Probability & Statistics

I.T3

9/2/2021
2:20 to 4:30
(1^{st} class)

Dr. B. R. Sreedhar
Dept of Mathematics

Unit-I

Topics:-

- → M.C.T — M, Md, Mo } formu
- → M. Disper — Range, Q.D, M.D, S.D, C.V } formula
- → Skewness — SkP, SkB
- → Kurtosis → Def
- → Curve fitting →
  - $y = a + bx$ [St line]
  - $y = a + bx + cx^2$ [2nd par]
  - $y = ae^{bx}$ [exp]
  - $y = ax^b$ [G.C]
  - $y = ab^x$ [P.C]

(data) — Primary data
(data) — Secondary data

Primary data, Secondary data => Research Gatteing Informa

data — Grouped data
data — unGrouped data

data — open data
data — closed data

Grouped data, unGrouped data, open data, closed data: Analysis part

Def. Statistics:-

↳ process

According Prof. R.A. Fisher, Statistics is defined as

Collection of data

Tabulation of data

Analysis of data

Interpretation of data

⟹ Information (or) Raw material

Primary data < Questionare
Interview

Secondary data $\longrightarrow$ Standare & Source

Grouped data:-

(100) Students

External marks → (0 – 70)
↳ random)

Closed data

C.I	freq
0 – 20	10
20 – 40	30
40 – 60	50
60 – 80	10

[0, 20) ← less 20

less 40

Less 60

Less 80

⇓ fixed

freq (variable)

frequency dis^n

$(X = 30)$ | $X = 40$

open data

0 9

10 19

20 29

20 39

40 49

Open data

Closed

0.5 - 9.5

9.5 - 19.5

19.5 - 29.5

29.5 - 39.5

(Ungrouped data | Raw data)

$$\left(\underbrace{\overbrace{-3, 10, 40, 120, 350, 1060, 2010}}\right)$$

— x —

Grouped data:- Is the data in the form of C.I & frequency together then the data is known as Grouped data.

(or)

Distributing the frequencies to their respective c.I's is also known as <u>frequency disn.</u>

<u>Closed data:-</u> If the class intervals are continous form, without any discontinuity then the data is known as closed data <u>~~ot~~ otherwise</u>

<u>open data</u>:

Ungrouped data / Raw data ?

If the data contains only observations without any C.I's, then the data is known as Raw data.

Measures of Central Tendency:

Mean :- (Arithmatic Mean):

$$\bar{x} = \frac{\text{Sum of the observations}}{\text{no of observations}} = \frac{\sum_{i=1}^{n} x_i}{n}$$

(UGD)

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{N}$$

(GD)

f_i :- frequency

x_i :- Mid Point $= \frac{UL + LL}{2}$

N : Sum of freq's

1
2
3
4
5

$n = \frac{5}{T}$

$N = 15$

find A.M ($\bar{x}$) for following observations

" −3, 6, 1, 7, 11, 2, 0, 8, 3 "

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

$$= \frac{-3+6+1+7+11+2+0+8+3}{9}$$

$$\bar{X} = \frac{35}{9} = 3.88 \,//$$

$$\boxed{\therefore \bar{X} = 3.88}$$

find mean for the following freq data ?

C.I	freq
0-2	3
2-4	9
4-6	12
6-8	11
8-10	5
10-12	2

Sol –

C·I	freq (f)	Midpoint (x_i)	$f_i x_i$
0–2	3	1	3
2–4	9	3	27
4–6	12	5	60
6–8	11	7	77
8–10	5	9	45
10–12	2	11	22
	42		234

$$\bar{X} = \frac{\Sigma f_i x_i}{N}$$

$$\therefore M = \bar{x} = \frac{234}{42}$$

$$= 5.57$$

Median (Middle Value)

eg ①: $-3, 7, 0, 1.2, 0.8, 2.6, 11, 14, -8, 1, 5.$

Sol: ✓ order (Ascending or desending)

$$\underbrace{-8, -3, 0, 0.8, 1,}_{5} \boxed{1.2,} \overbrace{2.6, 5, 7, 11, 14}^{5}$$

middle

$(n = 11)$
odd no

$$\boxed{\therefore Md = 1.2}$$

Eg (2): 120, 180, 4.3, 6.9, -1.9, 0, -11.5, -200, 10, 1

Sol:- (order)

$\Rightarrow$ -200, -11.5, -1.9, 0, $\boxed{1, 4.3}$, 6.9, 10, 120, 180

Middle observations

n = 10

$\Downarrow$

even

$$Md = \frac{1+4.3}{2} = \frac{5.3}{2} = 2.65$$

Grouped / freq dis^n :-

$$Md = l + \left(\frac{N/2 - m}{f} \right) c$$

Here

l : Lower value of the class

f : frequency "

m : Cumulative freq for above i.e. $\rightarrow [N/2]$

c : size of CI

N : sum of freq

find median for the freq distn.

C.I	0-5	5-10	10-15	15-20	20-25	25-30
freq	17	35	50	55	29	11

Sol: –

C.I	freq	Cumulati freq
0-5	17	17
5-10	35	52 (cm)
$\boxed{10\text{-}15}$	50 f	102
15-20	55	157
20-25	29	186
25-30	11	197

N=197

lc

$N = 197$

$\frac{N}{2} = \frac{197}{2} = \underline{\underline{98.5}}$

$\left[l = 10,\ f = 50,\ m = 52,\ N/2 = 98.5\right.$

$\left. c = 5 \right]$

$$\therefore Md = l + \left[\frac{N/2 - m}{f}\right] c$$

$$= 10 + \left(\frac{98.5 - 52}{50}\right) 5$$

$$= 10 + \left(\frac{46.5}{50}\right) 5$$

$$= 10 + \frac{46.5}{10}$$

$$= 10 + 4.65$$

$$\left(Md = \underline{\underline{14.65}}\right)$$

$$\Uparrow \qquad \Uparrow$$

(2) find the median for the following freq data ?

C.I	0-10	10-20	20-30	30-40	40-50	50-60
freq	15	20	25	32	18	10

S3
10/2/21
1-10 to
4-30

Sol:-

Sol:-

C.I	freq	Cumulative freq
0-10	15	15
10-20	20	35 m
20-30	25 f	60
30-40	32	92
40-50	18	110
50-60	10	120
	N = 120	

Ic l

$N = 120$

$N/2 = \frac{120}{2} = 60$

$$Md = l + \left(\frac{N/2 - m}{f}\right) c$$

$$= 20 + \left(\frac{60 - 35}{25}\right) 10$$

$$= 20 + \left(\frac{25}{25}\right) 10$$

$$\boxed{Md = 30}$$

find $\bar{x}$, & Md for the following freq data

C.I	0-5	5-10	10-15	15-20	20-25	25-30	30-35
freq	17	25	40	50	40	25	17

Sol:-

Sol:-

C.I	freq (f)	c.f	xi	fixi
0-5	17	17	2.5	42.5
5-10	25	42	7.5	187.5
10-15	40	82 m	12.5	500
15-20	50 f	132	17.5	875
20-25	40	172	22.5	900
25-30	25	197	27.5	687.5
30-35	17	214	32.5	552.5
	214			3745.0

L

Mean $= \bar{x} =$

$$= \frac{\Sigma fx}{N}$$

$$= \frac{3745}{214}$$

$$\boxed{\bar{x} = 17.5}$$

$$Md = l + \left(\frac{N/2 - m}{f}\right) c \qquad \frac{N}{2} = \frac{214}{2} = 107$$

Here $\left[\, l = 15,\ f = 50,\ m = 82,\ c = 5 \,\right]$

$$Md = 15 + \left(\frac{107 - 82}{50}\right) 5$$

$$= 15 + \left(\frac{25}{50}\right) 5$$

$$= 15 + 2.5 = 17.5$$

$$\therefore \left\{ \begin{array}{l} \bar{X} = 17.5 \\ Md = 17.5 \end{array} \right]$$

Mode :- [Model value]

$-3, 2, 8, 2, 1, 8, 3, 2, 11, 8, 17, 2, 4, 8$

$Mo = 2 \rightarrow$ Unimodel data

$Mo = 2, 8 \rightarrow$ Bimodel data

Def :- Most frequently repeated observation is known as Mode.

eg:- 7, 11, 14, -3, 0, 6, 4, 8, 9

$M_0 = \cancel{\phi}$ No mode

1. If data has one observation with highest freq then the data known as Unimodel data

2. If data has contains two or more observation with same number of repetition then the data is known as Multimodel data

③ If data do not have any repetation then the data is known as no mode.

