Code: ~~18K3ECC01~~ 18K3ECA31

1. Arithmetic Mean:

Arithmetic average is also called as mean. It is the Common type and widely ~~use~~ used measure of central tendency.

1. Calculate mean from the following data (Raw data)

X: 40, 50, 55, 78, 80

Solution:

$n = 5$ $\quad \bar{X} = \frac{\Sigma X}{n} = \frac{40+50+55+78+80}{5} = \frac{303}{5} = 60.5$

[Note: $\bar{X}$ = Mean, ΣX = The sum of variables, N = Number of observations]

2. Calculate Mean from the following data (discrete data)

x	20	30	35	45	50	55	60	70
f	5	3	6	8	12	7	5	4

Solution:

x	f	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
	50	2325

$\bar{X} = \frac{\Sigma fx}{N (\text{or}) \Sigma f}$ $\quad n = 50, \Sigma fx = 2325$

$\bar{X} = \frac{2325}{50} = 46.5$

3. Calculate mean from the following data (Continuous data)

X(C.I)	20–30	30–40	40–50	50–60	60–70	70–80
f	5	8	12	15	6	4

Solution:

x = m (mid value)

Class Interval	f	m	fm
20–30	5	$\frac{20+30}{2} = 25$	125
30–40	8	35	280
40–50	12	45	540
50–60	15	55	825
60–70	6	65	390
70–80	4	75	300
	50		2460

$\bar{X} = \frac{\Sigma fm}{N (\text{or}) \Sigma f}$ $\quad N = 50, \Sigma fm = 2460$

$= \frac{2460}{50} = 49.2$

Merits of Arithmetic Mean:

1, It is easy to understand, 2, It is easy to calculate.
~~3, It is a used~~ 3, It is used in further calculation.
4, It is rigidly defined, 5, It is based on the value of every item in the series. etc...

Demerits:

1. The mean is unduly affected by the extreme items.
2. It is unrealistic, 3, It may lead to a false conclusion.

Median:

"The median, as its name indicates, is the value of the middle in a series, when items are arranged according to magnitude."

1. find out the median of the following items (Raw data)

X : 10, 15, 9, 25, 19

Solution:

n = No. of obs value

S.No	Ascending Order(x)	Descending order (x)
1	9	25
2	10	19
3	15	15
4	19	10
5	25	9

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

n=5

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

= Size of $(6/2)$th item

= 3^{rd} item

Median = 15

2. Find out the median of the following items.

X: 12, 30, 22, 36, 17, 40

Solution:

Ascending order : 12, 17, 22, 30, 36, 40

Median = size of $\left(\frac{n+1}{2}\right)$th item = size of $\left(\frac{6+1}{2}\right)^{th}$ item = $7/2$ = 3.5 = $\frac{3^{th}\text{ item} + 4^{th}\text{ item}}{2}$

$= \frac{22+30}{2} = 26$

3. Locate median from the following: (discrete series)

x	5	5.5	6	6.5	7	7.5	8
f	10	16	28	15	30	40	34

Solution

∵ C.f = Cumulative frequency

x	f	C.f
5	10	10
5.5	16	26
6	28	54
6.5	15	69
7	30	99
7.5	40	139
8	34	173
	173	

Median = size of $\left(\frac{N+1}{2}\right)$th item

= size of $\left(\frac{173+1}{2}\right)^{th}$ item

= size of $(87)^{th}$ item

~~Median = 7~~

Median =

4. Calculate the median from the following data (Continuous Series)

x	50-60	60-70	70-80	80-90	90-100	100-110	110-120
f	5	8	15	25	21	9	7

Solution:

x	f	C.f
50-60	5	5
60-70	8	13
70-80	15	28 → c.f
L ← 80-90	25 → f	53
90-100	21	74
100-110	9	83
110-120	7	90
	90	

$$\text{Median} = L + \frac{N/2 - c.f}{f} \times c$$

$\frac{N}{2} = \frac{90}{2} = 45$, $L = 80$, $C = 10$, $c.f = 28$, $f = 25$

$$= 80 + \frac{45-28}{25} \times 10 = 86.8$$

~~Mode~~

Merits:

1. It is easy to understand and easy to Compute.
2. It is quite rigidly defined. 3, It eliminates the effect of entreme items

Demerits:

1. It ignores the extreme items.
2. It is more Continuous Series, the median ~~is~~ is estimated, but not Calculated.

—— x ——

Mode:

Mode is the common item of a series. Mode is the value which occurs the greatest number of frequency in a Series. Mode is the most fashionable avalue of a distribution, because it is repeated the highest numbe of time in the Series.

