14/02/24

Date ____
Page ____

* Arithmatic Mean:

Summation of all values $\frac{\Sigma x_i}{N}$ where, N = no. of Observation

Q.1) Find the avg. marks obtained by user,

64, 69, 72, 72, 75, 65

$$\text{Mean} = \frac{\Sigma x_i}{N} = \frac{64+69+72+72+75+65}{6} = 69.5$$

Q.2) Find the arithmatic mean for following :-

x_i	No. of days Spent	1	2	3	4	5	6	7	8
f_i	No. of patient	5	6	5	10	8	4	3	2

$$\frac{\Sigma x_i}{N} = \frac{5+12+15+40+40+24+21+16}{5+6+5+10+8+4+3+2}$$

$$= \frac{173}{43} = 4.02$$

mean = $\frac{\Sigma \text{ frequency} \times \text{midValue}}{N}$

* Median

Q.1) Find the median of given value.

65, 55, 89, 56, 35, 14, 56, 55, 87, 45, 92.

Sorting in asc :- 14, 35, 45, 55, 55, 56, 56, 65, 87, 89, 92

median is $\boxed{56}$

* $\boxed{\text{Mode}}$: Most repeatable element

find mode. 14, 35, 45, 55, 55, 56, 56, 65, 87, 89, 92

mode = 55, 56

Q.1] Find the median of following:-

monthly sal.	frequency.	cf. less than
100 - 200	15	15
120 - 140	35	50
$\boxed{140 - 160}$	~~15~~ $\boxed{50}$	100
160 - 180	60	160
180 - 200	30	190
200 - 220	10	200
	200 N.	

median $= l_1 + (l_2 - l_1)\dfrac{\left(\frac{N}{2} - c.f\right)}{f}$

l_1 = lower limit of median class
l_2 = upper limit of
cf = cumulative frequency of pre. median class
f = frequency of median class.

$\frac{N}{2} = \frac{200}{100}$. $l_1 = 140$, $l_2 = 160$, ~~cf = 200~~
~~f~~ $cf = 50$, $f =$ ~~200~~ 50

$l_1 + (l_2 - l_1)\dfrac{\left(\frac{N}{2} - cf\right)}{f}$

$= 140 + (160 - 140)\left(100 - 50\right)$

$= 140 + 20 \left(\frac{50}{50}\right)$

$= \underline{\underline{160}}$

$\boxed{\text{Ans:- Median is 160}}$

Exp :- Tossing of 2 coins.

$\{HH, HT, TH, TT\}$ = Random Experiment.
= Sample Space:

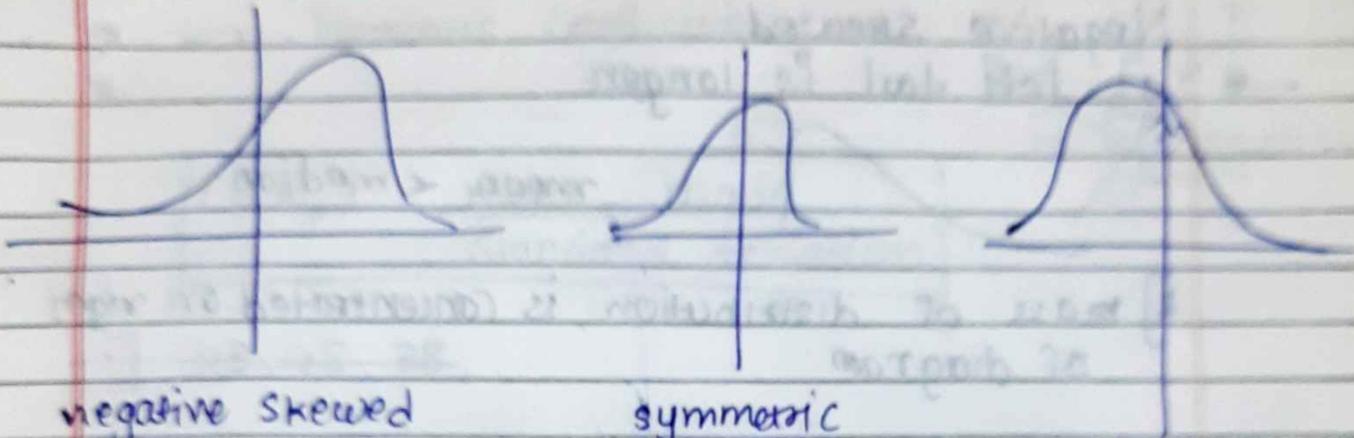
$\boxed{\text{Maths Online Lecture}}$

frequency distribution : representation of data.

class mark :- $100 - 120 = \frac{120-100}{2} = \cancel{\frac{20}{12}}\ \frac{20}{2} = 10$

at 10^{th} place $= 110$

UNIT TWO



negative skewed

symmetric skewed

positive

Skewness

skewness is the measure of asymmetry in probability distribution (symmetry about its mean)

~~Negative skewed~~

Positive skewd

Symmetric

Negative skewed

↳ left tail is longer

↳ mass of distribution is concentrated on right of diagram

1. Negative skewed

→ 1] left tail is longer

2]  mean < median

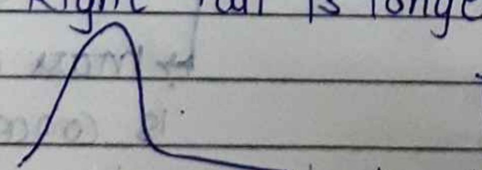
3] mass of distribution is concentrated on right of diagram.

