10/24

Dispersion:

Dispersion is the measures of variation of the observation from the average

No deviation / small range deviation	Slight deviation	wide range deviation
A	B	C
100	100	1
100	105	489
100	102	2
100	103	3
100 = 500	90 = 500	5 = 500
$\bar{x}_A = 100$	$\bar{x}_B = 100$	$\bar{x}_C = 100$

Here we observed that, AM for all three series is same,

- but in series A observations does not deviate from AM, so no despersion.
- In series B one observation is represented by AM and other obsesvations are slightly deviated from AM as compared with series C.
- In series C not a single observation is represented by AM and there is a wide range of deviation as compared with Series B.

In such cases central tendency are insufficient to conclude the data and hence, it is required to find dispersion by two ways:

(a) mean deviation
(b) Standard deviation.

For D.D

↳ 1) mean deviation : $\frac{1}{N} \Sigma\left(f_i \cdot |x_i - A|\right)$

where $N = \Sigma f_i$ and $|x_i - A|$ = deviation.
where A = mean/median/mode

2) std deviation
It is also known as root mean square deviation, and it is denoted by σ.

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

where $\bar{x}$ = actual AM
n = no. of observation.

ⓑ DD :

take h = 1

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

where A = Assumed mean
$u = x - A$

ⓒ GD :

$$\sigma = h \times \sqrt{\frac{\Sigma f_i u_i^2}{\Sigma f_i} - \left(\frac{\Sigma f_i u_i}{\Sigma f_i}\right)^2}$$

where $u = \frac{x - A}{h}$

A = Assumed mean
h = class width.

Variance:
variance is denoted and defined as $V = \sigma^2$

coefficient of variation

→ It is used to compare the variability of two or more than two series

→ The series for which the CV is greater is said to be more variable or less consistent or less uniform or less stable.

→ where $C.V = \frac{\sigma}{\bar{x}} \times 100$

Que) calculate Std deviation for the following freq. distribution and decide whether the AM is good avg or not.

1)

CI	X	f	$u=\frac{X-A}{h}$	fu	fu^2
0-10	5	5	-3	-15	45
10-20	15	12	-2	-24	48
20-30	25	30	-1	-30	30
30-40	$\boxed{35}$	45	0	0	0
40-50	45	50	1	50	50
50-60	55	37	2	74	148
60-70	65	21	3	63	189
		$\boxed{\Sigma fi = 200}$		118	510

$$\sigma = 10\sqrt{\frac{510}{200} - \left(\frac{118}{200}\right)^2}$$

$$= 10\sqrt{2.55 - 0.3481}$$

$$= 10\sqrt{2.2019}$$

$$= \boxed{14.83}$$

$$\overline{X} = A + h\frac{\Sigma fiui}{\Sigma fi} = 35 + 10 \times \frac{118}{200}$$

$$= 35 + 10 \cdot 0.59$$

$$= 35 + 5.9$$

$$= \boxed{40.9}$$

(for the transaction of 40.9 → change 14.83 नहीं → Not good at all)

$\sigma = 14.8$ for the AM $\boxed{40.9}$ and hence the AM is not a good avg for this data.

Que) calculate mean, median, mode for the following data

CI	midvalue X	f	u = X − A/h	fiui	c.f	
25-30	27.5	10	−4	−40	10	
30-35	32.5	13	−3	−39	23	
35-40	37.5	18	−2	−36	~~31~~ 41	
40-45	32.5	21	−1	−21	~~52~~ 62	
45-50	47.5	24	0	0	~~76~~ $\boxed{86}$	median
$\boxed{50-55}$	52.5	$\boxed{28}$	1	28	~~104~~ 114	modal class
55-60	57.5	20	2	40	~~124~~ 134	
60-65	62.5	11	3	33	~~135~~ 145	
65-70	67.5	8	4	32	~~143~~ 153	
70-75	72.5	7	5	35	~~150~~ 160	
x	x			32	928	

$\frac{160}{2} = 80 \leq$ Just 86

mean: $\bar{x} = A + h\frac{\Sigma f_i u_i}{\Sigma f_i}$

$= 47.5 + \frac{5 \cdot 32}{160} = \boxed{48.5}$

modal class → corresponding to max. freq.