Grouped data Mode:-

$$M_o = l + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right) c$$

Here

$\Delta_1 = f - f_{-1}$

$\Delta_2 = f - f_{+1}$

l. lower limit IC

eg:- find Mode for the given freq data

C.I	freq
0-2	35
2-4	47
4-6	52
6-8	29
8-10	11

Sol:-

C.I	freq	
0-2	35	
2-4	47	f_{-1}
$\boxed{4\text{-}6}$	$\boxed{52}$	f
6-8	29	f_{+1}
8-10	11	

L

Cal:-

$l = 4 \quad f = 52$

$f_{-1} = 47 \quad f_{+1} = 29$

$\Delta_1 = f - f_{-1}$

$= 52 - 47 = 5$

$\Delta_2 = f - f_{+1}$

$= 52 - 29 = 23$

Now

$$M_0 = l + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)c$$

$$= 4 + \left(\frac{5}{5+23}\right)2$$

$$= 4 + \frac{10}{28}$$

$$\boxed{M_0 = 4.35}$$

find M, Md & Mo for the given freq data

C.I	5.5 – 10.5	10.5 – 15.5	15.5 – 20.5	20.5 – 25.5	25.5 – 30.5
freq	47	52	80	71	39

Sol:-

C.I	Freq (f)	C.f	Midpoint (x_i)	$f.x_i$
5.5 – 10.5	47	47	8	376
10.5 – 15.5	52 f_{-1}	99 m	13	676
15.5 – 20.5	$\boxed{80}$ f	179	18	1440
20.5 – 25.5	71 f_{+1}	250	23	1633
25.5 – 30.5	39	289	28	1092
	289 $= N$			$5217 = \Sigma f_i x_i$

Md

Ic

$f_{+1} = 0$

$N = 289$

$\frac{N}{2} = \frac{289}{2}$

$= 144.5$

Now Mean $= \bar{x} = \frac{\Sigma fx}{N} = \frac{5217}{289} = 18.05$

$$Md = l + \left(\frac{N/2 - m}{f}\right) c$$

$$= 15.5 + \left(\frac{144.5 - 99}{80}\right) 5$$

$$= 15.5 + \left(\frac{45.5}{80}\right) 5$$

$$= 15.5 + 2.84$$

$$(Md = 18.3)$$

$$M_o = l + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right) c$$

$$= 15.5 + \left(\frac{28}{28+9}\right) 5$$

$$= 15.5 + 3.78$$

$$\boxed{M_o = 19.28}$$

$\Delta_1 = f - f_{-1}$

$= 80 - 52$

$= 28$

$\Delta_2 = f - f_{+1}$

$= 80 - 71$

$= 9$

Relation b/w Mean, Median & Mode

$$\left[Mo \approx 3Md - 2Mean \right]$$

M.C.T. :-

Mean $\Rightarrow \frac{\sum f_i x_i}{N}$ [all] → Best

Median $\Rightarrow l + \left(\frac{N/2 - m}{f}\right)c$ 50%

Mode $\Rightarrow l + \left(\frac{\Delta_1}{\Delta_2 + \Delta_1}\right)c$ → (Max freq)

Measures of Dispersion (or) Variabilities

(1) Range

(2) Quartile Deviation (Q.D)

(3) Mean Deviation (M.D)

(4) Standared deviation (S.d)

(5) Coefficient of Vaication (C.V)

$$
\begin{array}{c}
\underline{x} \\
x_1 \\
x_2 \\
x_3 \\
x_4 \\
\vdots \\
x_n
\end{array}
\quad \text{\underline{Within}}
$$

Range:-

$$\text{Range} = \text{Max} - \text{Min} \qquad [\text{Raw data}]$$
$$= GV - LV$$

$-3, 17, 21, 0, 11, 8, 12, 14, 2.6, 8$

max

Sol: Arrg :- $-3, 0, 2.6, 8, 8, 11, 12, 14, 17, 21$

Min

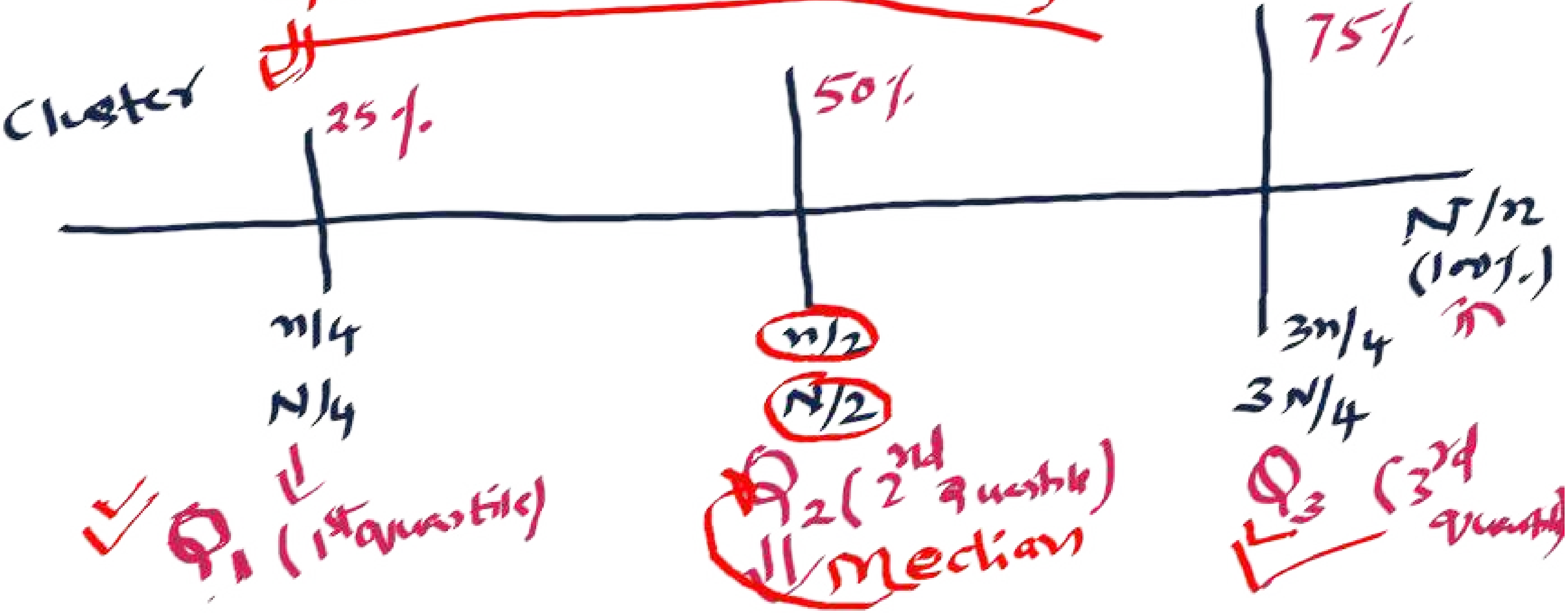
$$\text{Range} = \text{Max} - \text{Min}$$
$$= 21 - (-3) = 24$$

Coefficient of Range:

$$\left[\text{Coeff Range} = \frac{\text{Max} - \text{Min}}{\text{Max} + \text{Min}} \right]$$

$$\left(\text{Coeff Range} = \frac{21 - (-3)}{21 + (-3)} = \frac{24}{18} = 1.933 \right)$$

Quartile Devition [Q.D]



Mathematically

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

Q.D is also known as "Semi inter quartile Range"

Here $\Rightarrow$

$$\left[\begin{array}{l} Q_1 = n/4^{th} \text{ location | position} \\ Q_3 = 3n/4^{th} \text{ location | position} \end{array}\right] \Rightarrow \text{Rawdata}$$

Eg ①:- 5, 14, 22, 0, 6.8, 0.3, -1, 11, 17, 21

Sol: Arrange: [-1, 0, 0.3, 5, 6.8, 11, 14, 17, 21, 22]

Q_1 ... Q_3

$n = 10$

$Q_1 = \frac{n}{4}^{th} = \frac{10}{4} = 2.5 \simeq 3^{rd}$ observation $\Rightarrow$ Value of $\boxed{Q_1 = 0.3}$

↳ Location

$Q_3 = \frac{3n}{4} = 3\left(\frac{10}{4}\right) = 3(2.5)$

$= 7.5 \simeq 8^{th}$ observation $\Rightarrow$ Value of $\boxed{Q_3 = 17}$

↳ Location

$QD = \frac{Q_3 - Q_1}{2} = \frac{17 - (0.3)}{2} = 8.35$

Eg ②: find Q.D for the given data set?

