B.Sc., MATHEMATICS/CACS32

STATISTICAL METHODS AND THEIR APPLICATIONS I

E CONTENT

UNIT-2

METHODS OF CENTRAL TENDENCY OR AVERAGES

The various methods to find averages:

1.Arithmetic mean

2.Median

3.Mode

4.Geometric Mean

5.Harmonic Mean

ARITHMETIC MEAN:

Arithmetic mean is the most used measures of averages. It is defined as the sum of the values of all individual observations of a series divided by the number of observation of a series.

FORMULA:

X1,X2…..Xn are n observation of a series when the arithmetic mean denoted by $\bar{x}$

i. $\bar{x}=\frac{\Sigma x}{n}$ for individual observations.

ii. $\bar{x}=\frac{\Sigma fx}{\Sigma f}$ (or) $\bar{x}=A+\frac{\Sigma fd}{\Sigma f}$ for frequency distribution.

iii. Step deviation method or continuous distribution method

$\bar{x}=A+\frac{\Sigma fd}{\Sigma f}\times i$

PROBLEMS

1.Find the arithmetic mean of the following data 12,50,10,9,11,14,6.

Solution:

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

$$=\frac{12+50+10+9+11+14+6}{7}$$

$$=16$$

2.The following table gives the marks obtained by 10 students in a class. calculate the arithmetic mean.

rollno	1	2	3	4	5	6	7	8	9	10
marks	40	50	30	60	70	80	40	50	60	90

Solution:

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

$$=\frac{40+50+30+60+70+80+40+50+60+90}{10}$$

$$=57$$

3.From the following table find the mean height.

Height	60	61	62	63	64
No of children	2	3	5	8	7

Solution:

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

X	F	fx
60	2	120
61	3	183
62	5	310
63	8	504
64	7	448
	25	1565

$\overline{x}=\frac{\Sigma fx}{\Sigma f}$

$=\frac{1565}{25}$

=62.5

4.The following is the age distribution of 100 persons in a street. Calculate the arithmetic mean.

Age group	0-10	10-20	20-30	30-40	40-50	50-60
No of person	5	10	25	30	20	10

Solution:

$$\bar{x}=A+\frac{\Sigma fd}{\Sigma f}\times i$$

X	F	Mid x	$d=\frac{x-a}{i}$	Fd
0-10	5	5	$\frac{5-25}{10}$=-2	-10
10-20	10	15	-1	-10
20-30	25	25	0	0
30-40	30	35	1	30
40-50	20	45	2	40
50-60	10	55	3	30
	100			80

$$\bar{x}=A+\frac{\Sigma fd}{\Sigma f}\times i$$

$$\bar{x}=25+\frac{80}{100}\times 10$$

$$=25+8$$

$$\bar{x}=33$$

5.Find the missing frequency for the following distribution if the mean is 12.9.

Class interval	0-5	10-15	15-20	15-20	20-25
frequency	3	F	8	5	4

Solution:

X	F	Mid x	$d=\frac{x-a}{i}$	fd
0-5	3	2.5	-2	-6
5-10	F	7.5	-1	-f
10-15	8	12.5	0	0
15-20	5	17.5	1	5
20-25	4	22.5	2	8
	20+f			7-f

$$\bar{x}=A+\frac{\Sigma fd}{\Sigma f}\times i$$

$$12.9=12.5+\frac{7-f}{20+f}\times 5$$

$$\frac{12.9-12.5}{5}=\frac{7-f}{20+f}$$

$$\frac{0.4}{5}=\frac{7-f}{20+f}$$

$$0.4(20+f)=5(7-f)$$

$$8+0.4f=35-5f$$

$$0.4f+5f=35-8$$

$$5.4f=27$$

$$F=\frac{27}{5.4}$$

$$F=5$$

COMBINED ARITHMETIC MEAN:

The arithmetic mean of two or more groups with there number of items then we can compute the mean of the combined groups.

Combined mean of two groups is given by

$$\bar{x}=\frac{n1\bar{x}1+n2\bar{x}2}{n1+n2}$$

problem:

1.The mean height of 25 male workers in a factory is 61 cm and the meam height of 35 female workers in the same factory is 58cm.find the combined mean height of 60 workers in the factory.

$$\bar{x}=\frac{n1\bar{x}1+n2\bar{x}2}{n1+n2}$$

$$\bar{x}=\frac{25\times61+35\times58}{25+35}$$

$$\bar{x}=\frac{3555}{60}$$

$$\bar{x}=59.25$$

1.The mean marks of 100 students where found to be 40.later on it was discovered that a score of 53 was misread as 83.Find the correct mean corresponding to the correct score.

Solution:

n=100

40

correct value=53

wrong value=83

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

$$40=\frac{\Sigma x}{100}$$

$$4000=\Sigma x$$

Correct Σx=wrong Σx-wrong value+ correct value

=4000-83+53

=3970

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

$=\frac{3970}{100}$

Correct $\bar{x}$ =39.7

Merits and demerits of arithmetic mean:

Merits:

- It is easy to understand, it is easy to calculate
- It is based upon all the observation
- It is rigidly defined
- It is capable of algebraic treatment that it can be used to calculate the combine mean
- It is link affected by fluctuations

Demerits:

- It is affected very much by extreme values
- It cannot be accurately determined by even if one of the values is not known
- It cannot be calculated for distribution with open end class
- It cannot be located graphically

MEDIAN:

Median is the value which divides the distribution into two halves. Thus the median is the mid value of the distribution. Mefian does not depend on the values of all the items and it depends on the position of the values and hence it is called a position average.

Formula:

Individual and discrete series

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

Continuous series

Median=size of $(\frac{N}{2})$th item

The extract value of the median we use the formula

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

Problems:

1.Find the median marks of a students 70,60,75,90,65,80,42,65,75.

Ascending order

42,60,75,90,65,80,42,65,75

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

= size of$(\frac{(9+1}{2})$th item

= size of$(\frac{10}{2})$th item

=size of 5^{th} item

Median= 70

2.calculate the median of the following distribution.

X	10	15	8	20	18
F	24	6	30	16	26

Solution:

X	F	Cf
8	30	30
10	24	54
15	6	60
18	26	76
20	16	102

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

= size of($\frac{102+1}{2}$)th item

= size of($\frac{103}{2}$)th item

= size of(51.5)th item

median =10

Class interval	120-150	150-180	180-210	210-240	240-270	270-300	300-330	330-360
frequency	25	65	135	430	320	175	79	21

3. calculate the median for the following data.