—— x ——

1. Calculate mode from the following data.

850, 750, 600, 825, 850, 725, 600, 850, 640, 530.

Solution:

850, 750, 600, 825, 850, 725, 600, 850, 640, 530

850 repeats threee times,

∴ the mode is 850

2, 40, 44, 57, 78, 48

Solution

there is no mode value.

3, 45, 55, 50, 45, 40, 55, 45, 55 (Trimodal)

Solution

Mode i) 45, ii) 55

4 Calculate the mode from the following. (discrete series)

x	10	11	12	13	14	15	16	17	18
f	10	12	15	19	20	8	4	3	2

Solution: (1) Grouping Table.

x	f I	II	III	IV	V	VI
10	10	22		37		
11	12		27		46	
12	15	34				54
13	19		39	47		
14	20	28			32	
15	8		12			15
16	4	7		9		
17	3		5			
18	2					

(2) Analysis table.

	10	11	12	13	14	15	16	17	18
1									
2			I	I					
3				I	I				
4				I	I	I			
5		I	I	I					
6			I	I	I				
		1	3	5	3	1			

The mode is 13

5. Calculate the mode from the following! (Continuous series)

C.I	20-30	30-40	40-50	50-60	60-70	70-80	80-90
f	3	8	14	20	8	5	2

Solution:

C.I	f
20-30	3
30-40	8
40-50	14
50-60	20
60-70	8
70-80	5
80-90	2

formula: $\text{Mode} = L + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times c$

$L = 50$, $f_1 = 20$

$f_0 = 14$, $f_2 = 8$, $I = 10$

$\text{Model} = 50 + \frac{20-14}{2\times 20 - 14 - 8} \times 10$

$\text{Model} = 53.33$

L = lower limit of the model class.

f_1 = frequency of the model class

f_0 = frequency of the class preceding the model class

f_2 = frequency of the class succeeding the model class

C = class interval.

Geometric Mean:

Geometric mean is defined as the n^{th} root of the product of n items. The geometric mean larger than the arithmetic mean; on a ccasion it may turn out to be the same as the arithmetic mean, but usually it is smaller. If there are zeros or negative values in the series, the geometric mean cannot be used.

1. Calculate Geometric mean of the following.

50, 72, 54, 82, 93

Solution:

$n = 5$ $\quad GM = \sqrt[n]{x_1 x_2 \ldots x_n} = \sqrt[5]{50\times72\times54\times82\times93} = 68.26$

2. Calculate the Geometric mean of the following. (discrete series)

x	15	20	25	30	35	40	45	50
f	4	12	30	60	80	90	95	97

Solution:

x	f	log x	f log x
15	4	1.1761	4.7044
20	12	1.3010	15.612
25	30	1.3980	41.94
30	60	1.4771	88.626
35	80	1.5441	123.528
40	90	1.6021	144.189
45	95	1.6532	157.054
50	97	1.6990	164.803
	468		740.4564

$$GM = \text{Antilog} \frac{\Sigma f \log x}{N}$$

$$= \text{Antilog} \frac{740.4564}{468}$$

$$= \text{Antilog} (1.582171795)$$

$$= 38.2095$$

3. Calculate the geometric mean of the following (Continuous Series)

x	0-10	10-20	20-30	30-40	40-50
f	5	7	15	25	8

Solution:

x	f	m	log m	f log m
0-10	5	5	0.6991	3.4955
10-20	7	15	1.1761	8.2329
20-30	15	25	1.3979	20.9685
30-40	25	35	1.5441	38.6025
40-50	8	45	1.6532	13.2256
50-	60			84.5248

$$GM = \text{Antilog} \frac{\Sigma f \log m}{N}$$

$N = 60$, $\Sigma f \log m = 84.5248$

$$GM = \text{Antilog} \frac{84.5248}{60}$$

$$GM = 25.6299$$

Merits:

1. It is based on all observations, 2. It is rigidly defined. etc...
3. It is capable of further algebraic treatment.

Demerits:

1. It is difficult to understand.
2. Non-mathematical persons cannot do calculations.
3. It has restricted application.

Harmonic Mean:

Harmonic mean, like geometric mean is a measure of central tendency in solving special types of problems. Harmonic mean is the reciprocal of the arithmetic average of the reciprocal of values of various items in the variable. The reciprocal of a number is that value, which is obtained by dividing one by the value.