−17, −20, −35, 40, 19, 26, 0.8, 11, 26, 35, 2, 1, 0.2

Sol:- Arrange

Q_1 ... Q_3

−35, −20, −17, 0.2, 0.8, 1, 2, 11, 19, 26, 26, 35, 40

3rd ob ... 10th obsev

$n = 13$

$\frac{n}{4}^{th} = \frac{13}{4} = 3.25^{th}$ position $\simeq 3^{rd}$ Location

$\frac{3n}{4}^{th} = 3(3.25) = 9.75$ ↳ position $\simeq 10^{th}$ Location

$$QD = \frac{Q_3 - Q_1}{2} = \frac{26-(-17)}{2}$$

$$= \frac{43}{2} = 21.5$$

(3) find Md & QD for the given data set?

75, 120, 140, 11, 26, 4, 9, 3, 1, 0, 14, 15, 10, 3, 6

Sol:- (Arrg): 0, 1, 3, 3, 4, 6, 9, 10, 11, 14, 15, 26, 75, 120, 140

(Q_1) 3; Q_2 10 (Md); Q_3 15

$n = 15$

$n/4 = \frac{15}{4} = 3.75^{th}$ position $\simeq 4^{th}$ obser

$\frac{3n}{4} = 3\left(\frac{15}{4}\right) = 3(3.75)$

$= 11.25^{th}$ posn $\simeq 11^{th}$ obser

[$Md = Q_2 = 10$

$Q_1 = 3$; $Q_3 = 15$

$QD = \frac{15-3}{2} = \frac{12}{2} = 6$]

find the Q.D for given freq data ?

16/2/21
10.10 to 11.10

C.I	0-5	5-10	10-15	15-20	20-25	25-30	30-35
freq	15	40	65	80	55	35	10

Sol:-

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

Here

$$Q_3 = l_3 + \left(\frac{\frac{3N}{4} - m_3}{f_3}\right) C$$

$$Q_1 = l_1 + \left(\frac{N/4 - m_1}{f_1}\right) C$$

C.I	freq	Cumulative freq
0-5	15	15
5-10	40	55 m_1
l_1 ← 10-15	65 f_1	120 → Q_1
15-20	80	200 m_3
l_3 ← 20-25	55 f_3	255 → Q_3
25-30	35	290
30-35	10	300
	$N = 300$	

$$N/4 = \frac{300}{4} = 75$$

$$\frac{3N}{4} = 3(75) = 225$$

$$\begin{cases} l_1 = 10, \quad f_1 = 65, \quad m_1 = 55 \\ l_3 = 20 \quad f_3 = 55 \quad m_3 = 20 \end{cases}$$

$$Q_1 = l_1 + \left(\frac{N/4 - m_1}{f_1} \right) c$$

$$= 10 + \left(\frac{75 - 55}{65} \right) 5$$

$$= 10 + \left(\frac{20}{65} \right) 5$$

$$= 10 + \frac{100}{65}$$

$$= 10 + 1.52$$

$$= 11.53$$

$\therefore Q_1 = 11.53$

$$Q_3 = l_3 + \left(\frac{\frac{3N}{4} - m_3}{f_3}\right) c$$

$$= 20 + \left(\frac{225 - 200}{55}\right) 5$$

$$= 20 + \left(\frac{25}{55}\right) 5$$

$$= 20 + 2.27$$

$$Q_3 = 22.7$$

$$\therefore QD = \frac{Q_3 - Q_1}{2}$$

$$= \frac{22.7 - 11.53}{2}$$

$$\boxed{QD = 5.36}$$

find Md (Q_2), QD for given freq data ?

C.I	0-10	10-20	20-30	30-40	40-50	50-60
freq	40	65	90	110	75	50

Sol:-

$$Md = Q_2 = l_2 + \left(\frac{N/2 - m}{f}\right)c$$

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

$$Q_1 = l_1 + \left(\frac{N/4 - m_1}{f_1}\right)c$$

$$Q_3 = l_3 + \left(\frac{3N/4 - m_3}{f_3}\right)c$$

	C.I	freq	Cumulative freq
	0-10	40	40
	10-20	65	105
Q_1	20-30	90	195
Q_2	30-40	110	305
Q_3	40-50	75	380
	50-60	50	430
		$N = 430$	

$$N/4 = \frac{430}{4} = 107.5$$

$$N/2 = \frac{430}{2} = 215$$

$$\frac{3N}{4} = 3(107.5) = 322.5$$

$l_1 = 20$	$f_1 = 90$	$m_1 = 105$
$l_2 = 30$	$f_2 = 110$	$m_2 = 195$
$l_3 = 40$	$f_3 = 75$	$m_3 = 305$

$$Q_1 = l_1 + \left(\frac{N/4 - m_1}{f_1}\right)c$$

$$= 20 + \left(\frac{107.5 - 105}{90}\right)10$$

$$= 20 + \left(\frac{2.5}{90}\right)10$$

$$= 20 + 0.27$$

$$Q_1 = 20.27$$

$$Q_2 = l_2 + \left(\frac{N/2 - m_2}{f_2}\right)c$$

$$= 30 + \left(\frac{215 - 195}{110}\right)10$$

$$= 30 + \left(\frac{20}{110}\right)10$$

$$= 30 + 1.818$$

$$Q_2 = 31.818$$

$$Q_3 = L_3 + \left(\frac{3N/4 - m_3}{f_3}\right) c$$

$$= 40 + \left(\frac{322.5 - 305}{75}\right) 10$$

$$= 40 + \left(\frac{17.5}{75}\right) 10$$

$$= 40 + 2.333$$

$$Q_3 = 42.333$$

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

$$QD = \frac{42.333 - 20.27}{2}$$

$$\boxed{QD = \frac{22.06}{2} = 11.03}$$

Standard deviation (s.d)

$$\sqrt{\underset{(\sigma^2)}{\text{Variance}}} = \underset{(\sigma)}{s.d}$$

dataset $\rightarrow$ $x_1, x_2, x_3 \ldots x_n \Longrightarrow$ fixed value $(\bar{x})$

$$\Rightarrow (x_1 - \bar{x}_1)^2 + (x_2 - \bar{x}_2)^2 + \ldots + (x_n - \bar{x})^2 = n\sigma^2)$$

Variance : " difference or deviations within the data from its mean"

S.d → Minimizing the differences within the observation → It leads to measure the consistency / Uniformity / Regularity

Sometimes can also read as Root Mean Square error (RMS) eg :-

[Mean of const data set is const itself]

X_i (Const)	Y_i (Variable data)
3	1
3	2
3	3
3	4
3	5
3	6
3	7
$21 \Rightarrow \bar{x} = \frac{21}{7} = 3$	28

$$\bar{x} = \frac{28}{7} = 4$$

$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})$
$(3-3)^2$	$(1-4)^2$
$(3-3)^2$	$(2-4)^2$
⋮	⋮ +ve
$(3-3)^2$	$(7-4)^2$
0	[illegible]

Note:-

→ Var[const data] = 0 [zero]

→ var [variable data] = +ve [positive]

→ [Variance Never be negative]

(Online Classes (6,7)) SU 16/2/21 2.20 to 4.30

Mathematical formulae:

Raw data
$$\begin{cases} \sigma_x^2 = \frac{1}{n}\sum(x_i-\bar{x})^2 \\ \sigma_x^2 = \frac{1}{n}\left(\sum x_i^2\right) - (\bar{x})^2 \end{cases}$$

Sum of the squares of deviation from mean

freq distn:
$$\begin{cases} \sigma_x^2 = \frac{1}{N}\sum f_i(x_i-\bar{x})^2 \\ \sigma_x^2 = \frac{1}{N}\sum f_i x_i^2 - \left(\frac{1}{N}\sum f_i x_i\right)^2 \end{cases}$$

find the variance and s.d for the given observations ?