Solution:

CI	F	Cf
120-150	25	25
150-180	65	90
180-210	135	225
210-240	430	655
240-270	320	975
270-300	175	1150
300-330	79	1229
330-360	21	1250

Median=size of $(\frac{N}{2})$th item

=size of $(\frac{1250}{2})$th item

=Size of 625^{th} item

=210-240

L=210 n=1250

Cf=225 i=30

F=430

Median=L+$\frac{\frac{N}{2}-c.f}{f}$×i

=210+$\frac{\frac{1250}{2}-225}{430}$×30

=210+27.906

Median=227.906

4.calculate the median for the following data.

Saving(Rs) less than	10	20	30	40	50	60	70	80
Cumulative frequency	15	35	64	84	96	120	192	256

Solution:

Since the less than value are given we have to find that true class limits and the corresponding frequency

0-10=15

10-20=35-15=20

20-30=64-35=29

30-40=84-64=20

40-50=96-84=12

50-60=120-96=24

60-70=192-120=72

70-80=256-196=64

CI	F	Cf
0-10	15	15
10-20	20	35
20-30	29	64
30-40	20	84
40-50	12	96
50-60	24	120
60-70	72	192
70-80	64	256

Median=size of $(\frac{N}{2})$th item

= size of $(\frac{256}{2})$th item

=128th item

L=60 cf=120

F=72 n=256

I=10

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

$$\text{Median}=60+\frac{128-120}{72}\times 10$$

=60+1.111

Median=61.111.

Merits and Demerits:

Merits:

- It is easy to understand and easy to calculate in some cases it can be located by inspection
- It is rigidly defined
- It is not affected by extreme values
- It can be calculated for distribution with open end classes

Demerits:

- It is not based on all observations
- It is not capable of algebraic treatment
- It is affected more by sampling fluctuation as compared to the value of mean
- It is necessary to arrange the data to calculate the median

MODE

Mode is that value which occurs must often in the data that is with the highest frequency.

Formula:

Individual series = Repeated maximum number of items.

Discrete series = grouping table, analysis table

Analysis table formula mode = L+$\frac{f1-f0}{2f1-f0-f2}$ ×i

Empirical formula:

The empirical formula gives the relationship between the mean, median and mode

Mode = 3median-2mean

PROBLEM:

1. Find the mode for the sets of numbers 2,2,3,5,6,8,5,9,5

Solution:

5 appear maximum number of times

Mode=5

2. calculate the mode for the following data

X	3	5	7	9	11	13	15	17
F	2	5	7	8	15	7	5	1

Solution:

The value corresponding to the maximum frequency 15 is 11

Mode value=11

3.calculate the mode for the following distribution

X	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
F	3	6	10	20	15	5	4	2

The mode class of class corresponding highest frequency=20-25

L=20

F0=10

F1=20

F2=15

I=5

$$\text{mode}=L+\frac{f1-f0}{2f1-f0-f2}\times i$$

$$=20+\frac{20-10}{2(20)-10-15}\times 5$$

=20+0.6667×5

=20+3.3335

=23.3335

4.From the following data find out mode using empirical formula.

Class interval	3-4	4-5	5-6	6-7	7-8	8-9	9-10
frequency	83	27	25	50	75	38	18

Solution:

X	F	Mid x	D=x-a/i	fd	cf
3-4	83	3.5	-3	-249	83
4-5	27	4.5	-2	-54	110
5-6	25	5.5	-1	-25	135
6-7	50	6.5	0	0	185
7-8	75	7.5	1	75	260
8-9	38	8.5	2	76	298
9-10	18	9.5	3	54	316

Mean=A+$\frac{\Sigma fd}{\Sigma f}$×i

=6.5+$\frac{-123}{316}$×1

=6.5-0.3892

mean=6.1108

median=size of ($\frac{N}{2}$)th item

= size of ($\frac{316}{2}$)th item

=158th item

=6-7

Median=L+$\frac{\frac{N}{2}-c.f}{f}$×i

=6+$\frac{150-135}{50}$×1

=6+0.40

=6.46

Mode=3median-2mean

=3×6.46-2×6.110 =19.38-12.2216

Mode=7.1584

Merits and Demerits:

Merits:

- It is easy to calculate and in some cases it can be located by inspection
- It is not affected by extreme values
- It can be located the arithmetic with open end classes
- It can be determined graphically

Demerits:

- The value of mode cannot always be determined in some cases we may have bimodal or multi model series
- It is not capable of further algebraic treatment
- The value of the mode is not based on the each and every items of the series
- It is affected to greater extend by sampling fluexuation as compared to the value of mean

GEOMETRIC MEAN:

The geometric mean is defined as the nth root of the product of n items of the series.

Geometric mean= $\sqrt[n]{x1, x2, \ldots\ldots xn}$ Type equation here.

For individual observations

G.M=antilog of $[\frac{\Sigma logx}{n}]$

For discrete series

G.M=antilog of $[\frac{\Sigma flogx}{\Sigma f}]$

For continuous series

G.M=antilog of $[\frac{\Sigma flogm}{\Sigma f}]$

Problems:

1.Find the geometric mean of the following quantities 2,18,32,36,6.

(2×2×3×3×2×2×2×2×2×2×2×3×3×2×3)^1/5

= (2^10×3^5) ^1/5

= (4^5×3^5) ^1/5

= (4×3) ^5/5

=12

2.Find the geometric mean of the following data 82,93,50,54,72. Solution:

X	Log x
82	1.9138
93	1.9684
50	1.6989
54	1.7323
72	1.8573

G.M=antilog of $[\frac{\Sigma log x}{n}]$

=antilog of $[\frac{9.1707}{5}]$

=antilog of [1.8341]

=68.25.

3.compute the G.M from the following data given below.

Category	I.	11.	111	1V	V	V1	V11	V111
Monthly income	5000	3750	3000	750	600	400	300	200
No of employees	2	4	6	8	6	100	10	50

Case	Monthly income	No of employees	Log x	Flog x
I	5000	2	3.6990	7.398
II	3750	4	3.5740	14.396
III	3000	6	3.4771	20.8626
IV	750	8	2.8750	23.0000
V	600	6	2.7781	16.6686
VI	400	100	2.6020	260.2
VII	300	10	2.4771	24.471
VIII	200	50	2.3010	115.1
		186		481.9962

G.M=antilog of $[\frac{\Sigma flogx}{\Sigma f}]$

$=$antilog of $[\frac{481.9962}{186}]$

=Antilog of [2.5913]

=390.211

4.compute the geometric mean for the following data.