1. Calculate the harmonic mean from the following: (Raw data)
   X: 12, 10, 6, 8, 15, 5

Solution:

X	1/x
12	0.0833
10	0.1000
6	0.1666
8	0.1250
15	0.0666
5	0.2000
	0.7415

$$\text{Harmonic Mean} = \left[\frac{n}{\Sigma(1/x)}\right]$$

$n = 6$

$$= \frac{6}{0.7415} = 8.092$$

2. Calculate the harmonic mean from the following (Discrete Series)

x	6	7	8	9	10	11
f	4	6	9	5	2	8

Solution

x	f	1/x	f(1/x)
6	4	0.1667	0.668
7	6	0.1429	0.8568
8	9	0.125	1.1250
9	5	0.111	0.5555
10	2	0.1	0.2000
11	8	0.0909	0.7272
	34		4.1304

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

$$= \frac{34}{4.1304}$$

$= 8.2316$

Merits:

1. It is rigidly defined.
2. It is based on all the observation of the series. etc.

Demerits:

1. It is difficult to calculate and is not understandable.
2. It is not popular.

Dispersion UNIT-IV

Definition:- Dispersion is the measure of the variation of the items.

1, **Range:** The range is the simplest measure of dispersion. It is a rough measure of dispersion. It measure depends upon the extreme items and not ~~at~~ on all the items.

Problem:

1. Find the range of ~~wig~~ weights of 7 student from the following
27, 30, 35, 36, 38, 40, 43.

Solution:

27, 30, 35, 36, 38, 40, 43

L = Largest Value.
S = Smallest Value.

Range = L − S

= 43 − 27 = 16

Coefficient of range = $\frac{L-S}{L+S} = \frac{43-27}{43+27} = 0.23$

Merits:

1. It is simple to compute and understand
2. It gives a rough but quick answer.

Demerits:

1. It is not reliable, because it is affected by the extreme items.
2. It is not suitable for mathematical treatment.

Quartile Deviation:

By eliminating the lowest 25% and the highest 25% of items in series, we are left with the central 50% which are ordinarily free of extreme value. To obtain ~~the~~ a measure of variation, we use the distance between the first and the third quartiles, i.e., inner quartile range, but generally take half of this distance. Inter quartile range is computed by

Quartile Deviation:

Quartile deviation or semi-interquartile range Q is given by:

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

where Q_1 and Q_3 are the first and third quartiles of the distribution respectively.

1. Calculate the Quartile deviation from the following data
10, 15, 13, 18, 8, 16, 17, 14, 20, 6, 22,

Solution:

Ascending order.
6, 8, 10, 13, 14, 15, 16, 17, 18, 20, 22. $n = 11$

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

Q_3 = Value of $3\frac{(n+1)}{4}^{th}$ item = Value of $3\frac{(11+1)}{4}$th item = $\frac{36}{4} = 9^{th}$ item = 18

Quartile deviation = $\frac{Q_3 - Q_1}{2} = \frac{18-10}{2} = 4$

Coefficient Quartile deviation = $\frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{18-10}{18+10} = 0.2857$

2. Calculate the Quartile deviation from the following data.

x	20	30	40	50	60	70	80
f	3	61	132	152	140	51	3

Solution:

x	f	cf
20	3	3
30	61	64
40	132	196
50	152	349
60	140	489
70	51	540
80	3	543
	543	

Q_1 = Value of $\left(\frac{N+1}{4}\right)^{th}$ item = Value of $\left(\frac{543+1}{4}\right)^{th}$ item

= 136^{th} item = 40

Q_3 = Value of $3\frac{(N+1)}{4}$th item = Value of $3\frac{(543+1)}{4}$ "