classmate

Date

Page

median: $l + \frac{h}{f}\left(\frac{N}{2} - c\right)$

$45 + \frac{5}{24}\left(\frac{160}{2} - 62\right)$

$45 + 3.75 = \boxed{48.75}$

mode: $l + \frac{h(f_1 - f_0)}{2f_1 - f_0 - f_2}$

$50 + \frac{5(2\cancel{4}8 - 24)}{2 \cdot 28 - 24 - 20} = 50 + \frac{5 \cdot \cancel{4}}{\cancel{12}\,3}$

$= 50 + 1.67$

$= \boxed{51.67}$


classmate
Date
Page
Page No

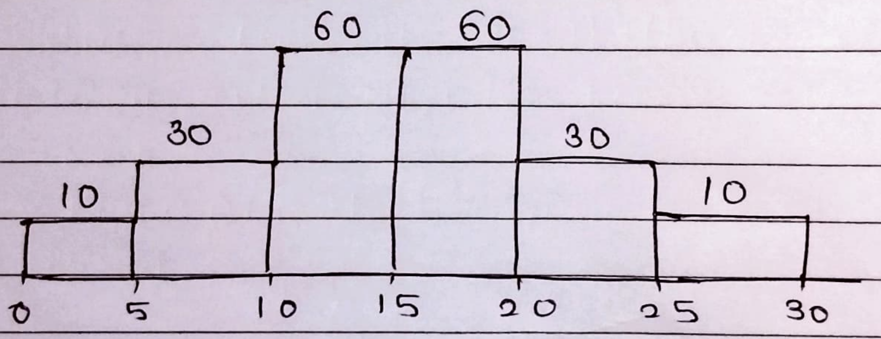
16/01/23

moment skewness and Kurtosis

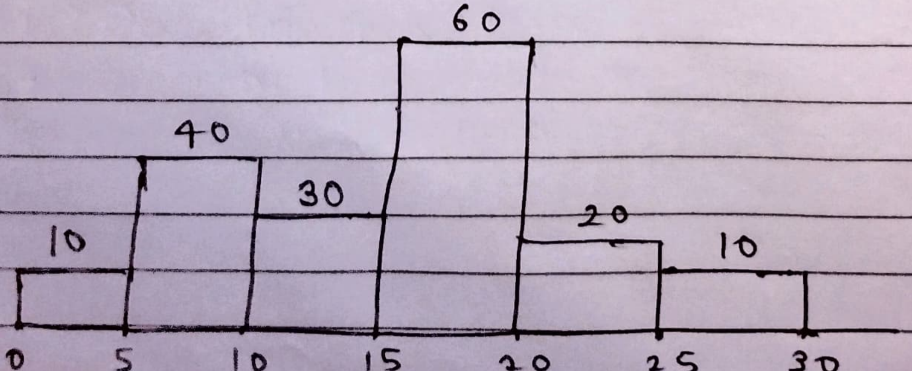
Skewness:
Like measures of central tendenry and Variability, there are 2 other comparble characteristic called: Skewness
Kurtosis
which helps us to understand distribution

Two distributions may have same mean and standard deviation but may differ widly in there overall appearnce

Histograph

1) 

$\overline{X} = 15$ and $\sigma = 6$

2) 

$\overline{X} = 15$ and $\sigma = 6$

(A.M)

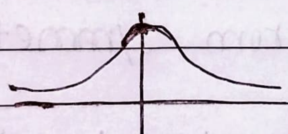
In both these distribution $\bar{x}$ and σ (S.D)

the distribution in 1) is symmetric distribution. and the distribution in 2) is not symmetrical i.e its skeweed distribution.

→ skewness refers to lack of symmetry, in this case distribution is called skeweed distribution.

ⓐ Symmetrical Distribution:

In symmetric distribution the values of mean, median and mode coincide i.e the spread of freq. distribution is same on both sides of the central point of the curve.



$\bar{x}$ = median = mode

$\beta_1 = 0$

bell shaped dig.

