UNIT - III

Average :-

A number expressing the central or typical value in a set of data, in particular the mode, median or the mean which is calculated by dividing the sum of the values in the set by their number.

$$\bar{x} = \frac{\Sigma x}{n}$$

Σx = Sum of x,

n = number of values.

The Scope calculates a rolling average, using two different techniques. In other words, if you have the oscilloscope set ~~of~~ to average four acquistions, then it will average each new acquisition into the total average displayed on the screen, as it is acquired.

Problem 1 :- The weights of 12 Students are given in Kilograms. See the combained average value for this.

x = 53, 65, 70, 48, 55, 72, 65, 52, 63, 58, 61, 70

Solution :-

$$\text{Average}, \ \bar{x} = \frac{\Sigma x}{n}$$

$$\bar{x} = \frac{732}{12}$$

$$\boxed{\bar{x} = 61}$$

Find the composite mean in a Continuous series :-

$$\bar{x} = \frac{\Sigma fx}{n}, \quad n = \Sigma f, \quad f = \text{Frequency}$$

Problem 2 :- The marks of Students are given. See the combained average value for this.

Marks	20	30	35	45	50	55	60	70
Number of Students	5	3	6	8	12	7	5	4

Solution:-

Marks (x)	Number of Students	fx
20	5	100
30	3	90
35	6	210
45	8	360
50	12	600
55	7	385
60	5	300
70	4	280
Total	n = 50	$\Sigma fx = 2,325$

$$\bar{x} = \frac{\Sigma fx}{n}$$

$$= \frac{2325}{50}$$

$$\boxed{\bar{x} = 46.5}$$

Find the composite mean in the following volume :-

$$\bar{x} = \frac{\Sigma fm}{n}, \quad m = \text{middle (mind) number of the } x$$

Marks	20-30	30-40	40-50	50-60	60-70	70-80
Number of Students	5	8	12	15	6	4

Solution :-

Marks (x)	Number of Students (f)	Mint value (M)	fx
20-30	5	25	125
30-40	8	35	280
40-50	12	45	540

50-60	15	55	825
60-70	6	66	390
70-80	4	75	300
Total	$n = 50$		$\Sigma fm = 2460$

$$\bar{x} = \frac{\Sigma fm}{n}$$

$$= \frac{2460}{50}$$

$$\boxed{\bar{x} = 49.2}$$

Intermediate :-

Being or occuring at the middle place, stage, or degree or between extremes, of or relating to an intermediate school an intermediate curriculum.

Intermediate $= \left(\frac{n+1}{2}\right)$ th value of the element

Problem 1 :- Find the intermediate of the following data.

$x = 8, 12, 20, 15, 18$

Solution :-

Ascending Order $= 8, 12, 15, 18, 20$

Intermediate $= \left(\frac{n+1}{2}\right)^{th}$ element

$= \left(\frac{5+1}{2}\right)^{th}$ item

$= \left(\frac{6}{2}\right)^{th}$ item

$= 3^{th}$ item

Intermediate $= 15$

Find the Composite Intermediate in a Countinaous ~~series~~ valume :-

Problem 2 :-

Marks	20	30	35	45	50	55	60	70
Number of Students	5	3	6	8	12	7	5	4

Solution :-

x	f	fx	cf
20	5	100	5
30	3	90	8
35	6	210	14
45	8	360	22
50	12	600	34
55	7	385	41
60	5	300	46
70	4	280	50

$= \left(\frac{n+1}{2}\right)^{th}$ item

$= \left(\frac{50+1}{2}\right)^{th}$ item $= \left(\frac{51}{2}\right)^{th}$ item

$= 25.5^{th}$ item

Intermediate = 45 //

Mode :-

The mode is the most commonly observed value in a set of data. For the normal distribution, the mode is also the same value as the mean and median.

Problem 1 :-

Find average, Intermediate, ridge in the following valume :-

Marks	20-30	30-40	40-50	50-60	60-70	70-80
Number of Students	5	8	12	15	6	4

Solution :-

x	f	~~fm~~ m	fm	cf
20-30	5	25	125	5
30-40	8	35	220	13
40-50	12	45	540	25
50-60	15	55	825	40
60-70	6	65	390	46
70-80	4	75	300	50
	50	300	2460	