2. Symmetric

1]

3] positive skewed.

1] Right tail is longer.

2]  $\boxed{\text{mean > median}}$

3] mass of distribution is concentrated on left of diagram.

* Karl-Pearson's coefficient
*

if $(Sk_p) < 0$
it is negative
Sk

$$\boxed{(Sk_p) = \frac{\text{mean} - \text{mode}}{\text{Standard deviation}}}$$

~~1] 43, 48, 38,~~

1] if $(Sk_p) < 0 \longrightarrow$ -vely skewed
2] if $Sk_p = 0 \longrightarrow$ Symmetric.
3] if $Sk_p > 0 \longrightarrow$ +vely skewed.

* mean - mode = 3 (mean - median).
* mode = 3 (median) - 2 (mean)

1] 43, 48, 38, 46, 50, 48, 47, 48, 62, 48

$\bar{x}$ = mode = 48
mean = 47.8

$$\boxed{SD = \sqrt{\frac{\sum f_i x_i^2}{n} - \bar{x}^2}}$$

for ungrouped data $f_i = 1$.

$$SD = \sqrt{\frac{\sum f_i x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{23178}{10} - (47.8)^2}$$

$$\boxed{SD = 5.74108}$$

By karl Pearson method,

$$= \frac{\text{mean} - \text{mode}}{SD}$$

$$= \frac{47.8 - 48}{5.74108}$$

$$= -0.03484$$

Ans - Data is negatively skewed.

Ex.2 Compute karl pearson coefficient for grouped data.

Daily wages	No. of workers (f_i)	x_i	$x_i f_i$	$(x_i f_i)^2$	x^2
400 - 500	8	450	3600		
500 - 600	16	550	8800		
⇒ 600 - 700	20	650	13000		
700 - 800	17	750	12750		
800 - 900	3	850	2550		

$\boxed{n = 64}$ $\boxed{\cancel{n=7}}$

~~$\bar{x}$ = mean~~

modal class → 600 - 700
$\qquad\qquad\qquad\quad l_1 \qquad l_2$

$$\text{mode} = l_1 + (l_2 - l_1)\left[\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right]$$

$f_1 \Rightarrow$ frequency of modal class = 20
$f_0 \rightarrow$ frequency of pre modal class = 16
$f_2 \rightarrow$ 11 post modal = 17

$$\text{mode} = 600 + (700-600)\left[\frac{(20-16)}{2(20-16-17)}\right]$$

$$= 707.1429$$

20 has highest frequency so its class will be modal class.

$$\boxed{\text{mode} = 707.1429}$$

$$SD = \sqrt{\frac{\Sigma f_i x_i^2}{n} - \bar{x}^2} =$$

$$\text{mean} = \frac{\Sigma f_i x_i}{n} = \cancel{0.65457}$$

$$\boxed{S_{kp} = -0.65457}$$

* Bowley's coefficient of skewness $(Sk)_B$

$$\boxed{(Sk)_B = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 \pm Q_2) + (Q_2 - Q_1)}}$$

Q.1] Calculate the Bowley's coefficient of skewness for the following.

Life in hours	< 10	20-30	30-40	40-50	50-60
No. of Bulbs	12	18	24	20	6

Ans:-

median class Q_1

Life in hours	< 10	20-30	30-40	40-50	50-60
No. of Bulbs	12	18 f.	24	20	6
Cummulative frequency	12 cf.	30	54	74	80

$N = 80$ To find Q_1 =

Q_1 Q_2 Q_3

for Q_1. $N/4 = \cancel{40}\ 20$

$N = 80$ $\frac{N}{2} = 40$ $\frac{40}{2} = 20$.