ⓑ Asymmetrical distribution:

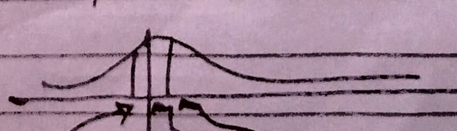
→ A distribution which is not symmetrical is called skeweed distribution

→ and it could be either positively skweed or negatively skweed

a) +ve skeweed distribution

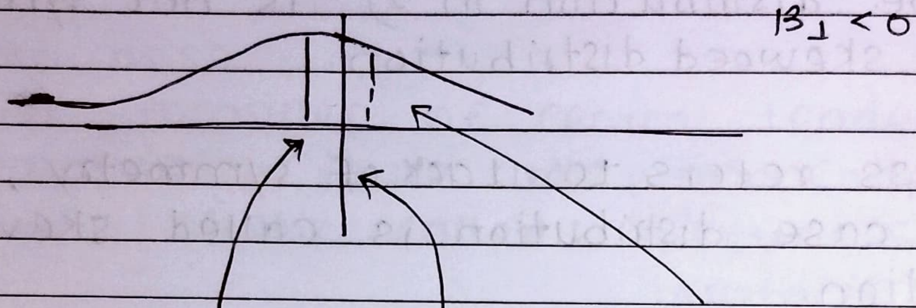
tail right side

$\beta_1 > 0$



mode < median < mean.

b) -vely skewed distribution.



mean < median < mode

***

→ Dispersion is concered with the amount of variation rather than with its derection and

→ skewness tells us about the direction of variation or departure from symmetry.

→ coefficient of skewness is denoted and defined as :

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

a) $\beta_1 > 0$ Distribution is +ve skewed

b) $\beta_1 < 0$ Distribution is -vely skewed

c) $\beta_1 = 0$ Distibution is symmetric

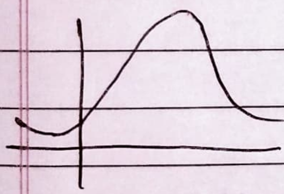
Kurtosis :

→ Kurtosis refers to the degree of flatness or peakedness of the curve

→ Or it refers to the bulginess of the curve.

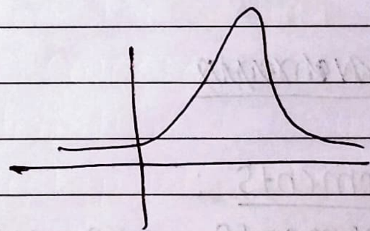
→ i.e kurtosis tells us extend to which a distribution is more peaked or flat toped than normal distribution.

ⓐ mesokurtic



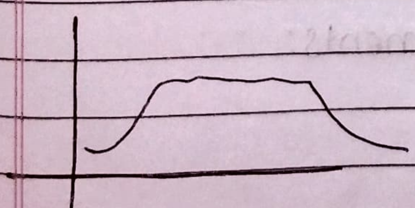
normal curve

ⓑ Leptokurtic



curve is more peaked than normal

ⓒ platykurtic



curve is more flat topped than normal curve.

* Kurtosis is denoted and defined as

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

a) $\beta_2 > 3$ the distribution is leptokurtotic
b) $\beta_2 < 3$ the —||— platykurtotic
c) $\beta_2 = 3$ the —||— mesokurtotic

→ ~~movements~~

moments :
moments are used to describe the various characteristic of the distribution like
- central tendency,
- Variation
- Kurtosis

There are two types of moments:

ⓐ central moments:
(moments about AM)
denoted by (μ)

⇒ zeroth central moment = $\boxed{\mu_0 = 1}$

⇒ 1st central moment $\boxed{\mu_1 = 0}$

(moments about AM)

(a) ID : $\mu_1 = \frac{\Sigma (x_i - \bar{x})}{n} = 0$

(b) DD : (for freq. distribution)

$\mu_1 = \frac{\Sigma f_i u_i}{\Sigma f_i} = \frac{\Sigma f_i u_i}{N} \neq 0$ where

$u = x - \bar{x}$

(c) GD :

$\mu_1 = \frac{\Sigma f_i u_i}{\Sigma f_i} = \frac{\Sigma f_i u_i}{N}$

$u = \frac{x - \bar{x}}{h}$

⇒ 2nd central moment μ_2

i) for ID : $\mu_2 = \frac{\Sigma (x_i - \bar{x})^2}{n}$

ii) for DD : $\mu_2 = \frac{\Sigma f_i u_i^2}{N}$; $u_i = x - \bar{x}$

iii) GD : $\mu_2 = \frac{\Sigma f_i u_i^2}{N}$; $u_i = \frac{x - \bar{x}}{h}$.