Average (mean) :-

$$\bar{x} = \frac{\Sigma fm}{n}$$

$$= \frac{2460}{50}$$

$$\bar{x} = 49.2$$

Intermediate (median) :-

i). $= n/2$

$= 50/2$

$= 25$

ii). $L + \frac{(n/2) - cf}{f} \times I$, $L = 40$, $I = 10$, $cf = 25$, $f = 12$

$$= 40 + \frac{50/2 - 25}{12} \times 10$$

$$= 40 + \frac{25 - 25}{12} \times 10$$

$$= 40 + 0 = 40$$

Mode :-

x	20-30	30-40	40-50	50-60	60-70	70-80
f	5	8	12	15	6	4

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

$$= 50 + \frac{15-12}{(2\times 15) - 12 - 6} \times 10$$

$$= 50 + \frac{3}{30-12-6} \times 10$$

$$= 50 + \frac{3}{12} \times 10$$

$$= 50 + 0.2 \times 10 \quad = 50 + 2.5$$

Mode $= 52.5$

Geometric Mean :-

The Geometric Mean (GM) is the average value or mean (GM) ~~is the~~ which signifies the central tendency of the set of numbers by finding the product of their values.

$$GM = \text{Anti log}\left(\frac{\Sigma \log x}{n}\right)$$

Find the Geometric mean for the following data.

$x = 125, 130, 75, 10, 45, 5, 0.5, 0.4, 500, 150$

x	log x
125	2.0969
130	2.1139
75	1.975
10	1.0
45	1.7
5	0.7
0.5	-0.3010
0.4	-0.3979
500	2.6989
150	2.1760
	13.6139

$$GM = \text{Anti log}\left(\frac{\Sigma \log x}{n}\right)$$

$$= \text{Anti log}\left(\frac{13.6139}{10}\right)$$

$$= \text{Anti log}\,(1.3613)$$

$$GM = 22.9773$$

Harmonic Mean :-

The harmonic mean is a type of numerical average. It is calculated by dividing the number of each number in the series. Thus, the harmonic mean is the reciprocal of the arithmetic mean of the reciprocals.

$$HM = \frac{n}{\Sigma(1/x)}$$

Problem 1 :- Find Harmonic mean :-

$x = 12, 10, 6, 8, 15, 5$

x	$(1/x)$
12	0.0833
10	0.1
6	0.1666
8	0.125
15	0.0666
5	0.2
	0.7415

$$HM = \frac{n}{\Sigma(1/x)}$$

$$= \frac{6}{0.7415}$$

$$= 8.0917$$

UNIT- IV

Measures of Dispersion :-

Range :-

The Range of set of data is the difference between the largest and smallest values.

$$Range(x) = Max(x) - Min(x) = L-S$$

$$= \frac{L-S}{L+S}$$

Problem 1 :-

Find the Range and of the following data.

$x = 100, 150, 50, 200, 250$

$$Range = L-S$$
$$= 250-50$$
$$= 200$$

$$\text{Coefficient of Range} = \frac{L-S}{L+S}$$
$$= \frac{250-50}{250+50}$$
$$= \frac{200}{300}$$
$$= 0.666$$

Problem 2 :-

Find the Rang and of the following frequency distribution.

x	50-60	60-70	70-80	80-90	90-100
f	5	15	20	25	5

x	f	m
50-60	5	55
60-70	15	65
70-80	20	75
80-90	25	85
90-100	5	95

$$\text{Range} = L - S$$
$$= 95 - 55$$
$$= 40$$

$$\text{Coefficient of Range} = \frac{L-S}{L+S}$$
$$= \frac{95-55}{95+55}$$
$$= \frac{40}{150}$$
$$= 0.2666$$

Quartile Deviation :-

When one takes half of the difference or variance between the 3rd and the 1st quartiles of a simple distribution or frequency distribution it is quartile deviation.

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

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

$$Q_3 = 3\left(\frac{n+1}{4}\right)$$

$$Q_1 = \left(\frac{n+1}{4}\right)$$

Problem 3 :-

Find the Quartile ~~Range~~ Deviation of the following data.

10, 15, 13, 18, 8, 16, 17, 14, 20, 6, 22

Quartile Deviation :- $Q_3 = 3\left(\frac{n+1}{4}\right)$

Ascending Order = 6, 8, 10, 13, 14, 15, 16, 17, 18, 20, 22

$$Q_3 = 3\left(\frac{11+1}{4}\right)$$
$$= 3\left(\frac{12}{4}\right)$$
$$= 3(3)$$
$$= 9\text{ th item}$$
$$\boxed{Q_3 = 18}$$

$$Q_1 = \left(\frac{n+1}{4}\right)$$
$$= \left(\frac{11+1}{4}\right) = \left(\frac{12}{4}\right)$$
$$= 3\text{ th item}$$
$$\boxed{Q_1 = 10}$$

$$\text{Quartile deviation} = \frac{Q_3 - Q_1}{2}$$
$$= \frac{18-10}{2}$$
$$= \frac{8}{2}$$
$$Q.D = 4$$

$$C.Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$
$$= \frac{18-10}{18+10}$$
$$= \frac{8}{28}$$
$$C.Q.D = 0.2857$$

Problem 4 :- Find Quartile and Coefficient quartile deviation.