median class = $\boxed{20-30}$

$$Q_1 = l_1 + \frac{(l_2 - l_1)\left(\frac{N}{4} - cf\right)}{f}$$

(Take previous class of median) $cf = 12$, $l_1 = 20$

$f = 18$ $l_2 = 30$

$$= \frac{20 + (30-20)\left(\frac{80}{4} - 12\right)}{18}$$

$$\boxed{Q_1 = 21.4444}$$

$\cancel{Q_2 = l_1 + l_2 -}$

for Q_2 ,

$\frac{N}{2} = 40$ $\boxed{\text{median class} = 30-40}$

$l_1 = 30$, $cf = 30$

$l_2 = 40$ $f = 24$

$$Q_2 = \frac{l_1 + (l_2 - l_1)\left(\frac{N}{2} - cf\right)}{f}$$

$$= \frac{30 + (40-30)(40-30)}{24}$$

$= \cancel{30 + (10)(}$

$Q_2 = 34.1667$

for Q_3,

$$Q_3 = \frac{l_1 + (l_2 - l_1)(3N/4 - cf)}{f}$$

$l_1 = 40$ $cf = 54$ $\frac{3N}{4} = 60$
$l_2 = 50$ $f = 20$

$$Q_3 = \frac{40 + (50-40)(60-54)}{20}$$

$= \cancel{0.749}$ 5

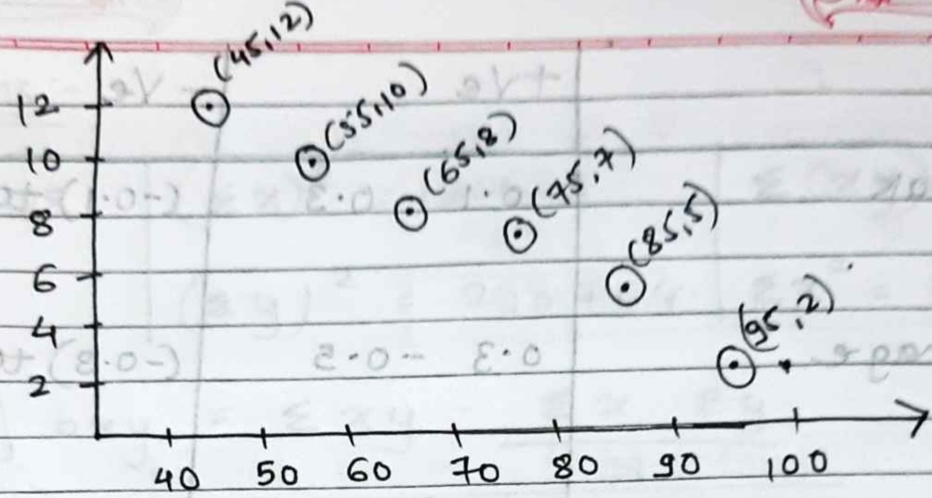
$$(SK)_B = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)}$$

$= -0.0479$

Ans :- Negatively skewed.

Q.1] Draw the scattered diagram of following and understand the type of correlation between them.

marks.	No. of students (y)	class mark (x)	
40-50	12	45	(45,12)
50-60	10	55	(55,10)
60-70	8	65	(65,8)
70-80	7	75	(75,7)
80-90	5	85	(85,5)
90-100	2	95	(95,2)



It is Negatively correlated

* Correlation Coefficients.

→ Karl pearson coeffi. of correlation

→ Denoted by (r_{xy})

$\Rightarrow$ Spearman's Rank Coeffi.

$$\boxed{R = 1 - \frac{6\Sigma d^2}{n(n^2-1)}}$$

$$\Rightarrow (r_{xy}) = \sqrt{b_{xy} \cdot b_{yx}} \text{ \} suffix.}$$

$\Rightarrow$ Properties 1] $r \in [-1, 1]$

2] r is unitless quantity

$$\boxed{b_{xy} = \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{N}}{\Sigma x^2 - \frac{(\Sigma x)^2}{N}}}$$

$$b_{yx} = \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{N}}{\Sigma y^2 - \frac{(\Sigma y)^2}{N}}$$

	+ve	-ve
Weak	0.1 - 0.3	(-0.1) to (-0.3)
Average	0.3 - 0.5	(-0.3) to (-0.5)
Strong	0.5 - 1.0	(-0.5) to (-1.0)

Example of karl pearson

(Q.1) An experiment is conducted on 9 different cigarette smoking subjects resulted in the following data. (karl pearson)

← Given → ← find out →

Cigarette smoked per week (x)	No. of years lived (y)	(x^2)	(y^2)	xy
25	63	625	3969	1575
35	68	1225	4624	2380
10	72	100	5184	720
40	62	1600	3844	2480
85	~~4~~65	7225	4225	5525
75	46	5625	2116	3450
60	51	3600	2601	3060
45	60	2025	3600	2700
50	55	2500	~~3136~~ 3025	2750

$\boxed{N=9}$ $\Sigma x = 425$ | $\Sigma y = 542$. | Σx^2 24525 | $\Sigma y^2 = 33188$ | Σxy 24640

Calculate the correlated coefficient between the no. of cigarette smoked and longevity of test subjects.