= 408^{th} item = 60

$Q.D = \frac{Q_3 - Q_1}{2} = \frac{60-40}{2} = 10$

Coefficient of $Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{60-40}{60+40} = 0.2$

3. Calculate the Quartile deviation from the following data:

x	30-32	32-34	34-36	36-38	38-40	40-42	42-44
f	12	18	16	14	12	8	6

Solution:

C.I	f	cf
30-32	12	12
32-34	18	30
34-36	16	46
36-38	14	60
38-40	12	72
40-42	8	80
42-42	6	86

$\boxed{\frac{N}{4} = \frac{86}{4} = 21.5}$ $\boxed{3 \times \frac{86}{4} = 64.5}$

$Q_1 = L + \frac{N/4 - c.f}{f} \times c = 32 + \frac{21.5 - 12}{18} \times 2 = 33.06$

$Q_3 = L + \frac{3N/4 - cf}{f} \times c = 38 + \frac{64.5 - 60}{12} \times 2 = 38.75$

$Q.D = \frac{Q_3 - Q_1}{2} = \frac{38.75 - 33.06}{2} = 2.85$

Coefficient Q.D = $\frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{38.75 - 33.06}{38.75 + 33.06} = 0.08$

Merits:

1, It is simple to understand and easy to compute
2, It is not influenced by the extreme values.

Demerits:

1. It ignores the first 25% of the items and the last 25% of the terms.
2. It gives only a rough measure.

Mean Deviation:

Mean deviation is the arithmetic mean of deviations of a series computed from any measure of central tendency. ie, the mean, median or mode; all the deviations are taken as positive

1. Calculate the mean deviation for ~~taken from deviation~~ median from the following data

12, 10, 9, 6, 3

Solution

$n=5$

$$\text{Mean} = \frac{12+10+9+6+3}{5} = \frac{40}{5} = 8 \Rightarrow \bar{x} = \frac{\Sigma x}{n}$$

X	$\lvert D\rvert = X - \bar{X}$	$\lvert D\rvert = X - \text{Median}$ $X - 9$
12	4	3
10	2	1
9	1	0
6	2	3
3	5	6
40	14	13

$$\text{Mean deviation from mean} = \frac{\Sigma|D|}{n} = \frac{14}{5} = 2.6$$

$= 9$

$$\text{Median} = \frac{13}{5} = 2.6$$

$$\text{Coefficieny M.D} = \frac{\text{M.D}}{\bar{X}} = \frac{14}{8} = 1.75$$

$$\text{Coefficient of M.D} = \frac{\text{M.D}}{\text{Median}} = \frac{2.6}{9} = 0.2888$$

2. Calculate mean deviation from median for the following data:

X: 400, 420, 440, 460, 480

Solu

$$\text{Median} = \text{Size of } \left(\frac{n+1}{2}\right)^{th} \text{ item} = \text{Size of } \left(\frac{5+1}{2}\right)^{th} \text{ item} = 3^{th} \text{ items}$$

$= 440$

$(X - \cancel{M}$ $|(X - \text{Median})| = |(X - 440)| = |d|$

$$|d| = 40 + 20 + 0 + 20 + 40 = \frac{\Sigma|d|}{N} = \frac{120}{5} = 24$$

3. Calculate the median from ~~med~~ mean deviation for the following data.

X	5	6	7	8	9	10
f	2	4	6	3	2	3

x	f	cf	|d|	f|d|
5	2	2	2	4
6	4	6	1	4
7	6	12	0	0
8	3	15	1	3
9	2	17	2	4
10	3	20	3	9
	20			24

Median = size of $\left(\frac{N+1}{2}\right)^{th}$ item

$= \text{''} \left(\frac{20+1}{2}\right)^{th}$ item

$= 10.5^{th}$ item $= 7$

$M.D = \frac{\Sigma f|d|}{N} = \frac{24}{20} = 1.2$

(4). Calculate mean deviation from the following data.