$-3, 9, 0, 11, 16, 5, 2, 6, 1$

Sol:-

x_i	$(x_i - \bar{x})^2$
-3	$(-3-5.22)^2 = 67.5$
0	$(0-5.22)^2 = 27.3$
1	$(1-5.22)^2 = 17.8$
2	$(2-5.22)^2 = 10.3$
5	$(5-5.22)^2 = 0.048$
6	$(6-5.22)^2 = 0.608$
9	$(9-5.22)^2 = 14.2$
11	$(11-5.22)^2 = 33.40$
16	$(16-5.22)^2 = 116.2$
47	287.496

Now

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{47}{9}$$

$$= 5.22$$

$$\text{Variance} = \sigma^2 = \frac{\Sigma(x_i - \bar{x})^2}{n} = \frac{287.496}{9} = 31.9$$

$$S.d_{(\sigma)} = \sqrt{\text{Variance}} = \sqrt{31.9} = 5.65$$

eg②:- find Q.D & S.d for the given data set?

2.8, 11.3, 0.6, 0.9, 0, −2, −0.8, 11.5, 6.9, −3.2, 1, 9

Sol

x_i	$(x_i-\bar{x})^2$	x_i	$(x_i-\bar{x})^2$
-3.2	$(-3.2-3.16)^2$	9	$(9-3.16)^2$
-2	$(-2-3.16)^2$	11.3	$(11.3-3.16)^2$
-0.8	$(-0.8-3.16)^2$	11.5	$(11.5-3.16)^2 =$
0	$(0-3.16)^2$		
0.6	$(0.6-3.16)^2$		
0.9	$(0.9-3.16)^2$		
1	$(1-3.16)^2$		
2.8	$(2.8-3.16)^2$		
6.9	$(6.9-3.16)^2$		

$$\Sigma x_i = 38$$

$$\bar{x} = \frac{\Sigma x_i}{12}$$

$$= \frac{38}{12} =$$

$$(\bar{x} = 3.16)$$

$$[40.44 + 26.62 + 15.68 + 9.98 + 6.55 + 5.1 + 4.64 + 0.129 + 13.98 + 34.1 + 66.25 + 69.55]$$

$$\Sigma(x_i-\bar{x})^2 = 293.01$$

$$\therefore \text{Variance} = \sigma_x^2 = \frac{\Sigma(x_i-\bar{x})^2}{n} = \frac{293.01}{12} =$$

$$= 24.4$$

$$\boxed{S.d = \sigma = \sqrt{24.4} = 4.94}$$

for Q.D:

Here $n = 12$

$Q_1 = \frac{n^{th}}{4} = \frac{12}{4} = 3^{rd}$ Location $\Rightarrow$ [$Q_1 = -0.8$ 3^{rd} observation]

$Q_3 = \frac{3n}{4} = 3(3) = 9^{th}$ Location $\Rightarrow$ [$Q_3 = 6.9$ 9^{th} observat]

$\therefore QD = \frac{Q_3 - Q_1}{2} = \frac{6.9 - (-0.8)}{2} = 3.85$

find S.D for the given freq data :

C.I	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
freq	25	46	54	70	80	55	29	11

Sol:-

$$\sigma_x^2 = \frac{1}{N}\Sigma f_i x_i^2 - \left(\frac{1}{N}\Sigma f_i x_i\right)^2$$

$$S.d = \sqrt{\sigma^2}$$

Sol:-

C.I	freq	Mid point	f.x	f.x^2
0-5	25	2.5	62.5	156.25
5-10	46	7.5	345	2587.5
10-15	54	12.5	675	8437.5
15-20	70	17.5	1225	21437.5
20-25	80	22.5	1800	40500
25-30	55	27.5	1512.5	41593.75
30-35	29	32.5	942.5	30631.25
35-40	11	37.5	412.5	15468.75
	370		6975	160812.5

Totals

$N = 370$

$\Sigma fx = 6975$

$\Sigma fx^2 = 160812.5$

$$\bar{X} = \frac{\Sigma fx}{N}$$

$$= \frac{6975}{370} = 18.85$$

Now $\bar{x} = 18.85$

$$\sigma_x^2 = \frac{1}{N}\Sigma f_i x_i^2 - (\bar{x})^2$$

$$= \left(\frac{160812.5}{370}\right) - (18.85)^2$$

$$= 434.62 - 355.32$$

$$\sigma_x^2 = 79.305$$

$\therefore$ s.d $=$

$$= \sqrt{\sigma^2}$$

$$= \sqrt{79.305}$$

$$(\sigma = 8.905)$$

Find sd & variance for the freq data

C.I	freq
0 – 10	25
10 – 20	50
20 – 30	65
30 – 40	70
40 – 50	35
50 – 60	15

Sol:-

$$\sigma_x^2 = \frac{1}{N} \Sigma f_i x_i^2 - (\bar{X})^2$$

$$\bar{X} = \frac{\Sigma f x}{N}$$

C.I	freq	mid pon	fx	fx^2
0–10	25	5	125	625
10–20	50	15	750	11250
20–30	65	25	1625	40625
30–40	70	35	2450	85750
40–50	35	45	1575	70875
50–60	15	55	825	45375
	260		7350	254500

Totals

$N = 260$

$\Sigma fx = 7350$

$\Sigma fx^2 = 254500$

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

$= \frac{7350}{260}$

$(\bar{x} = 28.26)$

$$\therefore \sigma_x^2 = \frac{1}{N} \Sigma f_i x_i^2 - (\bar{x})^2$$

$$= \frac{254500}{260} - (28.26)^2$$

$$\sigma_x^2 = 978.84 - 798.62$$

$$\sigma_x^2 = 180.22$$

$$\left[\sigma = s.d = \sqrt{180.22} = 13.42\right]$$

(8&9 class)
SO
17/2/21
1:10–3:20 PM

Coefficient of Variation [C.V]

$$C.V = \frac{S.d}{\text{Mean}} \times 100$$

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

Note:-

⇒ [Lesser C.V is more consistent or more uniform]

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

find C.V for the following, identify which is more consistent ?

GA (x_i)	GB (y_i)
17	120
26	115
35	145
47	165
11	200
9	175
145	

Sol:- (GA calculations)

x_i	17	26	35	47	11	9
$(x_i-\bar{x})^2$	51.36	3.36	117.50	521.66	173.18	229.82

$= 1096.8$

$$\bar{X}_{GA} = \frac{\Sigma x_i}{n} = \frac{145}{6} = 24.16$$

$$\Sigma(x_i - \bar{x})^2 = 1096.8$$

$$\sigma_x^2 = \frac{\Sigma(x_i - \bar{x})^2}{n} = \frac{1096.8}{6} = 182.8$$

$$\sigma = \sqrt{182.8} = 13.52$$

$$\left(CV_A = \frac{\sigma_x}{\bar{x}} \times 100 = \frac{13.56}{24.16} \times 100 = 55.9 \right) //$$

GB (Yi)

X_i	$(Y_i - \bar{Y})^2$
120	$(120 - 156.66)^2 = 1343.95$
115	$(115 - 156.66)^2 = 1735.55$
145	$(145 - 156.66)^2 = 135.95$
165	$(165 - 156.66)^2 = 69.55$
200	$(200 - 156.66)^2 = 1878.35$
195	$(195 - 156.66)^2 = 1469.95$
940	$\Sigma(Y_i - \bar{Y})^2 = 6633.3$

$$\bar{Y} = \frac{\Sigma Y_i}{n} = \frac{940}{6} = 156.66$$

$$\sigma_y^2 = \frac{\Sigma(Y_i - \bar{Y})^2}{n}$$

$$= \frac{6633.3}{6}$$

$$= 1105.5$$

$$\sigma = 33.2$$

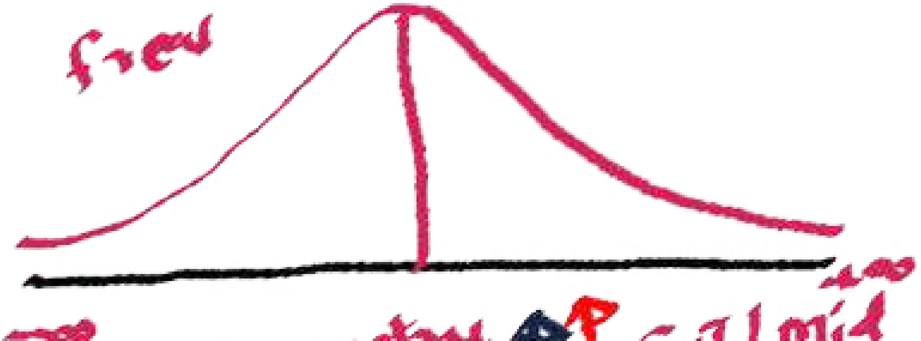
$$C.V_{GB} = \frac{\sigma_y}{\bar{y}} \times 100 = \frac{33.2}{156.66} \times 100$$

$$\boxed{\begin{aligned} CV_B &= 21.2 \;// \\ CV_A &= 55.9 \end{aligned}}$$

$\therefore \left[\begin{array}{l} CV_B < CV_A \\ \therefore G_B \text{ is more consistant //} \end{array} \right]$

Skewness :- X Asymmetry

Lack of symmetry.