Class	0-10	10-20	20-30	30-40	40-50
Frequency	5	7	15	25	8

Solution:

Class x	Frequency f	Mid value	Log m	Flog m
0-10	5	5	0.6989	3.4945
10-20	7	15	1.1760	8.232
20-30	15	25	1.3979	20.968
30-40	25	35	1.5440	38.6
40-50	8	45	1.6532	13.225
	60			84.519

Solution

G.M=antilog of $[\frac{\Sigma f logm}{\Sigma f}]$

=antilog of $[\frac{84.519}{60}]$

=antilog of [1.4086]

=25.621.

Merits and Demerits:

Merits:

- It is rigidly defined
- It is based upon all the observations
- It is suitable for further mathematical treatment
- It gives comparatively more weight to small items

Demerits:

- It is not easy to understand because of its abstract mathematical character
- It cannot be determined if one of the observations zero or negative

HARMONIC MEAN:

The harmonic mean is defined as the reciprocal of the arithmetic mean of the reciprocals of a items of a series.

1.If x1,x2,….xn are n items

H.M=$\frac{n}{\Sigma\frac{1}{x}}$

2.when frequencies are given

H.M=$\frac{\Sigma f}{\Sigma f\frac{1}{x}}$

3.continous series

H.M=$\frac{\Sigma f}{\Sigma f\frac{1}{m}}$

1.Find the harmonic mean for the following individual data 6,15,35,40,900,520,300,400,400,1800,2000.

X	$\frac{1}{x}$
6	0.1667
15	0.0667
35	0.0285
40	0.025
900	0.0011
520	0.0019
300	0.0033
400	0.0025
400	0.0025
1800	0.0006
2000	0.0005

0.2993

H.M=$\frac{n}{\Sigma\frac{1}{x}}$ = 36.7524

2.calculate the harmonic mean for the following data.

X	10	12	14	16	18	20
F	5	18	20	10	6	1

Solution:

X	F	$\frac{1}{x}$	$f\frac{1}{x}$
10	5	0.1	0.5
12	18	0.083	1.494
14	20	0.071	1.42
16	10	0.0625	0.625
18	6	0.0555	0.333
20	1	0.05	0.05
	60		4.422

$$H.M=\frac{\Sigma f}{\Sigma f\frac{1}{x}}$$

$$=\frac{60}{4.422}$$

$$=13.568$$

3.calculate the harmonic mean of the following data

Marks	15-25	25-35	35-45	45-55	55-65	65-75
No of students	4	11	19	14	0	2

Solution:

X	F	Mid m	$\frac{1}{m}$	$f\frac{1}{m}$
15-25	4	20	0.05	0.2
25-35	11	30	0.0333	0.3663
35-45	19	40	0.025	0.475
45-55	14	50	0.02	0.28
55-65	0	60	0.0166	0
65-75	2	70	0.0142	0.0284
	50			1.3497

H.M=$\frac{\Sigma f}{\Sigma f \frac{1}{m}}$

$=\frac{50}{1.3497}$

=37.0452.

Merits and Demerits:

Merits:

- It is rigidly defined
- It is based upon all the observations
- It is suitable for further mathematical treatment
- It gives comparatively more weight to small items

Demerits:

- It is not easy to understand because of its abstract mathematical character
- It cannot be determined if one of the observations zero or negative

UNIT -3

Methods of studying variable.

1. Range
2. Quartile deviation
3. Mean deviation
4. Standard deviation

Range:

Range is defined as the differences between the largest and smallest value of the distribution

Range=largest value-smallest value

Coeffient of Range:

L-S\L+S

PROBLEMS:

Individual series:

1. The profits earned by 10 public under taking or given below. 27,32,16,15,10,30,15,29,19,35.calculate the range and the coeffient of range.
   Solution:
   Range =L-S
   = 35-10
   =25
   Coefficient of range:
   = L-S/L+S
   = 35-10/35+10
   = 25/45
   =5/9

   Discreate distribution:

2. Calculate the range and its coefficient from the following.

X	4	6	8	10	12
F	15	25	12	36	30

Solution:

Range=L-S

=12-4

=8

Coefficient of range :

=L-S/L+S

=12-4/12+4

=1/2

3. Calculate the range from the following values.

Marks	10-20	20-30	30-40	40-50
No. of Students	5	8	10	7

Solution:

Range=L-S

=50-10

=40

Coefficient of range:

=L-S/L+S

=50-10/50+10

=2/3

Quartile deviation:

Quartile deviation is an absolute measure of dispersion and it is based upon upper quartile (Q3) and lower quartile (Q1) it represents the average difference between the two quartiles and is given by

Quartile deviation=$\frac{Q3-Q1}{Q3+Q1}$

Discrete series:

Q3=size of 3$(\frac{N+1}{4})$th item

Q1=size of $(\frac{N+1}{4})$th item

Continuous series:

Q1=size of $(\frac{N}{4})$th item

Q1= L+ $\frac{\frac{N}{4}-c.f}{f}\times$i

Q3=size of 3$\left(\frac{N}{4}\right)$

Q3= L+ $\frac{\frac{3N}{4}-c.f}{f}\times$i

Coefficient of quartile deviation:

$$\frac{Q3-Q1}{Q3+Q1}$$

PROBLEMS:

1. From the following data calculate quartile deviation and its coefficient.
   1490,692,777,335,582,488,753,384,407,
   672,522.
   Solution:
   355,384,407,488,522,582,672,692,753,777,
   1490.
   Q1=size of $(\frac{N+1}{4})th\ item$
   =size of $(\frac{11+1}{4})$
   =size of $(\frac{12}{4})$
   =3rd item
   Q1=407

Q3=size of $3(\frac{N+1}{4})$th item

=size of $3\left(\frac{11+1}{4}\right)$
=3(3)
=9th item

Q3=753

Q.D=$\frac{Q3-Q1}{2}$

=$\frac{753-407}{2}$

=$\frac{346}{2}$
=173

Coefficient of Q.D=$\frac{Q3-Q1}{Q3+q1}$

=$\frac{346}{1160}$

=0.2982

2. Compute the Quartile deviation.

weight	60	61	62	63	65	80	75	70
N0.of workers	1	3	5	7	10	1	3	1

Solution:

x	F	c.f
60	1	1
61	3	4
62	5	9
63	7	16
65	10	26
80	1	27
75	3	30
70	1	31

Q3=size of $3\left(\frac{N+1}{4}\right)$ *th item*

=size of $3(\frac{31+1}{4})$

=size of 3(8)

=24

Q3=65 item

Q1=size of ($\frac{N+1}{4}$)*item*

=size of($\frac{31+1}{4}$)

=size of 8^{th} item

$Q.D=\frac{Q3-Q1}{2}$

$=\frac{65-62}{2}$

$=\frac{3}{2}$

Q.D=1.5

Coefficient of $Q.D=\frac{Q3-Q1}{Q3+Q1}$

$=\frac{65-62}{65+62}$

=0.0236

3. Find the quartile deviation for the following deviation.

Marks	0-10	10-20	20-30	30-40	40-50	50-60
Frequency	8	20	25	30	12	5

Solution:

Marks	frequency	c.f
0-10	8	8
10-20	20	28
20-30	25Type equation here.	53
30-40	30	83
40-50	12	95
50-60	5	100

Q1=size of $\left(\frac{N}{4}\right)$

=size of $\left(\frac{100}{4}\right)$

=25item

Q1=10-20

Q1=L+ $\frac{\frac{N}{4}-c.f}{f}$×i

L=10, N=100,c.f=8,f=20,i=10

$= 10 + \frac{\frac{100}{4}-8}{20}$×10

=10+$\frac{17}{2}$

Q1=18.5

Q3=size of 3$\left(\frac{N}{4}\right)$

=size of 3($\frac{100}{4}$)

=3×25

=75th item.