C.I	2-4	4-6	6-8	8-10
f	3	4	2	1

Solution:

$d = x - \bar{x}$

C.I	f	m	fm	|d|	f|d|
2-4	3	3	9	2.2	6.6
4-6	4	5	20	0.2	0.8
6-8	2	7	14	1.8	3.6
8-10	1	9	9	3.8	3.8
	10		52		14.8

$\bar{x} = \frac{\Sigma fm}{N} = \frac{52}{10} = 5.2$

Mean deviation $= \frac{\Sigma f|d|}{N} = \frac{14.8}{10} = 1.48.$

Coefficien of M.D $= \frac{M.D}{Mean} = \frac{1.48}{5.2} =$

$= 0.2846$

Merits:

1, It is simple to understand and easy to compute.
2. M.D is a Calculated Value. (3), It is a better measure for Comparison.

De Merits

1. It is a non-algebraic treatment.
2. It is not a very accurate measure of dispersion.

Standrad Deviation:

Karl Pearson introduced the concept of standard deviation in 1893. It is most important measure of dispersion and is widely used in many statistical formula, Standard deviation is also called Root mean square Deviation or Mean Error or Mean Square Error. The reason is that it is the square root of the means of the Square dieviation from the arithmetic mean. It provids accurate result. The standard is denoted by Greek Letter (σ) Sigma.

1. Calculate the standard deviation from the following data. (Raw data) (11)

12, 10, 9, 6, 3

Solution:

$$\text{Mean}(\bar{x}) = \frac{\Sigma x}{n} = \frac{12+10+9+6+3}{5} = 40/5 = 8$$

$\boxed{\bar{x} = 8}$ $n=5$

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

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

$$= \sqrt{\frac{50}{5}} \Rightarrow 3.16$$

2. Calculate the standard deviation from the following data.

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

Solution:

x	f	$d=(x-A)$, $A=3$	fd	fd^2
1	3	-2	-6	12
2	7	-1	-7	7
3	10	0	0	0
4	3	1	3	3
5	2	2	4	8
	25		-6	30

$$\sigma = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2} = \sqrt{\frac{30}{25} - \left(\frac{-6}{25}\right)^2}$$

$$\sigma = 2.083$$

3. Calculate standard deviation from the following data.

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

Solution:

$$d = \frac{m-A}{c} = \frac{m-25}{10}$$

C.I	f	m	~~d~~	d	fd	fd^2
0-10	3	5	~~-20~~	-2	-6	12
10-20	4	15	~~-10~~	-1	-4	4
20-30	7	25	~~0~~	0	0	0
30-40	6	35	~~10~~	1	6	6
40-50	5	45	~~20~~	2	10	20
	25				6	42

$N = 25$

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

$$= \sqrt{\frac{42}{25} - \left(\frac{6}{25}\right)^2} \times 10$$

$$\sigma = 12.74$$

Merits :

1. It is rigidly defined and its value is always define and based on all the observations and the actual signs of deviations are used.
2. It is possible for further algebraic treatment. etc.

Demerits

1. It is not easy to understand, and it is difficult to calculate.
2. It gives more weight to extreme values because the values are squared up. etc.

Unit-5

Skewness:

"Measures of Skewness tell us the direction and the extent of skewness. In symmetrical distribution the mean, median and mode are identical. The more the mean moves away from the mode, the larger the asymmetry or skewness."

—Simpon and kafka.

The measures of central tendency and dispersion are inadequate to characterise a distribution completely: they may be supported and supplemented by two more measures, viz Skewness and kurtosis. Thus the following four measures are sufficient to describe a frequency distribution completely!

Moment	What it measures
1. first moment about the origin	Mean
2. Second moment about the mean	Variance.
3. Third moment about the mean	Skewness.
4. fourth moment about the mean	Kurtosis.

Skewness	Symmetry (No skewness)	Asymmetry	
		Positive	negative
Average	$\bar{X} = M = Z$ $40 = 40 = 40$	$\bar{X} > M > Z$ $32.1 > 30 > 10$	$\bar{X} < M < Z$ $30 < 50 < 70$
Quartile	$(Q_3 - M) = (M - Q_1)$	$(Q_3 - M) > (M - Q_1)$	$(Q_3 - M) < (M - Q_1)$
Curve	Normal	Skewed to the right	Skewed to the left.