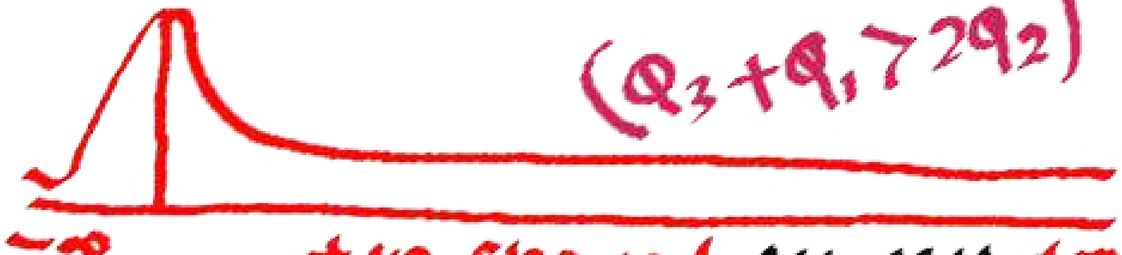
Symmetry ⇒ (c-3) | mid

(1) $M = Md = Mo$

$[Q_3 + Q_1 = 2Q_2]$

$(Q_3 + Q_1 < 2Q_2)$

$-$ve skewed $(M < Md < Mo)$



$(Q_3 + Q_1 > 2Q_2)$

$+$ve skewed $M > Md > Mo$

Mathematical Formulae:

(1) Karl pearsons coefficient of Skewness (Skp)

(2) Bowley's coefficient of Skewness (SkB)

(1) Karl pearson's coefficient of Skewness:-

$$Skp = \frac{M - Mo}{\sigma} = \frac{3(M - Md)}{\sigma}$$

$$\boxed{-3 \leq Skp \leq +3}$$

$$\begin{bmatrix} -3 \le Sk_P < 0 & \rightarrow & \text{-ve skewed} \\ Sk_P = 0 & \rightarrow & \text{Symmetry} \\ 0 < Sk_P \le +3 & \rightarrow & \text{+ve skewed} \end{bmatrix}$$

(2) Bowley's coefficient of skewness (Sk_B):

$$\boxed{Sk_B = \frac{(Q_3 + Q_1) - 2Q_2}{Q_3 - Q_1}}$$

75% 25%

$$\boxed{-1 \le Sk_B \le +1}$$

find Skp and SkB for the given freq data?

C.I	0-10	10-20	20-30	30-40	40-50
freq	15	26	35	14	9

Sol:-

$$Sk_P = \frac{M - M_o}{\sigma} = \frac{3(M - Md)}{\sigma}$$

$$Sk_B = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

$$Me = \bar{x} = \frac{\Sigma fx}{N}$$

$$Md = l + \left(\frac{N/2 - m}{f}\right)c$$

$$Q_1 = l + \left(\frac{N/4 - m}{f}\right)c$$

$$Q_3 = l + \left(\frac{3N/4 - m}{f}\right)c$$

	C.I	freq	c.f	Mid point (x_i)	$f_i x_i$	$f_i x_i^2$
	0–10	15	15	5	75	375
Q_1 ←	10–20	26	41	15	390	5850
Q_2 / Q_3	20–30	35	76	25	875	21875
	30–40	14	90	35	490	17150
	40–50	9	99	45	405	18225
		$N = 99$			2235	63475

$N = 99$

$\frac{N}{4} = 24.7$

$\frac{N}{2} = 49.5$

$\frac{3N}{4} = $

$= 74.25$

Calculations:-

$$\text{Mean} = \bar{x} = \frac{\Sigma f x}{N} = \frac{2235}{99} = 22.57$$

$$\sigma^2 = \frac{1}{N}\Sigma f x^2 - (\bar{x})^2 = \frac{63475}{99} - (22.57)^2$$

$$= 641.16 - 509.40$$

$$\sigma^2 = 131.76$$

$$\sigma = \sqrt{131.76} = 11.47$$

$$Q_1 = 4 + \left(\frac{N/4 - m_1}{f_1}\right) c$$

$$= 10 + \left(\frac{24.7 - 15}{26}\right) 10$$

$$= 13.73$$

$$Q_2 = 12 + \left(\frac{N/2 - m_2}{f}\right) c$$

$$= 20 + \left(\frac{49.5 - 41}{35}\right) 10 = 22.42$$

$$Q_3 = l_3 + \left(\frac{3N/4 - m_3}{f_3}\right) c$$

$$= 20 + \left(\frac{74.25 - 41}{35}\right) 10$$

$$Q_3 = 29.5$$

$\bar{x} = 22.57$	$Q_1 = 13.73$	$Q_3 = 29.5$
$\sigma = 11.47$	$Q_2 = 22.42$	

$$Skp = \frac{3(m - md)}{\sigma} = \frac{3(22.57 - 22.42)}{11.47}$$

$$= 0.03$$

$$\therefore \boxed{Skp = 0.03}$$

Inference / comment = Given freq data is +ve skewed

$$Sk_B = \left(\frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}\right)$$

$$= \frac{29.5 + 13.73 - 2(22.42)}{29.5 - 13.73}$$

$$\boxed{Sk_B = -0.10}$$

Comment:- Given freq data is <u>-ve skewed</u>

find (Skp & SkB) for the following freq data

C.I	0-5	5-10	10-15	15-20	20-25	25-30	30-35
freq	55	83	95	110	95	83	55

Sol:-

$$\left[\begin{aligned} Sk_p &= \frac{3(M - Md)}{\sigma} \\ Sk_B &= \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1} \end{aligned}\right]$$

Table

C.I	freq	c.f	Mid point (x)	fx	fx^2	
0-5	55	55	2.5	137.5	343.75	
5-10	83	138	7.5	622.5	4668.75	
10-15	95	233	12.5	1187.5	14843.75	Q_1
15-20	110	343	17.5	1925	33687.5	Q_2
20-25	95	438	22.5	2137.5	48093.75	Q_3
25-30	83	521	27.5	2282.5	62768.75	
30-35	55	576	32.5	1787.5	58093.75	
	N=576			10080	222500.00	