30-40

Q3= L+ $\frac{\frac{3N}{4}-c.f}{f}$×i

=30+$\frac{\frac{3\times100}{4}-53}{30} \times 10$

=30+$\frac{22}{3}$

=30+7.333

Q3 =37.333

Q.D=$\frac{Q3-Q1}{2}$

=$\frac{37.333-18.5}{2}$

=$\frac{18.8333}{2}$

Q.D=9.4167

Coefficient of Q.D=$\frac{Q3-Q1}{Q3+Q1}$

$=\frac{37.33-18.5}{37.33+18.5}$

=0.3373

Mean deviation:

Mean deviation is an absolute measure of dispersion defined as an arithmetic mean of ─the deviations of the individual values from the average of the given data.so it is also called as average deviation.

1. frequency distribution:

Mean deviation $=\frac{\sum f|D|}{N}$

Where $|D| = |x\text{-mean}|$

Where $|D| = |x\text{-median}|$

$N=\sum f$

2.Individual series:

Mean deviation $=\frac{\sum |D|}{n}$

Where $|D| = |x\text{-mean}|$

Where $|D| = |x\text{-median}|$

3.coefficient of mean deviation:

$\frac{mean\ deviation}{mean}$

Or

$\frac{mean\ deviation}{median}$

1. calculate the mean deviation about the mean and about the median for the following data.
15,25,32,46,80,95,98
Solution:

Mean=$\frac{15+25+32+46+80+95+98}{7} = \frac{391}{7}$

=55.857

Median=size of$(\frac{N+1}{2})$

= $(\frac{7+1}{2})$

=4^{th} item.

Median=46

X	|D| =| x-mean |	D| =| x-median |
15	40.8571	3
25	30.8571	21
32	23.8571	14
46	9.8571	0
80	24.1429	34
95	39.1429	49
98	42.1429	52

Mean deviation about mean =$\frac{\sum|D|}{n}$

=$\frac{210.8571}{7}$

=30.1224

Mean deviation about median=$\frac{\sum|D|}{n}$

=$\frac{201}{7}$ = 28.7142

2. calculate the mean deviation about the median from the following data.

X	10	11	13	14	12
F	3	12	12	3	18

Solution:

X	F	C.f	D| =| x-median |	f |D|
10	3	3	2	6
11	12	15	1	12
13	12	33	0	0
14	3	45	1	12
12	18	48	2	6

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

=size of$\left(\frac{48+1}{2}\right)$

=24.5th item

=12

Mean deviation about median=$\frac{\sum f|D|}{N}$

=$\frac{36}{48}$

=0.75

3.calculate the mean deviation about the mean for the following data.

No.of calls	2	3	4	5	6	7
frequency	1	5	8	4	2	1

Solution:

Mean $=\frac{\sum fx}{N}$

x	F	fx	|D| =| x-mean |	f |D|
2	1	2	2.19	2.19
3	5	15	1.19	5.95
4	8	32	0.19	1.52
5	4	20	0.81	3.24
6	2	12	1.81	3.62
7	1	7	2.81	2.81

$=\frac{88}{21}$

$=4.19$

Mean deviation about mean$=\frac{\sum f|D|}{N}$

$=\frac{19.33}{21}$

$=0.9204$

Coefficient of mean deviation$=\frac{mean\ deviation}{mean}$

$=\frac{0.9204}{4.19}$

$=0.2196$

4.calculate the mean deviation and its coefficient from the following.

Income	No. of person
Less than 10	10
Less than 20	25
Less than 30	40
Less than 40	63
Less than 50	85
Less than 60	104
Less than 70	116
Less than 80	120

Solution:

x	F	c.f	midx	D| =| x-median |	f |D|
0-10	10	10	5	33.695	336.95
10-20	15	25	15	23.695	355.425
20-30	15	40	25	13.695	205.425
30-40	23	63	35	3.695	84.985
40-50	22	85	45	6.305	138.71
50-60	19	104	55	16.305	309.795
60-70	12	116	65	26.305	315.66
70-80	4	120	75	36.305	145.22

Median=size of $\left(\frac{N}{2}\right)$

=size of $\left(\frac{120}{2}\right)$

=60th item

=30-40

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

$= 30+\frac{\frac{120}{2}-40}{23}\times 10$

$= 30+\frac{20}{23}\times 10$

$=38.695$

Mean deviation$=\frac{\sum f|\ D\ |}{N}$

$=\frac{1892.17}{120}$

$=15.76$

Coefficient if M.D$=\frac{mean\ deviation}{median}$

$=\frac{15.768}{38.695}$

$=0.407$

Standard deviation:

Standard deviation is the square root of the mean of the square deviation from the arithmetic mean

Coefficient of the variance:

Coefficient of the variation is said be variable or more consistent ,more uniform or more stable.

c.v$=\frac{\sim}{\bar{x}}\times 100$

where ∾is standard deviation

$\bar{x}$=mean

Individual observations:

$$\sim=\sqrt{\frac{\sum x^2}{N}-(\frac{\sum x}{N})^2}$$

Derivations from actual form:

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

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{N}}$$

N=∑f

Derivations for assumed mean:

$$\sigma=\sqrt{\frac{\sum fd^2}{N} - (\frac{\sum fd}{N})^2}$$

$$\sigma=\sqrt{\frac{\sum fd^2}{N} - (\frac{\sum fd}{N})^2}$$

Step deviation method:

$$\sigma=\sqrt{\frac{\sum fd^2}{N} - (\frac{\sum fd}{N})^2}\times i$$

Combined standard mean:

$$\sigma 12=\sqrt{\frac{N1\sigma 1^2+N2\sigma 2^2+N1d1^2+N2d2^2}{N1+N2}}$$

σ12=combined standard deviation

σ1=standard deviation 1st group

σ2=standard deviation 2nd group

D1=x1-x12‾ D2=x2-x12‾

1.compute the standard deviation and coefficient of the variation for the following data.1,5,4,2,3,8,6,2,8