C.I	0-10	10-20	20-30	30-40	40-50	50-60	60-70
f	1	6	15	20	14	6	2

x	f	cf
0-10	1	1
10-20	6	7
20-30	15	22
30-40	20	42
40-50	14	56
50-60	6	62
60-70	2	64

$$Q_1 = \left(\frac{n+1}{4}\right)$$

$$= \left(\frac{64+1}{4}\right) = \frac{65}{4}$$

$$= 16.25$$

$$Q_1 = 20$$

$$Q_3 = 3\left(\frac{n+1}{4}\right)$$

$$= 3\left(\frac{64+1}{4}\right)$$

$$= 3(16.25)$$

$$Q_3 = 48.75$$

$$Q_1 = l + \frac{n/4 - cf}{f} \times I = 20 + \frac{64/4 - 7}{15} \times 10$$

$$= 20 + \frac{16-7}{15} \times 10 = 20 + 9/15 \times 10$$

$$= 20 + 0.6 \times 10 = 20+6 = 26$$

$$Q_3 = l + \frac{3(n/4) - cf}{f} \times I = 20 + \frac{3(64/4) - 7}{15} \times 10$$

$$= 20 + \frac{3(16) - 7}{15} \times 10 = 20 + \frac{41}{15} \times 10$$

$$= 20 + 2.7333 \times 10 = 20 + \cancel{2.73}\ 27.33$$

$$Q_3 = 47.33$$

$$QD = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

$$= \frac{47.333 - 26}{47.333 + 26}$$

$$= \frac{21.333}{73.333} = 0.2909$$

$$\boxed{QD = 0.2909}$$

Standard Deviation :-

The standard deviation is a statistic that measures the dispersion of a data set relative to its mean and is calculated as the square root of the variance. The standard deviation is calculated as the square root of variance by determining each data point's deviation relative to the mean.

$$\sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$

Problem 1 :-

Find the standard deviation.

$x = 12, 10, 9, 6, 3$

Solution :-

$$\bar{x} = \frac{\Sigma x}{n} = \frac{40}{5} = 8$$

x	$x - \bar{x}$	$(x-\bar{x})^2$
12	4	16
10	2	4
9	1	1
6	-2	4
3	-5	25
		50

$$\sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}} = \sqrt{\frac{50}{5}} = \sqrt{10}$$

$$\sigma = 3.16$$

Problem 2 :-

Find the standard deviation.

x	1	2	3	4	5
f	3	7	10	3	2

x	f	fx	d = x - A	fd	$(fd)^2$
1	3	3	-2	-6	36
2	7	14	-1	-7	49
3	10	30	0	0	0
4	3	12	1	3	9
5	2	10	2	4	16
	25			-6	110

$$\sigma = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2}$$

$$= \sqrt{\frac{110}{25} - \left(\frac{-6}{25}\right)^2}$$

$$= \sqrt{4.4 - (-0.24)^2} = \sqrt{4.4 - 0.058}$$

$$= \sqrt{4.34}$$

$$\sigma = 2.083$$

Problem 3 :- Find the Standard deviation :-

x	0-10	10-20	20-30	30-40	40-50
f	3	4	7	6	5

x	f	m	d	fd	fd^2
0-10	3	5	-20	-60	3600
10-20	4	15	-10	-40	1600
20-30	7	25	0	0	0
30-40	6	35	10	60	3600
40-50	5	45	20	100	10000
	25			60	18,800

$$\sigma = \sqrt{\frac{\Sigma f d^2}{n} - \left(\frac{\Sigma f d}{n}\right)^2} \times I$$

$$= \sqrt{\frac{18800}{25} - \left(\frac{60}{25}\right)^2} \times 10$$

$$= \sqrt{752 - (2.4)^2} \times 10$$

$$= \sqrt{752 - 5.76} \times 10$$

$$= \sqrt{752 - 57.6}$$

$$= \sqrt{694.4}$$

$$\sigma = 26.35$$

Measures of Skewness :-

In probability theory and statistics, skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. The skewness value can be positive, zero, negative, or undefined.

Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same to the left and right of the center point.

Kurtosis is measure of whether ~~tha~~ the data are heavy-tailed or light-tailed relative to a normal distribution.