Calculations:-

$N = 576$

$\frac{N}{2} = 288$

$$\text{Mean} = \bar{x} = \frac{\Sigma fx}{N} = \frac{10080}{576} = 17.5$$

$$Md = l + \left(\frac{N/2 - m}{f}\right)c = 15 + \left(\frac{288 - 233}{110}\right)5$$

$$= 17.5$$

$$\sigma^2 = \frac{1}{N}\Sigma fx^2 - (\bar{x})^2 = \frac{222500}{576} - (17.5)^2$$

$$= 80.03 \Rightarrow \boxed{\sigma = 8.94}$$

$$N = 576$$

$$\frac{N}{4} = \frac{576}{4} = 144$$

$$Q_1 = l + \left(\frac{N/4 - m}{f}\right) c$$

$$= 10 + \left(\frac{144 - 138}{95}\right) 5$$

$$= 10.31$$

$$\frac{3N}{4} = 3(144) = 432$$

$$Q_3 = l + \left(\frac{3N/4 - m}{f}\right) c$$

$$= 20 + \left(\frac{432 - 343}{95}\right) 5 = 24.68$$

$\bar{x} = 17.5$; $Q_2 = 17.5$; $\sigma = 8.94$

$Q_3 = 24.68$ $Q_1 = 10.31$

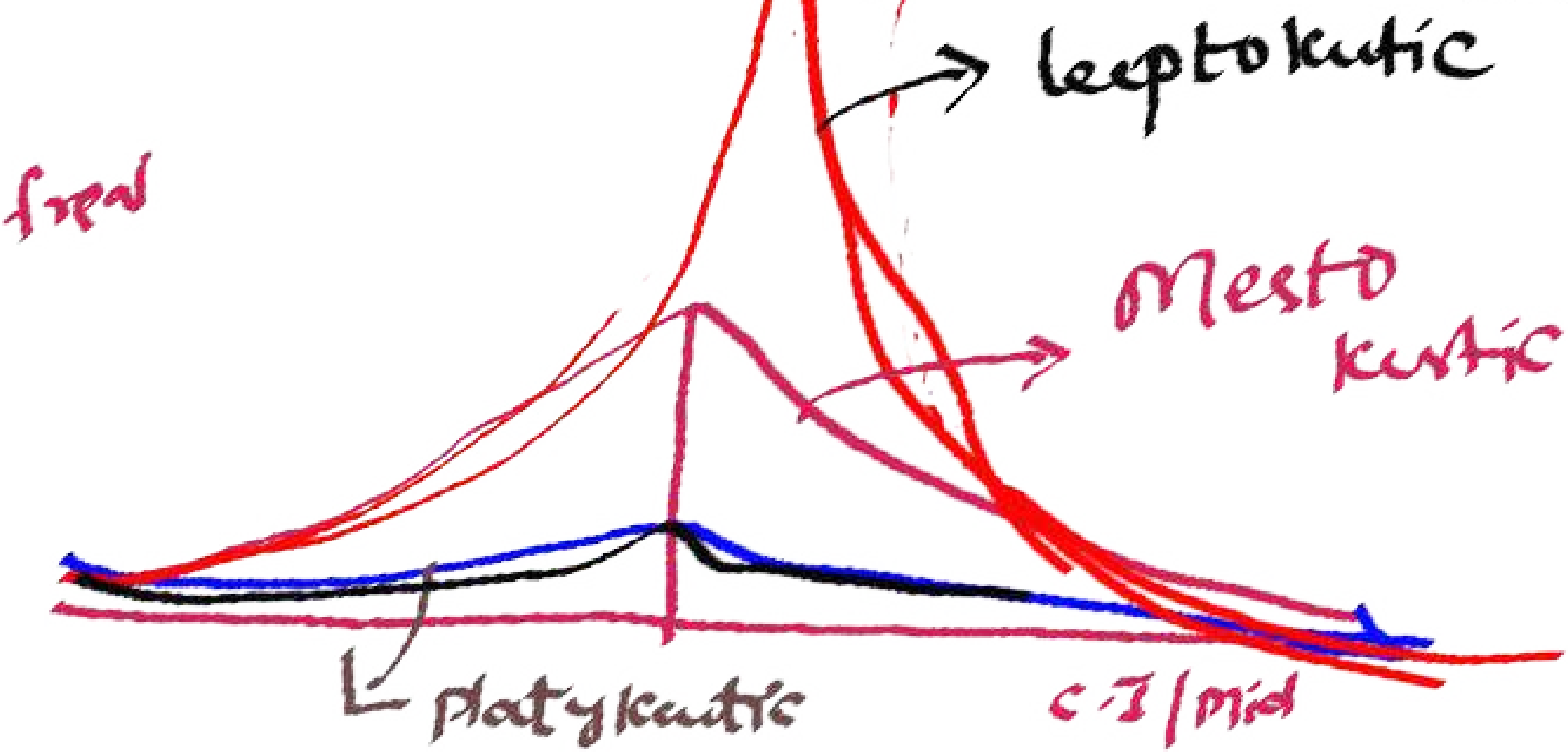
$$Sk_P = \frac{3(M - Md)}{\sigma} = \frac{3(17.5 - 17.5)}{8.94} = 0$$

Given data is symmetry

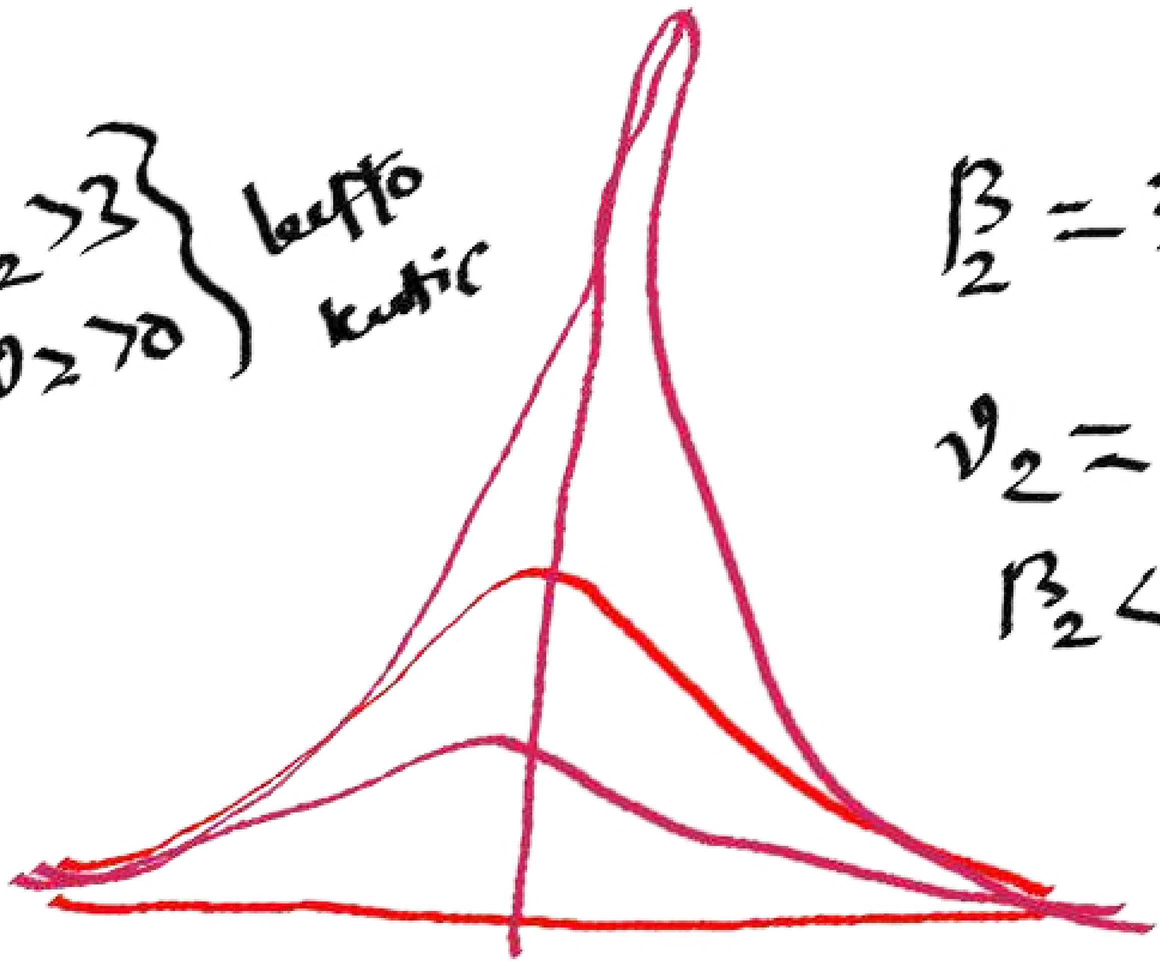
$$Sk_B = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1} = \frac{24.68 + 10.31 - 2(17.5)}{24.68 - 10.31}$$

$= 0$ //

Kurtosis:- Bulkginess (or) Peakedness.



