–1	–10
20 – 30	25	15	0	0	0
30 – 40	35	8	10	1	8
40 – 50	45	10	20	2	20
		$N = \sum f_i = 50$			$\sum f_iu_i = 4$

We have,

$A = 25$, $h = 10$, $N = 50$ and $\sum f_iu_i = 4$.

$\therefore$ Mean $= A + h\left\{\frac{1}{N}\sum f_iu_i\right\}$ $\Rightarrow$ mean $= 25 + 10 \times \frac{4}{50} = 25.8$

Example # 13 : Find the mean marks of students from the following cumulative frequency distribution :

Marks	Number of students	Marks	Number of students
0 and above	80	60 and above	28
10 and above	77	70 and above	16
20 and above	72	80 and above	10
30 and above	65	90 and above	8
40 and above	55	100 and above	0
50 and above	43		

Solution : Here we have, the cumulative frequency distribution. So, first we convert it into an ordinary frequency distribution. We observe that there are 80 students getting marks greater than or equal to 0 and 77 students have secured 10 and more marks. Therefore, the number of students getting marks between 0 and 10 is 80 – 77 = 3.

Similarly, the number of students getting marks between 10 and 20 is 77 – 72 = 5 and so on. Thus, we obtain the following frequency distribution.

Marks	Number of students	Marks	Number of students
0 – 10	3	50 – 60	15
10 – 20	5	60 – 70	12
20 – 30	7	70 – 80	6
30 – 40	10	80 – 90	2
40 – 50	12	90 – 100	8

Now, we compute arithmetic mean by taking 55 as the assumed mean.
Computation of Mean

Marks	Mid-value (x_i)	Frequency (f_i)	$u_i = \frac{x_i - 55}{10}$	f_iu_i
0 – 10	5	3	–5	–15
10 – 20	15	5	–4	–20
20 – 30	25	7	–3	–21
30 – 40	35	10	–2	–20
40 – 50	45	12	–1	–12
50 – 60	55	15	0	0
60 – 70	65	12	1	12
70 – 80	75	6	2	12
80 – 90	85	2	3	6
90 – 100	95	8	4	32
Total		$\sum f_i = 80$		$\sum f_iu_i = -26$

We have,

$N = \sum f_i = 80$, $\sum f_iu_i = -26$, $A = 55$ and $h = 10$

$$\therefore \quad \overline{X} = A + h\left\{\frac{1}{N}\sum f_iu_i\right\} \Rightarrow \overline{X} = 55 + 10 \times \frac{-26}{80} = 55 - 3.25 = 51.75 \text{ Marks}$$

(ii) Median :

(a) Median of continuous frequency distribution :

Let the number of observation be N. Prepare the cumulative frequency table. Find the median class i.e. the class in which the observation whose cumulative frequency is equal to or just greater than $\frac{N}{2}$ lies.

The median value is given by the formula : Median

$$(M) = \ell + \left[\frac{\left(\frac{N}{2}\right) - c}{f}\right] \times h \text{ where}$$

N = total frequency = $\sum f_i$

ℓ = lower limit of median class

f = frequency of the median class

c = cumulative frequency of the class preceding the median class

h = class interval (width) of the median class

Example # 14 : Calculate the median from the following distribution :

Class	5 – 10	10 – 15	15 – 20	20 – 25	25 – 30	30 – 35	35 – 40	40 – 45
Frequency	5	6	15	10	5	4	2	2

Solution :

Class	Frequency	Cumulative Frequency
5 – 10	5	5
10 – 15	6	11
15 – 20	15	26
20 – 25	10	36
25 – 30	5	41
30 – 35	4	45
35 – 40	2	47
40 – 45	2	49
	N = 49	

We have, N = 49 ∴ $\frac{N}{2} = \frac{49}{2}$ = 24.5

The cummulative frequency just greater than $\frac{N}{2}$ is 26 and the corresponding class is 15-20. Thus 15-20 is the median class such that ℓ = 15, f = 15, F = 11 and h = 5

$$\therefore \text{ Median} = \ell + \frac{\frac{N}{2} - F}{f} \times h = 15 + \frac{24.5 - 11}{15} \times 5 = 15 + \frac{13.5}{3} = 19.5$$

(iii) **Mode :**

(a) **Mode for continuous frequency distribution :**

First find the modal class i.e. the class which has maximum frequency. The modal class can be determined either by inspecting or with the help of grouping.