⇒ 3rd central moments μ_3

i) ID $\mu_3 = \Sigma (x_i - \bar{x})^3 / n$

ii) DD $\mu_3 = \Sigma f_i (u_i^3) / N$ $u_i = x_i - \bar{x}$

iii) GD $\mu_3 = \frac{\Sigma f_i u_i^3}{N}$ $u_i = \frac{x_i - \bar{x}}{h}$

→ 4th central moment

i) ID $\mu_4 = \frac{(x_i - \bar{x})^4}{n}$

ii) DD $\mu_4 = \frac{\Sigma f_i (u_i^4)}{N}$ $u_i = x - \bar{x}$

iii) GD $\mu_4 = \Sigma \frac{f_i u_i^4}{N}$ $u_i = x - \bar{x} / n$

Note for $\bar{x}$

→ GD for

$\bar{X} = A + h \frac{\Sigma f_i u_i}{\Sigma f_i}$

㊀ raw moments: μ'

moments abt arb. origin)

↳ arb. value/assumed mean.

1) $\mu_0' = 1$

2) $\mu_1' =$

a) ID $\frac{\Sigma (x_i - A)}{n} = \bar{X} - A$

b) DD $\frac{\Sigma f u_i}{\Sigma f_i} = \frac{\Sigma f_i u_i}{N}$ $u = x - A$

c) GD $h \frac{\Sigma f_i u_i}{\Sigma f_i}$

3) 2nd raw moment μ_2'

ID $\mu_2' = \frac{\Sigma (x_i - A)^2}{n}$

DD $\mu_2' = \frac{\Sigma f_i u_i^2}{\Sigma f_i} = \frac{\Sigma f_i (u_i^2)}{N}$; $u_i = x - A$

GD $\mu_2' = h^2 \frac{\Sigma f_i u_i^2}{\Sigma f_i} = h^2 \frac{\Sigma f_i u_i^2}{N}$; $u_i = \frac{x - A}{h}$

4) 3rd row moment μ_3'

ID $\mu_3' = \frac{\Sigma (x_i - A)^3}{n}$

DD $\mu_3' = \frac{\Sigma f_i u_i^3}{\Sigma f_i} = \frac{\Sigma f_i u_i^3}{N}$; $u_i = x - A$

GD $\mu_3' = h^3 \frac{\Sigma f_i u_i^3}{\Sigma f_i}$

5) 4th row moment μ_4'

ID $\frac{\Sigma (x_i - A)^4}{n}$

DD $\frac{\Sigma f_i (u_i^4)}{\Sigma f_i}$

GD $\frac{h^4 \times \Sigma f_i (u_i)^4}{\Sigma f_i}$

Relation between row and central moment.

→ $\boxed{\mu_0 = 1}$

→ $\boxed{\mu_1 = 0}$

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1

→ $\mu_2 = \mu_2' - (\mu_1')^2$

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

→ $\mu_4 = 1\mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$

- μ_2 is the 2nd moment about mean which gives you AM which gives us variance
- μ_3 3rd moment about mean; it measures skewness
- μ_4 4th moment about mean, it measures about kurtosis

moment	what is measures
μ_2	variance : $\mu_2 = \sigma^2$ $\quad \sigma = S.D = \sqrt{\mu_2}$
μ_3	skewness: $\beta_1 = \mu_3^2 / \mu_2^3$
μ_4	kurtosis: $\beta_2 = \frac{\mu_4}{\mu_2^2}$

solve the ex. analyse the freq distribution by method of moments.

x :	2	3	4	5	6
f :	1	3	7	3	1

x	f	c.f	$f_i x_i$	$u = x - \bar{x}$	fu	fu^2	fu^3	fu^4
2	1	1	2	-2	-2	4	-8	16
3	3	4	9	-1	-3	3	-3	3
4	7	11	28	0	0	0	0	0
5	3	13	15	1	3	3	3	3
6	1	14	6	2	2	4	8	16
	N=15		$\Sigma f_i x_i$ 60		0	14	0	38

∴ Here we observed the given distribution is discrete freq. distribution

$$\frac{\Sigma f_i x_i}{\Sigma f_i} = \bar{x} = \frac{60}{15} = 4 \text{ int. value}$$

Therefore directly ~~co~~ central mommements are obtained from the table.