$\left.\begin{array}{l}\beta_2 > 3\\ \vartheta_2 > 0\end{array}\right\}$ lepto kurtic



$\left.\begin{array}{l}\beta_2 = 3\\ \vartheta_2 = 0\end{array}\right\}$ Mesto kurtic

$\beta_2 < 3$; $\vartheta_2 < 0$

↳ platy kurtic

10ट11

23/2/21

2:30 to 4:30

Curve fitting:-

→ S.t line $y = ax + b$

→ parabola $y = ax^2 + bx + c$

→ Growth curve

- E.C $[y = ae^{bx}]$
- G.C $[y = ax^b]$
- P.C $[y = ab^x]$

$\boxed{a, b}$

Parameters

Fitting St. line : $[y = ax + b]$

General form st line $y = ax + b + \epsilon$ ↳ Error term

<u>Technique</u> :- Method of least squares Approximation (LSA) → method

Using L.S.A method we can minimize the error sum of squares, such way that, we can able to compute the best values of the parameters (a, b) : [using given Normal equations]

Now, let us consider the S.t line

$$y = ax + b \rightarrow (1)$$

multiply eq (1) both sides with "Σ"

$$\Sigma y = \Sigma(ax + b)$$

$$\Sigma y = \Sigma ax + nb$$

$$\boxed{\Sigma y = a\Sigma x + nb} \rightarrow (2)$$

2
2
2
2
2
2

12

6(2) = 12

Again multiply eq ① with 'Σx' on both sides

$$\Sigma x \cdot y = \Sigma x (ax + b)$$

$$= \Sigma a \cdot x^2 + \Sigma x \cdot b$$

$$\boxed{\Sigma x \cdot y = a \Sigma x^2 + b \Sigma x} \longrightarrow ③$$

Here eq ② & eq ③ are known as "Normal equations"

By solving eq ② & eq ③ we get the values of " a, b "

Now the fitted equation is

estimated value

$$\boxed{\hat{y} = ax + b}$$

fit a second degree parabola:-

$\{Y = ax^2 + bx + c\}$

The General form of eqn is

$$Y = ax^2 + bx + c \rightarrow (1)$$

By the Method of L.S.A, we can write normal eqnts for form (1).

Multiply eq (1) with 'Σ' on both sides

$$\Sigma Y = \Sigma(ax^2 + bx + c)$$

$$\Sigma y = \Sigma a x^2 + \Sigma b x + nc$$

$$\boxed{\Sigma y = a \Sigma x^2 + b \Sigma x + nc} \rightarrow (2)$$

Again multiply eq (1) with 'Σx' on both sides

then

$$\Sigma x \cdot y = \Sigma x (a x^2 + b x + c)$$

$$= \Sigma a x^3 + \Sigma b x^2 + \Sigma c x$$

$$\boxed{\Sigma x y = a \Sigma x^3 + b \Sigma x^2 + c \Sigma x} \rightarrow (3)$$

Again multiply eq 1 with "Σx^2" on both sides

$$\Sigma x^2 y = \Sigma x^2 (ax^2 + bx + c)$$

$$= \Sigma a x^4 + \Sigma b x^3 + \Sigma c x^2$$

$$\boxed{\Sigma x^2 y = a \Sigma x^4 + b \Sigma x^3 + c \Sigma x^2} \rightarrow 4$$

By solving eq 2, eq 3 & eq 4 we get "a, b & c"

$\therefore$ The fitted equation is

$$\boxed{\hat{y} = ax^2 + bx + c}$$

Problem:-

fit a St-line for the following data

x	1	3	5	7	9	11
y	45	70	110	150	180	200

estimate $y(10)$?

output

independent

Sol:- $y = ax + b \rightarrow (1)$

Normal eq's are $\Sigma y = a\Sigma x + nb \rightarrow (2)$

$$\Sigma xy = a\Sigma x^2 + b\Sigma x \rightarrow (3)$$

Table:-

x	y	$x \cdot y$	x^2
1	45	45	1
3	70	210	9
5	110	550	25
7	150	1050	49
9	180	1620	81
11	200	2200	121
36	755	5675	286

We know that normal eq's are

$$\Sigma y = a\Sigma x + nb \rightarrow ①$$

$$\Sigma xy = a\Sigma x^2 + b\Sigma x \rightarrow ②$$

Sub values in eq ① & ②

$$755 = 36a + 6b$$

$$5675 = 286a + 36b$$

$a = 16.35$, $b = 27.69$

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 16.35 \\ 27.69 \end{bmatrix}$$

$\therefore$ The fitted equation is

$$\boxed{\hat{y} = (16.35)x + 27.69}$$

and $y(10)$ is

$$\hat{y} = (16.35).10 + 27.69$$

$$= \underline{\underline{191.19}}$$

(2) fit a equation of form $y = a + bx$ for the following observations?

x	11	14	16	18	20
y	50	110	200	350	500

Sol:- Equation of form $y = a + bx \rightarrow (1)$

Normal equations of (1) are

$$\Sigma y = na + b\Sigma x \rightarrow (2)$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \rightarrow (3)$$

Table:

x	y	xy	x^2
11	50	550	121
14	110	1540	196
16	200	3200	256
18	350	6300	324
20	500	10,000	400
79	1210	21590	1297

Normal equation

$$\Sigma y = na + b\Sigma x$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2$$

Sub all values in N.Eq

$$1210 = 5a + 79b$$

$$21590 = 79a + 1297b$$

$a = -558.4$, $b = 50.66$

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -558.4 \\ 50.66 \end{bmatrix}$$

Now the fitted equation is

$$\boxed{\hat{y} = (-558.4) + (50.66)x}$$

(3) Fit a equation of form $y = ax^2 + bx + c$ for the following data:

x	1	3	5	7	9	11
y	15	30	45	60	75	90

Sol:-

$$y = ax^2 + bx + c \rightarrow (1)$$

Normal equations are

$$\Sigma y = a\Sigma x^2 + b\Sigma x + nc \rightarrow (2)$$

$$\Sigma xy = a\Sigma x^3 + b\Sigma x^2 + c\Sigma x \rightarrow (3)$$

$$\Sigma x^2 y = a\Sigma x^4 + b\Sigma x^3 + c\Sigma x^2 \rightarrow (4)$$

Table:

x	y	xy	x^2	x^2y	x^3	x^4
1	15	15	1	15	1	1
3	30	90	9	270	27	81
5	45	225	25	1125	125	625
7	60	420	49	2940	343	2401
9	75	675	81	6075	729	6561
11	90	990	121	10890	1331	14641
36	315	2415	286	21315	2556	24310

$\Sigma x = 36$

$\Sigma y = 315$

$\Sigma xy = 2415$

$\Sigma x^2 = 286$

$\Sigma x^2 y = 21315$

$\Sigma x^3 = 2556$

$\Sigma x^4 = 24310$

Normal eqs are

$\Sigma y = a\Sigma x^2 + b\Sigma x + nc$

$315 = 286a + 36b + 6c \rightarrow$ (2)

$\Sigma xy = a\Sigma x^3 + b\Sigma x^2 + c\Sigma x$

$2415 = 2556a + 286b + 36c \rightarrow$ (3)

$\Sigma x^2 y = a\Sigma x^4 + b\Sigma x^3 + c\Sigma x^2$

$21315 = 24310a + 2556b + 286 \rightarrow$ (4)

By solving eq ②, eq ③ & eq ④ we get
a, b & c

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 7.5 \\ 7.5 \end{bmatrix}$$

∴ fitted eq is

$$\left[\hat{y} = 0.x^2 + 7.5x + 7.5\right]$$

(4) fit second degree parabola for the given data

x	8	11	15	16	20
y	45	50	75	100	110

Sol:- $y = ax^2 + bx + c \rightarrow (1)$

Normal eq

$$\Sigma y = a\Sigma x^2 + b\Sigma x + nc \rightarrow (2)$$

$$\Sigma xy = a\Sigma x^3 + b\Sigma x^2 + c\Sigma x \rightarrow (3)$$

$$\Sigma x^2 y = a\Sigma x^4 + b\Sigma x^3 + c\Sigma x^2 \rightarrow (4)$$

Table:

x	y	xy	x^2	x^2y	x^3	x^4
8	45	360	64	2880	512	4096
11	50	550	121	6050	1331	14641
15	75	1125	225	16875	3375	50625
16	100	1600	256	25600	4096	65536
20	110	2200	400	44000	8000	160,000
70	380	5835	1066	95405	17314	294898