The mode is given by the formula :

$$\text{Mode} = \ell + \frac{f_m - f_{m-1}}{2f_m - f_{m-1} - f_{m+1}} \times h$$

where ℓ = lower limit of the modal class

h = width of the modal class

f_m = frequency of the modal class

f_{m-1} = frequency of the class preceding modal class

f_{m+1} = frequency of the class succeeding modal class

Example # 15 : Compute the mode for the following frequency distribution :

Size of items	0 – 4	4 – 8	8 – 12	12 – 16	16 – 20	20 – 24	24 – 28	28 – 32	32 – 36	36 – 40
Frequency	5	7	9	17	12	10	6	3	1	0

Solution : Here, the maximum frequency is 17 and the corresponding class is 12-16 So 12-16 is the modal class.

We have, ℓ = 12, h = 4, f = 17, f_1 = 9 and f_2 = 12

$$\therefore \text{Mode} = \ell + \frac{f - f_1}{2f - f_1 - f_2} \times h \Rightarrow \quad \text{Mode} = 12 + \frac{17 - 9}{34 - 9 - 12} \times 4$$

$$\Rightarrow \text{Mode} = 12 + \frac{8}{13} \times 4 = 12 + \frac{32}{13} = 12 + 10.66 = 32.66$$

(iv) **Mean deviation :**

(a) **Mean deviation of continuous frequency distribution :** For calculating mean deviation of a continuous frequency distribution, the procedure is same as for a discrete frequency distribution. The only difference is that here we have to obtain the midpoints of the various classes and take the deviations of these mid point from the given average A.

Example # 16 : Find the mean deviation about the median of the following frequency distribution :

Class	0 – 6	6 – 12	12 – 18	18 – 24	24 – 30
Frequency	8	10	12	9	5

Solution : Calculation of mean deviation about the median

Class	Midvalues (x_i)	Frequency (f_i)	Cumulative Frequency (c.f.)	$\|x_i - 14\|$	$f_i\|x_i - 14\|$
0 – 6	3	8	8	11	88
6 – 12	9	10	18	5	50
12 – 18	15	12	30	1	12
18 – 24	21	9	39	7	63
24 – 30	27	5	44	13	65
			$N = \sum f_i = 44$		$\sum f_i\|x_i - 14\| = 278$

Here N = 44, so $\frac{N}{2}$ = 22 and the cumulative frequency just greater than $\frac{N}{2}$ is 30. Thus 12-18 is the median class.

Now Median = $\ell + \frac{N/2 - F}{f} \times h$, where ℓ = 12, h = 6, f = 12, F = 18

or Median = $12 + \frac{22 - 18}{12} \times 6 = 12 + \frac{4 \times 6}{12} = 14$

Mean deviation about median = $\frac{1}{N} \sum f_i \, |x_i - 14| = \frac{278}{44} = 6.318$

Example # 17 : Find the mean deviation from the mean for the following data :

Classes	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70	70 – 80
Frequencies	2	3	8	14	8	3	2

Solution : We prepare the table as follows :

Computation of mean deviation from mean

Classes	Mid–values (x_i)	frequencies f_i	$f_i\, x_i$	$\|x_i - \overline{X}\| = \|x_i - 45\|$	$f_i\|x_i - \overline{X}\|$
10 – 20	15	2	30	30	60
20 – 30	25	3	75	20	60
30 – 40	35	8	280	10	80
40 – 50	45	14	630	0	0
50 – 60	55	8	440	10	80
60 – 70	65	3	195	20	60
70 – 80	75	2	150	30	60
		$n = \sum f_i = 40$	$\sum f_i x_i = 1800$		$\sum f_i\|x_i - \overline{X}\| = 400$

We have, N = 40 and $\sum f_i \; x_i = 1800$

$\therefore \; \overline{X} = \frac{\sum f_i \; x_i}{N} = \frac{1800}{40} = 45$

Now $\sum f_i \; |x_i - \overline{X}| = 400$ and $N = \sum f_i = 40$

$\therefore$ M.D. = $\frac{1}{N} \sum f_i \; |x_i - \overline{X}| = \frac{400}{40} = 10$

(v) Variance and standard deviation :

(a) **Variance of a grouped or continuous frequency distribution :** In a grouped or continuous frequency distribution, any of the formulae discussed in discrete frequency distribution can be used.

Example # 18 : Calculate the mean and standard deviation for the following distribution :

Marks	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70	70 – 80	80 – 90
No. of Students	3	6	13	15	14	5	4

Solution : Calculation of Standard deviation

Class interval	Frequency f_i	Mid-values x_i	$u_i = \frac{x_i - 55}{10}$	$f_i u_i$	u_i^2	$f_i u_i^2$
20 – 30	3	25	–3	–9	9	27
30 – 40	6	35	–2	–12	4	24
40 – 50	13	45	–1	–13	1	13
50 – 60	15	55	0	0	0	0
60 – 70	14	65	1	14	1	14
70 – 80	5	75	2	10	4	20
80 – 90	4	85	3	12	9	36
	$N = \sum f_i = 60$			$\sum f_i u_i = 2$		$\sum f_i u_i^2 = 134$