$\mu_2 = 0.93$

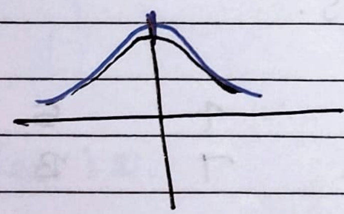
→ $\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{0}{14} = 0$ symmetric

slightly platykurtic.

→ $\beta_2 = \cancel{\frac{3/8}{14}} = \cancel{2.7} < 3 \quad \cancel{\frac{2/9}{13/06}} = \frac{38/15}{(14/15)^2} = \frac{2.5}{0.88} = \boxed{2.8 < 3}$

→ $\sigma^2 = \mu_2 = 14/15 = 0.93$

$\sigma = \sqrt{0.93}$



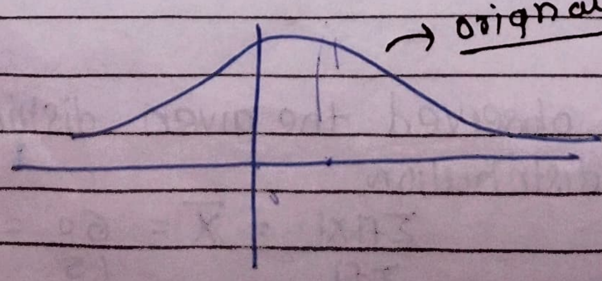
2) $\mu_1 = 0$ $\quad \mu_2 = 2.83$ Comment on skewness & kurtosis.

$\mu_3 = 3.38$ $\quad \mu_4 = 30.295$

$\beta_1 = \frac{11.4244}{22.66} = \boxed{0.5}$ slight +vly skewed

$\beta_2 = \frac{30.295}{8.0089} = \boxed{3.8}$ Leptokurtic

→ orignal



19/01/24

3) 4 moments of a distribution abt $x = 2$ are 1, 2.5, 5.5, 16.
Comment on β_1 and β_2
also find $\bar{x}$, variance and S.D.

→ ~~AM~~ ~~$\bar{x} = A\mu_1'$~~

Above data is raw data.

$A = 2$
$\mu_1' = 1$
$\mu_2' = 2.5$
$\mu_3' = 5.5$
$\mu_4' = 16$

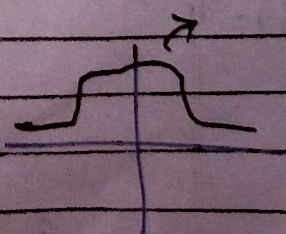
1
1 1
1 2 1
1 3 3 1
1 4 6 4 1

$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$ $\beta_2 = \frac{\mu_4}{\mu_2^2}$

$\mu_2 = 2.5 - 1 = \boxed{1.5}$
$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3$
$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1)^4$

$\mu_3 = 5.5 - 7.5 + 2 = \boxed{0}$
$\mu_4 = 16 - 22 + 6(2.5) - 3$
$= 16 + 15 - 22 - 3$
$= \boxed{6}$

slightly flat



$\beta_1 = 0$ $\beta_2 = \frac{6}{1.5 \times 1.5} = \boxed{2.67}$

↳ symmetric ↳ platykurtic.

$= \mu_1' = 1$

$1 = \bar{x} - A$

$2 + 1 = \bar{x}$

$\boxed{\bar{x} = 3}$

$V = \sigma^2$

$\mu_2 = \text{variance} = \boxed{1.5}$

$S.D = \sqrt{1.5} = \boxed{1.22}$

Que) 4 calculate 1st 4 moments about mean comment of skewness and kurtosis

a) also find σ^2, S.D, $\bar{x}$.

mks	x (midvalue)	F	$u = \frac{x-A}{h}$	f_iu_i	$f_iu_i^2$	$f_iu_i^3$	$f_iu_i^4$
0-10	5	8	-3	-24	72	-216	648
10-20	15	12	-2	-24	48	-96	192
20-30	25	20	-1	-20	20	-20	20
30-40 ~~40-50~~	$\boxed{35}$	30	0	0	0	0	0
40-50 ~~50-60~~	45	15	1	15	15	15	15
~~50~~ ~~70-8~~							
50-60 ~~60-70~~	55	10	2	20	40	80	160
60-70	65	5	3	15	45	~~95~~	~~285~~ 405
		$\boxed{100}$		$\boxed{-18}$		135	~~1680~~
					$\boxed{240}$	$\boxed{-102}$	$\boxed{1440}$

A = 35 centralvalue.