X	x^2
1	1
5	25
4	16
2	4
3	9
8	64
6	36
2	4
8	64

$$\sigma=\sqrt{\frac{\sum x^2}{N}-\left(\frac{\sum x}{N}\right)^2}$$

$$\sigma=\sqrt{\frac{223}{9}-\left(\frac{39}{9}\right)^2}$$

$$=\sqrt{24.778-(4.333)^2}$$

$$=\sqrt{24.778-18.774}$$

$$=\sqrt{6}$$

$$=2.4494$$

$$c.v=\frac{\sigma}{\bar{x}}\times 100$$

$$=\frac{2.4494}{4.333}\times 100$$

$$=0.5652\times 100$$

$$c.v=56.52$$

2. *calculate the coefficient if variation for the* following data.

Size of item	3.5	4.5	5.5	6.5	7.5	8.5	9.5
frequency	3	7	22	60	85	32	8

X	F	D=x-a	d^2	fd	fd^2
3.5	3	-3	9	-9	27
4.5	7	-2	4	-14	28
5.5	22	1	1	-22	22
6.5	60	0	0	0	0
7.5	85	1	1	85	85
8.5	32	2	4	64	128
9.5	8	3	9	25	72

$$\propto=\sqrt{\frac{\sum fd^2}{N}-\left(\frac{\sum fd}{N}\right)^2}$$

$$\propto=\sqrt{\frac{362}{217}-\left(\frac{128}{217}\right)^2}$$

$$=\sqrt{1.6682-0.3478}$$

$$=\sqrt{1.3204}$$

$$=1.1490$$

$$\bar{x}=A+\frac{\sum fd}{\sum f}$$

$$=6.5+\frac{128}{217}$$

$$=7.0898$$

$$c.v=\frac{\propto}{\bar{x}}\times 100$$

$$=\frac{1.1490}{7.0898}\times 100$$

$$=16.20$$

3. from the following data find the which product is more stable in prices.

Price of productA(Rs)	20	22	19	23	16
Price of productB(RS)	10	20	18	12	15

Solution:

X	$(x-\bar{x})$	$(x-\bar{x})^2$	Y	$(y-\bar{y})$	$(y-\bar{y})^2$
20	0	0	10	-5	25
22	2	4	20	5	25
19	-1	1	18	3	9
23	3	9	12	-3	9
16	-4	16	15	0	0

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

$=\frac{100}{5}$

$\bar{x}=20$

$\bar{y}=\frac{\sum y}{n}$

$=\frac{75}{5}$

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

$= \sqrt{\frac{30}{5}}$

$=2.4449$

$c.v=\frac{\sigma}{\bar{x}} \times 100$

$=\frac{2.449}{20} \times 100$

$=12.25$

$$\sigma y=\sqrt{\frac{\sum(y-\bar{y})^2}{n}}$$

$=\sqrt{\frac{68}{5}}$

$=3.6878$

c.v$=\frac{\sigma}{\bar{Y}} \times 100$

$=\frac{3.6878}{15} \times 100$

$=24.58$

Since the coefficient of variation of the prices of product A is less than that of prices of product B. we conclude that product A is more stable in prices.

4.coefficient of variation 2 different distributions are 58% and 69% there standard deviations are 21.2 and 15.6 respectively. What are the arithmetic mean.

Solution:

c.v$=\frac{\sigma}{\bar{x}} \times 100$

c.v=58% $\sigma=21.2$

$58=\frac{21.2}{\bar{x}} \times 100$

$\bar{x}=\frac{21.2}{58} \times 100$

$\bar{x}=36.55$

c.v=69% ∝=15.6

c.v=$\frac{\propto}{\bar{x}} \times 100$

69=$\frac{15.6}{\bar{x}} \times 100$

$\bar{x}$=$\frac{15.6}{69} \times 100$

$\bar{x}$=22.66

5.A consegment of 180 articles is classified according to the size of article has below find the standard deviation and its coeeficien.

Measurement	No. of articles
More than 80	5
More than 70	14
More than 60	34
More than 50	65
More than 40	110
More than 30	150
More than 20	170
More than 10	176
More than 0	180

Solution:

X	F	midx	D	d^2	fd	fd^2
0-10	4	5	-4	16	-16	64
10-20	6	15	-3	9	-18	54
20-30	20	25	-2	4	-40	80
30-40	40	35	-1	1	-40	40
40-50	45	45	0	0	0	0
50-60	31	55	1	1	31	31
60-70	20	65	2	4	40	80
70-80	9	75	3	9	27	81
80-90	5	85	4	16	20	80

$\sigma=\sqrt{\frac{\sum fd^2}{N}-(\frac{\sum fd}{N})^2}\times i$

$\sigma=\sqrt{\frac{510}{180}-(\frac{4}{180})^2}\times 10$

$=\sqrt{2.8333-0.00049}\times 10$

$=16.830$

$c.v=\frac{\sigma}{\bar{x}}\times 100$

$\bar{x}=A+\frac{\sum fd}{\sum f}\times i$

$=45+\frac{4}{180}\times 10$

$\bar{x}=45.2222$

$c.v=\frac{\sigma}{\bar{x}}\times 100$

$c.v=\frac{16.830}{45.222}\times 100$

$c.v=37.2162$

6. The mean and standard deviation of 200 items are found to be 60 and 20 resoectively. If at that time calculations of 2 items where wrongly taken as 3 and 67 instead of 13 and 17.find the the correct mean and standard deviation what is the correct coefficient of variation?

Solution:

$\bar{x}=60$

$\sigma=20$ N=200

Wrong values: 3 and 67

Correct values: 13 and 17

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

$60=\frac{\sum x^2}{200}$

$\sum x=12000$

This value is wrong

correct$\sum$x=wrong$\sum$x-wrongvalues+correct values

=12000-3-67+13+17

=11960

Correct standard deviation:

$\sigma=\sqrt{\frac{\sum x^2}{N}-(\bar{x})^2}$

$20=\sqrt{\frac{\sum x^2}{200}-(60)^2}$

$400=\frac{\sum x^2}{200}$

wrong$\sum$x=800000

correct$\sum x^2 = wrong\sum x^2 - (wrong\ values)^2 + (correct\ values)^2$

$=800000-3^2-67^2+13^2+17^2$

=795960

correct$\sigma=\sqrt{\frac{correct\sum x^2}{N}-(correct\bar{x})^2}$

$=\sqrt{\frac{795960}{200}-(59.8)^2}\times$

Correct ∝=20.094

Correct c.v=$\frac{correct\propto}{correct\bar{x}}\times 100$

$=\frac{20.094}{59.8}\times 100$

=33.602

UNIT-IV

MEASURES OF SKEWNESS

Measures of skewness tell us about the direction and the extend of symmetric or asymmetric in a distribution. It describes the shape of the distribution. If a distribution is not symmetrical it is skewed. In a perfect symmetrical distribution mean, median and mode.