Normal eqs

$$\Sigma y = a\Sigma x^2 + b\Sigma x + nc$$

$$380 = 1066a + 70b + 5c \rightarrow (2)$$

$$\Sigma xy = a\Sigma x^3 + b\Sigma x^2 + c\Sigma x$$

$$5835 = 17314a + 1066b + 70c \rightarrow (3)$$

$$\Sigma x^2 y = a\Sigma x^4 + b\Sigma x^3 + c\Sigma x^2$$

$$95405 = 294898a + 17314b + 1066c \rightarrow (4)$$

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0.06 \\ 4.22 \\ 3.35 \end{bmatrix}$$

$\therefore$ fitted eq is

$$\boxed{\hat{y} = 0.06x^2 + 4.22x + 3.35}$$

Method – III [Growth Curves]

Fit a exponential curve $y = ae^{bx}$

Procedure:- Equation of form $y = ae^{bx} \rightarrow (1)$

Take $\log_e$ on both sides of eq (1)

$$\log_e y = \log_e (ae^{bx})$$

$$Y = \log_e a + \log_e e^{bx}$$

$$\boxed{Y = A + bx} \rightarrow (2)$$

Method – III [Growth Curves]

Fit a exponential curve $y = ae^{bx}$

Procedure:- Equation of form $y = ae^{bx} \rightarrow (1)$

Take $\log_e$ on both sides of eq (1)

$$\log_e y = \log_e (ae^{bx})$$

$$Y = \log_e a + \log_e e^{bx}$$

$$\boxed{Y = A + bx} \rightarrow (2)$$

Reduced equation form is

$$Y = A + bX \rightarrow (2)$$

Prepare normal equations for eqn (2) using LSA, are

$$\Sigma Y = nA + b\Sigma X \rightarrow (3)$$

$$\Sigma X \cdot Y = A\Sigma X + b\Sigma X^2 \rightarrow (4)$$

By solving eqn (3) & (4) we get "A, b"

Now $\left[a = e^{A}, \quad b = b \right]$

$\therefore$ fitted equation is

$$\left[\hat{y} = a e^{bx} \right]$$

Equation of form $Y = ab^x$

<u>Procedure:-</u> Given form $Y = ab^x \rightarrow (1)$

Take $\log_e$ on both sides

$$\log_e Y = \log_e (ab^x)$$

$$\log_e y = \log_e a + \log_e b^x$$

$$\boxed{Y = A + x \cdot B} \rightarrow (2)$$

$$\left[\begin{array}{l} Y = \log y \\ A = \log a \\ B = \log b \end{array}\right]$$

Reduced form is

$$Y = A + xB \rightarrow (2)$$

Normal equations are of eq (2)

$$\Sigma Y = nA + B\Sigma x \rightarrow (3)$$

$$\Sigma x \cdot Y = A\Sigma x + B\Sigma x^2 \rightarrow (4)$$

By solving eq (3) & (4) we get "A & B"

Now $a = e^{A}, \quad b = e^{B}$

$\therefore$ The fitted equation is

$$\boxed{\hat{y} = a b^{x}}$$

Equation of form $y = ax^b$

Procedure:. Given form $y = ax^b \rightarrow (1)$

Take $\log_e$ on both sides

$$\log_e y = \log_e (ax^b)$$

$$\log_e y = \log_e a + \log_e x^b$$

$$\boxed{Y = A + bX} \rightarrow (2)$$

Here
$Y = \log y$
$A = \log a$
$X = \log x$

Reducible form is

$$Y = A + bX \rightarrow (2)$$

Normal equations

$$\Sigma Y = nA + b\Sigma x \rightarrow (3)$$

$$\Sigma xY = A\Sigma x + b\Sigma x^2 \rightarrow (4)$$

By solving (3) & (4) we get "A, b"

Now, $\left[a = e^{A}, \quad b = b \right]$

finally the fitted equation is

$$\boxed{\hat{y} = a\,x^{b}}$$

fit a exponential curve $y = ae^{bx}$ 24/2/21

x	5	20	35	70	80
y	115	240	350	700	900

$$\log y = \log a + bx$$

$$Y = A + bx$$

$$\boxed{\begin{aligned} \Sigma Y &= nA + b\Sigma x \\ \Sigma xY &= A\Sigma x + b\Sigma x^2 \end{aligned}}$$

x	y	$Y = \log y$ $= \ln$	xY	x^2
5	115	4.7449	23.7	25
20	240	5.48	109.6	400
35	350	5.85	204.7	1225
70	780	6.65	465.5	4900
80	900	6.80	544	6400
210		29.52	1347.55	12950

Totals

$n = 5$

$\Sigma x = 210$

$\Sigma Y = 29.52$

$\Sigma x^2 = 12950$

$\Sigma xY = 1347.5$

The Normal eq's

$$\Sigma Y = nA + b\Sigma x \rightarrow (3)$$

$$\Sigma xY = A\Sigma x + b\Sigma x^2 \rightarrow (4)$$

Sub all value is normal eq's

$$29.52 = 5A + 210b$$

$$1347.5 = 210A + 12950b$$

$A = 4.80$

$b = 0.02$

$a = e^A = 121.51$

$\therefore$ The fitted eq is

$$\boxed{y = (121.51)e^{(0.02)x}}$$

fit the equation of form $y = a x^b$

x	1	9	11	17	25
y	50	67	85	110	172

Sol:
$$y = a x^b$$
$$\log_e y = \log a + b \log x$$
$$Y = A + bX$$

$$\Sigma Y = nA + b\Sigma X$$
$$\Sigma XY = A\Sigma X + b\Sigma X^2$$

Table:

x	y	$X = \log_e x$	$Y = \log_e y$	X^2	$X.Y$
1	50	0	3.9	0	0
9	67	2.1	4.2	4.41	8.82
11	85	2.3	4.4	5.29	10.12
17	110	2.8	4.7	7.84	13.16
25	172	3.2	5.1	10.24	16.32
		10.47	22.31	27.78	48.42

Totals:

$n = 5$, $\Sigma x = 10.47$, $\Sigma y = 22.31$,

$\Sigma x^2 = 27.78$, $\Sigma xy = 48.42$

Normal eq's $\Sigma y = nA + b\Sigma x$

$22.31 = 5A + 10.47b \rightarrow (3)$

$\Sigma xy = A\Sigma x + b\Sigma x^2$

$48.42 = 10.47A + 27.78b \rightarrow (4)$

$A = 3.8$, $b = 0.29$

$A = 3.8, \quad b = 0.29$

$a = e^A = 44.7$

$b = 0.29$

fitted eq is $\boxed{\hat{y} = (44.7)\, x^{0.29}}$

fit the equation form $y = ab^x$

x	1.7	3.5	9.6	11.5	15
y	20.6	41.3	54.1	65.6	90.5

Sol:- $y = ab^x$

$$\log y = \log a + x \log b$$

$$Y = A + xB$$

$$\Sigma Y = nA + B\Sigma x \rightarrow (2)$$

$$\Sigma xY = A\Sigma x + B\Sigma x^2 \rightarrow (3)$$

Table :-

x	y	$Y = \log_e y$	xY	x^2
1.7	20.6	3.02	5.13	2.89
3.5	41.3	3.72	13.06	12.25
9.6	54.1	3.99	38.30	92.16
11.5	65.6	4.18	48.07	132.25
15	90.5	4.50	67.5	225
41.3		19.41	172.02	464.55

Totals

Totals:

$\Sigma x = 41.3$, $\Sigma Y = 19.41$ $n = 5$

$\Sigma x^2 = 464.55$ $\Sigma xY = 172.02$

Sub in normal equations

$\Sigma Y = nA + B\Sigma x$

$\Sigma xY = A\Sigma x + B\Sigma x^2$

$19.41 = 5A + 41.3B$

$172.02 = 41.3A + 464.55B$

$A = 3.09$ | $a = 21.97$

$B = 0.094$ | $b = 1.09$

$\therefore$ fitted eqn

$\boxed{\hat{y} = (21.97)(1.09)^x}$

Assign - 1

1. Fit the equation

(a) $y = ae^{bx}$

(b) $y = ax^{b}$

(c) $y = ab^{x}$

x	17.6	29.5	34.2	48.1	60.5	73.4
y	116.2	210.5	250.1	310.6	360.5	400.8

(2) fit the equation $y = a + xb$

x	15	35	50	66	72	80
y	85	110	140	165	190	210

(3) fit the equation $y = ax^2 + bx + c$

x	1	3	8	11	15	
y	25	40	55	70	90	