1, Calculate Karl pearson's Coefficient of Skewness for the following data.

x	58	59	60	61	62	63	64	65
f	10	18	30	42	35	28	16	8

Solution

x	f	fx	$d=\frac{x-A}{c}$	d^2	fd	fd^2
58	10	580	0	0	0	0
59	18	1062	1	1	18	18
60	30	1800	2	4	60	120
61	42	2562	3	9	126	378
62	35	2170	4	16	140	560
63	28	1764	5	25	140	700
64	16	1024	6	36	96	576
65	8	520	7	49	56	329
	187	11482			636	2744

$$\bar{X} = \frac{\Sigma fx}{N} = \frac{11482}{187} = 61.40$$

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} = \sqrt{\frac{2744}{187} - \left(\frac{636}{187}\right)^2} = 1.7635$$

$$\text{Coefficient of Skewness} = \frac{\bar{x} - \text{mode}}{\sigma} = \frac{61.40 - 61}{1.7635} = 0.2268$$

2. Calculate Karl pearson's coefficient of Skewness from the following data:

C.I	300-400	400-500	500-600	600-700	700-800	800-900	900-1000	1000-1100	1100-1200
f	14	46	58	76	68	62	48	22	6

Solution !

$A = 350$

C.I	f	c.f	m	$d = \frac{m-A}{c}$	fd	fd^2	fd^2
300-400	14	14	350	0	0	0	0
400-500	46	60	450	1	46	1	46
500-600	58	118	550	2	116	4	232
600-700	76	194	650	3	228	9	682
700-800	68	202	750	4	272	16	1088
800-900	62	324	850	5	310	25	1550
900-1000	48	372	950	6	288	36	1728
1000-1100	22	391	1050	7	154	49	1078
1100-1200	6	400	1150	8	48	64	384
	400				1462		6790

$\bar{x} = A + \frac{\Sigma fd}{N} \times C = \Rightarrow 350 + \frac{1462}{400} \times 100 = 715.5$

$\frac{N}{2} = \frac{400}{2} = 200$

$\text{Median} = L + \frac{(N/2 - c.f)}{f} \times C = 700 + \frac{(200-194)}{68} \times 100$

$= 708.82$

$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times C = \sqrt{\frac{9790}{400} - \left(\frac{1462}{400}\right)^2} \times 100 = 190.2$

$\text{Coefficient of Skewness} = \frac{3(\bar{x} - \text{mode})}{\sigma} = \frac{3(715.5 - 708.82)}{190.2}$

$= 0.1054$

x

2. Bowley's Coefficient of Skewness !

Prof. Bowley's has suggested a formula based on relative ~~posti~~ position of quartiles. In a symmetrical distribution, the quartiles are equidistant from the value of the mean: i.e, Median $- Q_1 = Q_3 -$ Median. This means the value of the median is the mean of Q_1 and Q_2. But in a skewed distribution, the quartiles will not be equidistant from the median.

C.I	f	cf
4-8	6	6
8-12	10	16
12-16	18	34
16-20	30	64
20-24	15	79
24-28	12	91
28-32	10	101
32-36	6	107
36-40	2	109

Q_1 = size of $\left(\frac{N}{4}\right)^{th}$ item

= size of $\left(\frac{109}{4}\right)^{th}$ item.

= 27.25 (12-16)

$Q_1 = 12 + \frac{(27.25-16)}{18} \times 4$

= 14.5

Q_2 = size of $(N/2)^{th}$ item

= size of $\left(\frac{109}{2}\right)^{th}$ item

= 54.5^{th} item.

= (16 - 20)

$\text{Median} = L + \frac{(N/2 - cf)}{f} \times C = 16 + \frac{(54.5-34)}{30} \times 4$

= 18.73

Q_3 = size of $3(N/4)^{th}$ item = size of $\frac{3(109)}{4}^{th}$ item

= 31.75^{th} = (24-28)