If the frequency curve as s long tail to the right it is skewed to the right. This means that mean is greater than the mode.so the distribution is positively skewed.

If the frequency curve as the long tail to the left it is said to be skewed to the left and mean is less than the mode. so the distribution is negatively skewed.

IMPORTANT METHODS OF SKEWNESS:

I. Karl Pearson's coefficient of skewness
II. Bowley's coefficient of skewness

KARL PEARSON'S COEFFICIENT OF SKEWNESS:

The Pearson's coefficient of skewness is based upon the difference between mean and mode. This differences is divided by standard deviation to give a relative measure.

COEFFICIENT OF SKEWNESS=$\frac{mean-mode}{standard\ deviation}$

(or)

COEFFICIENT OF SKEWNESS=$\frac{3(mean-median)}{standrad\ deviation}$

This value is usually lies between +1 and -1

Problems:

1.calculate Karl Pearson coefficient of skewness from the data given below.

Size(x)	1	2	3	4	5	6	7
Frequency(f)	10	18	30	25	12	3	2

Solution:

Mode=3

X	F	Fx	D=x-a	d^2	fd	fd^2
1	10	10	-3	9	-30	90
2	18	36	-2	4	-36	72
3	30	90	-1	1	30	30
4	25	100	0	0	0	0
5	12	60	1	1	12	12
6	3	18	2	4	6	12
7	2	14	3	9	6	18
	100	328			-72	234

Mean=$\frac{\Sigma fx}{\Sigma f}$

=$\frac{328}{100}$

=3.28

s.d=$\sqrt{\frac{\Sigma fd^2}{N} - (\frac{\Sigma fd}{N})^2}$

=$\sqrt{2.34 - (0.72)^2}$

=$\sqrt{1.8216}$

s.d=1.3496

Skp=$\frac{mean-mode}{standard\ deviation}$

=$\frac{0.28}{1.3496}$

Skp=0.2074

2.calculate the karl pearson coefficient of skewness.

x	12.5	17.5	22.5	27.5	32.5	37.5	42.5	47.5
f	28	42	54	108	129	61	45	33

Solution:

Mean $\overline{x}$=A+$\frac{\Sigma fd}{\Sigma f} \times i$

=27.5+0.52×5

=27.5+2.96

=30.46

S.d=$\sqrt{\frac{\Sigma fd^2}{N} - (\frac{\Sigma fd}{N})^2}$×i

=$\sqrt{3.56 - (0.592)^2} \times 5$

=$\sqrt{3.56 - 0.3504 \times 5}$

=$\sqrt{3.2096 \times 5}$

=1.7915×5

Sd=8.9575

Mode=32.5

Skp=$\frac{mean - mode}{standard\ deviation}$

=$\frac{30.46-32.5}{8.9575}$

=$\frac{-2.04}{8.9575}$

=-0.228

3.calculate the karl pearson coefficient of skewnessof the data given below.

Daily expenditure	0-20	20-40	40-60	60-80	80-100
No of families	13	25	27	19	16

Solution:

X	F	Mid m	d=$\frac{x-a}{i}$	d^2	fd^2	Fd
0-20	13	10	-2	4	52	-26
20-40	25	30	-1	1	25	-25
40-60	27	50	0	0	0	0
60-80	19	70	1	1	19	19
80-100	16	90	2	4	64	32
	100				160	0

Mean $\overline{x}$=A+$\frac{\Sigma fd}{\Sigma f} \times i$

=50+$\frac{0}{100}$×20

=50

S.d=$\sqrt{\frac{\Sigma fd^2}{N} - (\frac{\Sigma fd}{N})^2}$×i

=$\sqrt{\frac{160}{100} - (\frac{0}{100})^2}$×20

=1.2649×20

=25.298

mode=L+$\frac{f1-f0}{2f1-f0-f2}$ ×i

=40+$\frac{2}{54-25-19}$×20

=40+4

=44

Skp=$\frac{mean-mode}{standard\ deviation}$

=$\frac{50-44}{25.298}$

Skp =0.2371

4.calculate from the following data karl pearson coefficient of skewness

Marks more than	0	10	20	30	40	50	60	70	80
No of students	150	140	100	80	80	70	30	14	0

Solution:

X	F	Mid x	d=$\frac{x-a}{i}$	d^2	fd	fd^2	cf
0-10	10	5	-3	9	-30	9	10
10-20	40	15	-2	4	-80	160	50
20-30	20	25	-1	1	-20	20	70
30-40	0	35	0	0	0	0	70
40-50	10	45	1	1	10	10	80
50-60	40	55	2	4	80	160	120
60-70	16	65	3	9	48	144	136
70-80	14	75	4	16	56	224	150
	150				64	808	

Mean $\bar{x}$=A+$\frac{\Sigma fd}{\Sigma f} \times i$

=35+0.4266×10

=35+4.266

=39.266

Median=size of($\frac{N}{2}$)th item

= size of($\frac{150}{2}$)th item

=size of 75th item

=40-50

Median=L+$\frac{\frac{N}{2}-c.f}{f}$×i

=40+$\frac{5}{10}$×10

=40+0.5×10

=40+5

=45

S.d=$\sqrt{\frac{\Sigma fd^2}{N}-(\frac{\Sigma fd}{N})^2}$×i

=$\sqrt{5.3866-0.1819}$×10

=$\sqrt{5.2047}$×10

=2.2813×10

=22.813

COEFFICIENT OF SKEWNESS=$\frac{3(mean-median)}{standrad\ deviation}$

=$\frac{3(39.266-45)}{22.813}$

=$\frac{-17.202}{22.813}$

Skp=-0.7540

BOWLEY'S COEFFICIENT OF SKEWNESS(Skb):

The bowley's coefficient of skewness is based on quartiles.

Skb=$\frac{Q3+Q1-2median}{Q3-Q1}$

This is also called as quartile measure of skewness. And it varies between -1 and +1.

Problems:

1.find the bowley's coefficient of skewness for the following frequency distribution.