Karl Pearson coefficient of Skewness :-

Pearson's coefficient of skewness (second method) is calculated by multiplying the difference between the mean and median, multiplied by three. The result is divided by the standard deviation. You can use excel functions Average, Median and stdev.P to get a value for this measure.

$$S_{kp} = 3\left(\frac{\bar{x} - Me}{\sigma}\right)$$

Problem 1:-

Find the Karl Pearson's skewness coefficient :-

x	0-7	7-14	14-21	21-28	28-35	35-42	42-49	49-56
f	26	31	36	42	82	71	54	19

x	f	m	fm	cf
0-7	26	3.5	91	26
7-14	31	10.5	325.5	51
14-21	36	17.5	630	93
21-28	42	24.5	1029	135
28-35	82	31.5	2083	217
35-42	71	38.5	2733.5	288
42-49	54	45.5	2457	342
49-56	19	52.5	997.5	361
	361		10,846.5	1519

i). Mean :-

$$\bar{x} = \frac{\Sigma fm}{n}$$

$$= \frac{10,846.5}{361}$$

$$= 30.045$$

ii). Median :-

$$= \frac{L + n/2 - cf}{f} \times c$$

$$= \frac{28 + 361/2 - 217}{82} \times 7$$

$$= \frac{28 + 180.5 - 217}{82} \times 7$$

$$= \frac{28 + (-36.5)}{82} \times 7$$

$$= 28 + (-0.445) \times 7$$

$$= 28 - 3.115$$

$$= 24.88$$

iii). Standard deviation:-

x	f	m	d	fd	fd^2
0-7	26	3.5	-26.5	-689	-474.721
7-14	31	10.5	-19.5	-604.5	365420.25
14-21	36	17.5	-12.5	-450	202,500
21-28	42	24.5	-5.5	-231	53,361
28-35	82	31.5	1.5	123	15,129
35-42	71	38.5	8.5	603.5	364,212.25
42-49	54	45.5	15.5	839	700,569
49-56	19	52.5	22.5	427.5	182,756.25
				3965.5	2,358,668.75

$$\sigma = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2} \times c$$

$$= \sqrt{\frac{2,358,668.75}{361} - \left(\frac{3965.5}{361}\right)^2} \times 7$$

$$= \sqrt{6,533.70 - 10.984 \times 7}$$

$$= \sqrt{6,533.70 - 76.88}$$

$$= \sqrt{6,456.82} = 80.354$$

$$S_{kp} = 3\left(\frac{\bar{x} - Me}{\sigma}\right)$$

$$= 3\left(\frac{30.04 - 24.88}{80.354}\right)$$

$$= 3(0.0642)$$

$$= 0.1926$$

Bowley's Coefficient of Skewness :-

According to Business Statistics, Bowley recognized that the Bowley Skewness formula could not be used to compare different distributions with different units.

$$S_{kB} = \frac{Q_3 + Q_1 - 2Me}{Q_3 - Q_1}$$

Problem 1 :-

Find the Bowley's Co-efficient of skewness:-

x	below 200	200-400	400-600	600-800	800-1000	above 1000
f	25	40	80	75	20	16

x	f	cf
Below 200	25	25
200-400	40	65
400-600	80	145
600-800	75	220
800-1000	20	240
above-1000	16	256
	256	

i). $Q_1 = \left(\frac{n+1}{4}\right)$

$= \left(\frac{256+1}{4}\right) = \frac{257}{4}$

$= \cancel{25}\ 64.25$

ii). $Q_3 = 3\left(\frac{n+1}{4}\right)$

$= 3\left(\frac{256+1}{4}\right)$

$= 3(64.25)$

$= 192.75$

iii). $\text{Median} = L + \frac{n/2 - cf}{f} \times I$

$= 400 + \frac{128-65}{80} \times 200$

$= 400 + \frac{63}{80} \times 200$

$= \cancel{4}\ 400 + 0.7876 \times 200$

$= 400 + 157.5$

$= 557.5$

$\cancel{iv}$ $S_{KB} = \frac{Q_3 + Q_1 - 2Me}{Q_3 - Q_1}$

$= \frac{192.75 + 64.25 - 2(557.5)}{192.75 - 64.25}$

$= \frac{257 - 1{,}115}{128.5} = \frac{-858}{128.5}$

$= -6.6677$

Kurtosis Concept :-

Kurtosis is a statistical measure used to describe the degree to which scores cluster in the tails or the peak of a frequency distribution. The peak is the tallest part of the distribution, and the tails are the ends of the distribution.