$Q_3 = L + \frac{3N/4 - c.f}{f} \times C = 24 + \frac{81.75-79}{12} \times 4$

= 24.91

Coefficient of Skewness = $\frac{Q_3 + Q_1 - 2\,\text{Median}}{Q_3 - Q_1}$

$= \frac{24.91 + 14.5 - 2(18.73)}{24.91 - 14.5}$

= ~~0.18~~

= 0.1873

1. Calculate bowely's coefficient of skewness for the following data:

x	20	30	40	50	60	70	80
f	3	61	132	153	140	51	3

Solution:

x	f	cf
20	3	3
30	61	64
40	132	196
50	153	349
60	140	489
70	51	540
80	3	543
	543	

$Q_1 = \text{Size of } \left(\frac{N+1}{4}\right)^{th} \text{ item.}$

$= \text{Size of } \left(\frac{543+1}{4}\right)^{th} \text{ item.}$

$= 136^{th} \text{ item} = 40$

$Q_2 = \text{Size of } \left(\frac{N+1}{2}\right)^{th} \text{ item.}$

$= \text{Size of } \left(\frac{543+1}{2}\right)$ "

$= 272^{th} \text{ item}$

$= 50 = \text{Median}$

$Q_3 = \text{Size of } 3\left(\frac{N+1}{4}\right)^{th} \text{ item} = \text{Size of } 3\left(\frac{543+1}{4}\right)^{th} \text{ item}$

$= 408^{th} \text{ item} = 60$

$$\text{Coefficient of Skewness} = \frac{Q_3 + Q_1 - 2\,\text{median}}{Q_3 - Q_1}$$

$$= \frac{60 + 40 - 2 \times 50}{60 - 40}$$

$$= 0$$

2. Calculate Bowley's coefficient of Skewness of for the following data.

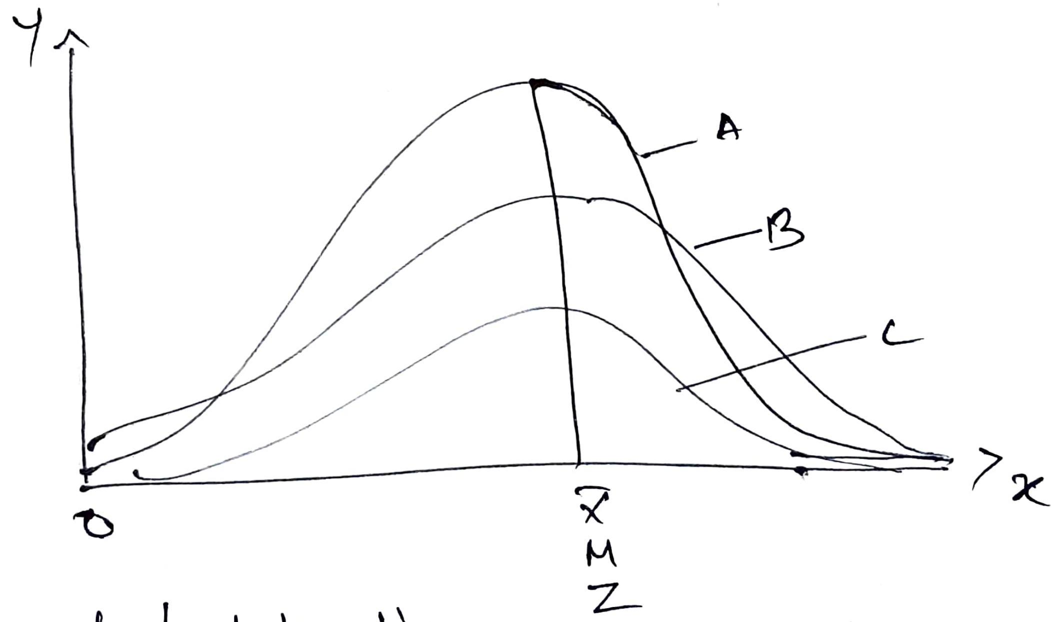
C.I	4-8	8-12	12-16	16-20	20-24	24-28	28-32	32-36	36-40
f	6	10	18	30	15	12	10	6	2

Kurtosis

"The degree of kurtosis of a distribution is measured relative to the peakedness of a normal"

— Simpson and Kafka.

The expression "kurtosis" is used to describe the peakedness of Curve. The three measures - central tendency, dispersion and skewness, describe the characteristics of frequency distribution. But these studies will not give us a clear picture of the characteristics of a distribution.



A → Peaked Leptokurtic , B → Normal Mesokurtic

C → (Flat topped) platykurtic.