No of children	0	1	2	3	4	5	6
No of families	7	10	16	25	18	11	8

Solution:

X	F	Cf
0	7	7
1	10	17
2	16	33
3	25	58
4	18	76
5	11	87
6	8	95

Q3=size of $(\frac{3N+1}{4})$th item

=size of $3(\frac{96}{4})$th item

=size of 72th item

=4

Q1= size of $(\frac{N+1}{4})$th item

=size of $(\frac{96}{4})$th item

=24th item

=2

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

=48th item

Median= 3

Skb=$\frac{Q3+Q1-2median}{Q3-Q1}$

=$\frac{4+2-6}{4-2}$

Skb=0

2.calculate bowley's coefficient of skewness for the following distribution.

x	10-20	20-30	30-40	40-50	50-60	60-70	70-80
F	358	2417	976	129	62	18	10

Solution:

X	F	Cf
10-20	358	358
20-30	2417	2775
30-40	976	3751
40-50	129	3880
50-60	62	3942
60-70	18	3960
70-80	10	3970

Q3=Size of $(\frac{3N}{4})$th item

=Size of $(\frac{3\times3970}{4})$th item

=2977.5th item

30-40

Q3=L+$\frac{\frac{3N}{4}-c.f}{f}$×i

$=30+\frac{202.5}{976}\times 10$

$=30+2.20747$

Q3=32.0747

Q1=Size of $(\frac{N}{4})$th item

=size of $(\frac{3970}{4})$th item

Q1=992.5

$Q1=L+\frac{\frac{N}{4}-c.f}{f}\times i$

$=20+\frac{992.5-358}{2417}\times 10$

$=20+0.2625\times 10$

$=20+2.6251$

Q1=22.6251

Median= size of $(\frac{N}{2})$th item

= size of $(\frac{3970}{2})$th item

=1985

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

$=20+\frac{1627}{2417}\times 10$

$=20+6.7314$

$=26.7314$

$Skb=\frac{Q3+Q1-2median}{Q3-Q1}$

$=\frac{32.0747+22.6251-2(26.7314)}{32.0747-22.6251}$

$=\frac{1.237}{9.4496}$

=0.1309

3.calculate the bowley's coefficient of skewness from the data given below.

Profit (Rs in lakhs)	Less than 10	20	30	40	50	60	70
No of companies	8	20	40	50	56	59	60

Solution:

X	F	Cf
0-10	8	8
10-20	12	20
20-30	20	40
30-40	10	50
40-50	6	56
50-60	3	59
60-70	1	60

Q3=Size of $(\frac{3N}{4})$th item

=size of $3\times\frac{60}{4}^{th}$ item

=45th item

30-40

Q3=L+$\frac{\frac{3N}{4}-c.f}{f}\times$i

=30+$\frac{45-40}{10}\times$10

=30+5

=35

Q1=Size of $(\frac{N}{4})$th item

=size of $(\frac{60}{4})$th item

=15th item

Q1=L+$\frac{\frac{N}{4}-c.f}{f}$×i

=10+$\frac{15-8}{12}$×10

=10+5.5833

=15.8333

Median= size of $(\frac{N}{2})$th item

=size of$(\frac{60}{2})$ th item

=30th item

Median=L+$\frac{\frac{N}{2}-c.f}{f}$×i

=20+$\frac{30-20}{20}$×10

=20+5

=25

Skb=$\frac{Q3+Q1-2median}{Q3-Q1}$

=$\frac{35+15.8333-2(25)}{35-15.8333}$

Skb=0.0434

4.In a frequency distribution the coefficient of skewness based on quartiles 0.6 if the sum of the upper and lower quartile is 100 and the median is 38.find the value of upper quartile.

Skb=$\frac{Q3+Q1-2median}{Q3-Q1}$

Q3-Q1=$\frac{100-76}{0.6}$

=$\frac{24}{0.6}$

Q3-Q1=40

Q3=40+Q1

Q3+Q1=100

40+Q1+Q1=100

40+2Q1=100

2Q1=100-40

2Q1=60

Q1=30

Q3=40+30

Q3=70

5.You are given the skp=0.8,mean=40 and mode=36.Find the value of standard deviation.

Skp=$\frac{mean-mode}{standard\ deviation}$

0.8=$\frac{40-36}{standard\ deviation}$

s.d=$\frac{4}{0.8}$

s.d=5

6.a frequency distribution showed the following measyres of location mean=45,median=48 coefficient of skewness=-0.4 estimate its standard deviation.

COEFFICIENT OF SKEWNESS=$\frac{3(mean-median)}{standrad\ deviation}$

$$sd=\frac{3(45-48)}{-0.4}$$

$$=\frac{3(-3)}{-0.4}$$

$$=\frac{-9}{-4}$$

Sd=22.5

Moments:

The arithmetic mean of the various of the deviation from mean in any distribution is called the moments about the mean or central moments of the distribution.

The rth moment about the mean denoted by $\mu_{r}=\frac{\Sigma(x-\bar{x})^r}{n}$

For a frequency distribution $\mu_{r}=\frac{\Sigma f(x-\bar{x})^r}{\Sigma f}$

When the actual mean $\bar{x}$ is in a fraction it is difficult to calculate moments about the mean by applying the above formula. In such case we first compute moments about and orbitrary value(A) called row moments and then convert these moments into moment about mean.

Th rth moment about any point A is given by $\mu_{r}=\frac{\Sigma(x-A)^r}{n}$

For a frequency $\mu_{r}=\frac{\Sigma(x-A)^r}{n}$

Conversions of moments about AX moments about mean:

To obtain moments about mean, we applying the following relationship

$$\mu^1 = \mu_1' - \mu_1' = 0$$

$$\mu^2 = \mu_2' - (\mu_1')^2$$

$$\mu^3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1^1)^3$$

$$\mu^4 = \mu_4' - \mu_1'\mu_3' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$$

Skewness:

A measure of skewness is obtained by making use of the second and third moments about the mean and it is denoted by β1

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

Kurtosis:

Kurtosis is a measure which studies about the platness or peakness of the frequency curve of the distribution.

The measure of kurtosis is denoted by β2 is also used as a measure of kurtosis and is defined by

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

Y2 obtained from β2 is also used as a measure of kurtosis and is defined by

Y2=β2-3

For a normal curve Y2=0 that is β2=3.Then the curve is called mesokurtic.

When Y2 is positive that β2>3 then the curve is more peaked than the normal curve and it is called leptokurtic.

When Y2 is negative that is β2<3 then curve is less peaked than the normal curve and it is called platykurtic.

Problem:

1.compute the first four central moments for the following data.

8,10,11,12,14.