Here N = 60, $\sum f_i u_i = 2$, $\sum f_i u_i^2 = 134$ and h = 10

∴ Mean = $\bar{X} = A + h\left(\frac{1}{N}\sum f_i u_i\right)$ ⇒ $\bar{X} = 500 + 10\left(\frac{2}{60}\right) = 55.333$

Var(X) = $h^2\left[\frac{1}{N}\sum f_i u_i^2 - \left(\frac{1}{N}\sum f_i u_i\right)^2\right] = 100\left[\frac{134}{60} - \left(\frac{2}{60}\right)^2\right] = 222.9$

∴ S.D. = $\sqrt{\text{Var}(X)} = \sqrt{222.9} = 14.94$

Example # 19 : Suppose that samples of polythene bags from two manufactures, A and B are tested by a prospective buyer for bursting presure, with the following results :

Bursting presure in kg	Number of bags manufactued by manufactuer A	B
5 – 10	2	9
10 – 15	9	11
15 – 20	29	18
20 – 25	54	32
25 – 30	11	27
30 – 35	5	13

Which set of the bag has the highest average brusting pressure ? Which has more uniform pressure ?

Solution: For determining the set of bags having higher average bursting pressure, we compute mean and for finding out set of bags having more uniform pressure we compute coefficient of variation.

Manufacturer A :

Computation of mean and standard deviation

Bursting pressure	Mid-values x_i	f_i	$u_i = \frac{x_i - 17.5}{5}$	$f_i u_i$	$f_i u_i^2$
5 – 10	7.5	2	– 2	– 4	8
10 – 15	12.5	9	– 1	– 9	9
15 – 20	17.5	29	0	0	0
20 – 25	22.5	54	1	54	54
25 – 30	27.5	11	2	22	44
30 – 35	32.5	5	3	15	45
		$N = \sum f_i = 110$	$\sum u_i = 3$	$\sum f_i u_i = 78$	$\sum f_i u_i^2 = 160$

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

$\Rightarrow \quad \bar{X}_A = 17.5 + 5 \times \frac{78}{110}$ $\quad$ [∵ h = 5, a = 17.5]

$\Rightarrow \quad \bar{X}_A = 17.5 + 3.5 = 21$ $\quad \Rightarrow \quad \sigma_A^2 = h^2\left[\frac{1}{N}\sum f_i u_i^2 - \left(\frac{1}{N}\sum f_i u_i\right)^2\right]$

$\Rightarrow \quad \sigma_A^2 = 25\left[\frac{160}{110} - \left(\frac{78}{110}\right)^2\right]$ $\quad \Rightarrow \quad \sigma_A^2 = 25\left(\frac{17600 - 6084}{110 \times 110}\right) = 23.79$

$\Rightarrow \quad \sigma_A = \sqrt{23.79} = 4.87$

∴ Coefficient of variation = $\frac{\sigma_A}{\bar{X}_A} \times 100 = \frac{4.87}{21} \times 100 = 23.19$

Manufacurer B :

Bursting pressure	Mid-value x_i	f_i	$u_i = \frac{x_i - 17.5}{5}$	$f_i u_i$	$f_i u_i^2$
5 – 10	7.5	9	– 2	– 18	36
10 – 15	12.5	11	– 1	– 11	11
15 – 20	17.5	18	0	0	0
20 – 25	22.5	32	1	32	32
25 – 30	27.5	27	2	54	108
30 – 35	32.5	13	3	39	117
		$N = \sum f_i = 110$	$\sum u_i = 3$	$\sum f_i u_i = 96$	$\sum f_i u_i^2 = 304$

$\bar{X}_B = a + h\left(\frac{\sum f_i u_i}{N}\right) \Rightarrow \bar{X}_B = 17.5 + 5 \times \frac{96}{110} = 17.5 + 4.36 = 21.81$

$$\sigma_B^2 = h_2\left[\frac{1}{N}\left(\sum f_i u_i^2\right) - \left(\frac{1}{N}\sum f_i u_i\right)^2\right]$$

$\Rightarrow \quad \sigma_B^2 = 25\left[\frac{304}{110} - \left(\frac{96}{110}\right)^2\right] \Rightarrow \sigma_B^2 = 25\left(\frac{33440 - 9216}{110 \times 110}\right) = 50.04$

$\sigma_B = 7.07$

∴ Coefficient of variation = $\frac{\sigma_B}{\bar{X}_B} \times 100 = \frac{7.07}{21.81} \times 100 = 32.41$

We observe that the average brusting pressure is higher for manufacturer B. So, bags manufactured by B have higher bursting pressure.

The coefficient of variation is less for manufacturer A. So bags manufactured by A have more uniform pressure.