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

$=\frac{8+10+11+12+14}{5}$

=11

X	$(x-\bar{x})^1$	$(x-\bar{x})^2$	$(x-\bar{x})^3$	$(x-\bar{x})^4$
8	-3	9	-27	81
10	-1	1	-1	1
11	0	0	0	0
12	1	1	1	1
14	3	9	27	81
55	0	20	0	164

$\mu_r=\frac{\Sigma(x-\bar{x})^r}{n}$

$\mu_1=\frac{\Sigma(x-\bar{x})^1}{n}$

=0/5

=0

$\mu_2=\frac{\Sigma(x-\bar{x})^2}{n}$

=20/5

=4

$\mu_3=\frac{\Sigma(x-\bar{x})^3}{n}$

=0/5

=0

$\mu_4 = \frac{\Sigma(x-\bar{x})^4}{n}$

=164/5

=32.8

2.The first four momentums of the distribution abut the value 4 of a variable are -1.5,17,-30 and 108.Find the central moment β1 and β2.

Solution:

Moment about the value 4

-1.5,17,-30 and 108

Moments about mean

$$\mu^1 = \mu_1' - \mu_1' = 0$$

$$\mu^2 = \mu_2' - (\mu_1')^2$$

=17-2.25

=14.75

$$\mu^3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1^1)^3$$

=-30-3(-1.5)(17)+2$(-1.5)^3$

=39.75

$$\mu^4 = \mu_{4-}' 4\mu_1'\mu_3' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$$

=108-4(-1.5)(-30)+6(17)$(-1.5)^2$-3$(1.5)^4$

=142.3125

$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$

$=\frac{1580.0625}{3209.046}$

=0.4923

$\beta_2 = \frac{\mu_4}{\mu_2^2}$

$=\frac{142.3125}{217.5625}$

=0.6541

3.calculate the first four central moments for the following data

X	2	3	4	5	6
F	1	3	7	3	1

Solution:

X	F	fx	(x-x̅)	f(x-x̅)	f(x-x̅)2	f(x-x̅)3	f(x-x̅)4
2	1	2	-2	-2	4	-8	16
3	3	9	-1	-3	9	-27	81
4	7	28	0	0	0	0	0
5	3	15	1	3	9	27	81
6	1	6	2	2	4	8	16
	15	60		0	26	0	194

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

=60/15

=4

$\mu_1=\frac{\Sigma f(x-\bar{x})^1}{\Sigma f}$

=0

$\mu_2 = \frac{\Sigma f(x-\bar{x})^2}{\Sigma f}$

=26/15

=1.733

$\mu_3 = \frac{\Sigma f(x-\bar{x})^3}{\Sigma f}$

=0

$\mu_4 = \frac{\Sigma f(x-\bar{x})^4}{\Sigma f}$

=194/15

=12.933

4.Analysis the following distribution by the following distribution by the method of moments.

X	2	4	6	8	10	12	14
F	4	11	18	27	20	16	8

Solution:

X	F	d=$\frac{x-a}{i}$	d^2	d^3	d^4	Fd	fd^2	fd^3	fd^4
2	4	-3	9	-27	81	-12	36	-108	324
4	11	-2	4	-8	16	-22	44	-88	176
6	18	-1	1	-1	1	-18	18	-18	18
8	27	0	0	0	0	0	0	0	0
10	20	1	1	1	1	20	20	20	20
12	16	2	4	8	16	32	64	128	256
14	8	3	9	27	81	24	72	216	648
	104					24	254	150	1442

$\mu_1' = \frac{\Sigma fd}{\Sigma f} \times i$

$=\frac{24}{104}\times 2$

$=0.4615$

$\mu_2'=\frac{\Sigma f d^2}{\Sigma f}\times i^2$

$=\frac{254}{104}\times 4$

$=9.7692$

$\mu_3'=\frac{\Sigma f d^3}{\Sigma f}\times i^3$

$=\frac{150}{104}\times 8$

$=11.538$

$\mu_4'=\frac{\Sigma f d^4}{\Sigma f}\times i^4$

$=\frac{1442}{104}\times 16$

$=221.85$

$$\mu^1 = \mu_1' - \mu_1' = 0$$

$$\mu^2 = \mu_2' - (\mu_1')^2$$

$=9.7692-0.4615^2$

$=9.5562$

$$\mu^3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1^1)^3$$

$=11.538-3(0.4615)(9.7692)+2(0.4615)^3$

$=-1.74$

$$\mu^4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$$

$=221.85-4(0.4615)(11.538)+6(9.7692)(0.4615)^2-3(0.4615)^4$

=212.88

Skewness:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

$$=\frac{-1.74^2}{9.56^3}$$

=0.00346

Kurtosis:

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$= \frac{212.88}{9.5562^2}$$

=2.329

5.calculate the skewness and kurtosis for the following

Marks	0-10	10-20	20-30	30-40	40-50	50-60
Frequency	5	20	15	45	10	5

Solution:

X	f	Mid x	d=$\frac{x-a}{i}$	fd	d^2	fd^2	d^3	fd^3	d^4	fd^4
0-10	5	5	-2	-10	4	20	-8	-40	16	80
10-20	20	15	-1	-20	1	20	-1	-20	1	20
20-30	15	25	0	0	0	0	0	0	0	0
30-40	45	35	1	45	1	45	1	45	1	45
40-50	10	45	2	20	4	40	8	80	16	160
50-60	5	55	3	15	9	45	27	135	81	405
	100			50		170		200		710

$\mu_1' = \frac{\Sigma fd}{\Sigma f} \times i$

$= \frac{50}{100} \times 10$

$=5$

$\mu_2' = \frac{\Sigma fd^2}{\Sigma f} \times i^2$

$= \frac{17000}{100}$

$=170$

$\mu_3' = \frac{\Sigma fd^3}{\Sigma f} \times i^3$

$=200000/100$

$=2000$

$\mu_4' = \frac{\Sigma fd^4}{\Sigma f} \times i^4$

$= \frac{7100000}{100}$

$=71000$

$$\mu^1 = \mu_1' - \mu_1' = 0$$

$$\mu^2 = \mu_2' - (\mu_1')^2$$

$=170-25$

$=145$

$$\mu^3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1^1)^3$$

$=2000-3(5)(170)+250$

$=-300$

$$\mu^4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$$

=71000-4(5)(2000)+6(170)(25)-1875

=54625

Skewness:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

$$=\frac{90000}{3048625}$$

=0.0295

Kurtosis:

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$=\frac{54625}{21025}$$

=2.5980

6.The first four central moments about the distribution are 2,6,12 and 100.Find β1 and $\beta 2$.

Solution:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

$$=\frac{12^2}{6^3}$$

=0.6666

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$=\frac{100}{6^2}$$

=2.7777