Ans :-

$\Sigma y^2 = 33188$	$(\Sigma x)^2 = 180625$	$\Sigma(xy) = 24640$
	$(\Sigma y)^2 = 293764.$	$\Sigma x^2 = 24525$

i] $b_{xy} = \dfrac{\Sigma xy - \dfrac{\Sigma x \cdot \Sigma y}{N}}{\Sigma x^2 - \dfrac{(\Sigma x)^2}{N}}$

$= \dfrac{24640 - \dfrac{(425)(542)}{9}}{24525 - \dfrac{(425)^2}{9}}$

$\boxed{b_{xy} = -0.2142}$

for.

ii] $b_{yx} = \dfrac{\Sigma xy - \dfrac{\Sigma x \cdot \Sigma y}{N}}{\Sigma y^2 - \dfrac{(\Sigma y)^2}{N}}$

$= \dfrac{24640 - \dfrac{(425)(542)}{9}}{33188 - \dfrac{(542)^2}{9}}$

$\boxed{b_{yx} = -1.7431}$

$(r_{xy}) = \sqrt{(-0.2142)(-1.7431)} = 0.61$

$(r_{xy}) = 0.61$, so it is strong positive correlation

* Spearman's Rank Correlation coefficient:

$$\boxed{R = 1 - \frac{6 \Sigma d^2}{n(n^2-1)}}$$

Q.1] The following table provides the data about the percentage of students who have free university mess (meals) and their CGPA Scores Calculate the Spearsman's rank Correlation between the two and interpret the result

State University	(x) % of Stud having free	(y) % of Stud Scoring abv 8.5 CGPA	d_x rank	d_y rank	$(d_x - d_y)^2$
Pune	14.4	54	3	4	$(-1)^2 = 1$
Chennai	7.2	64	2	6	$(-4)^2 = 16$
Delhi	27.5	44	5	3	$(2)^2 = 4$
Kanpur	33.8	32	6	1	$(5)^2 = 25$
Ahmedabad	38.8	37	7	2	$(5)^2 = 25$
Indore	15.9	68	4	7	$(-3) = 9$
Guwahati	4.9	62	1	5	$(-4)^2 = 16$

$\boxed{N = 7}$

$\Sigma (d_x - d_y)^2 = 96$

Arranged dx and dy in increasing order as their occurrence in x and y respectively.

$$R = 1 - \frac{6 \sum d^2}{n(n^2-1)}$$

$$\boxed{\sum d^2 = 96}$$

$$R = 1 - \frac{6(96)}{7(7^2-1)}$$

$$\boxed{R = -0.714}$$

Ans:- It has strong negative correlation

* Regression Analysis.

Q.1] The experimental values relating to centripetal force (N) and radius for a mass travelling at constant velocity

Force (N)	5	10	15	20	25	30	35	40
Radius	55	30	16	12	11	9	7	5

Determine the eqn. of regression line of force on radius and hence calculate the force at radius of 40 cm.

$\boxed{y = mx + c}$ — regression line of Y on X

slope Intercept

$y = a_0 x + a_1$ slope Intercept

$\cancel{\Sigma y = a_0}$

$$\boxed{\begin{aligned} \Sigma y &= a_0 \cdot N + a_1 \cdot \Sigma x \\ \& \quad \Sigma xy &= a_0 \Sigma x + a_1 \Sigma (x^2) \end{aligned}}$$

Ans:-

Radius (x)	Force (y)	x^2	y^2	xy
55	5	3025	25	275
30	10	900	100	300
16	15	256	225	240
12	20	144	400	~~275~~ ⇒ 240
11	25	121	625	275 ⇒ 275
9	30	81	900	~~245~~ ⇒ 270
7	35	49	1225	245
5	40	25	1600	200
$\Sigma x =$ $\boxed{145}$	$\Sigma y =$ $\boxed{180}$	Σx^2 $\boxed{4601}$	$\Sigma y^2 =$ $\boxed{5100}$	$\Sigma xy =$ $\boxed{2045}$

$$\boxed{\begin{aligned} \therefore \Sigma y &= a_0 N + a_1 \Sigma x \\ \Sigma xy &= a_0 \Sigma x + a_1 \Sigma (x^2) \end{aligned}}$$

$\therefore 180 = 8 a_0 + (145) a_1$

$\therefore 2045 = (145) a_0 + (4601) a_1$

Solving two eqn. ~~usin~~ we get,

$a_0 = 33.7 \quad , a_1 = -0.617$

$y = a_0 + a_1 \cdot x$

Regression, line of y on x

$y = a_0 + a_1 \cdot x$

$$\boxed{y = 33.7 - (0.617)x}$$

Take 40 cm as value.

$y = 33.7 - (0.617)40$

$y = 9.0$

Wednesday

Linear Regression

Q.1 In study between the amount of rainfall and quantity of the air pollution removed the following data were collected

Daily rainfall 0.01 cm (x)	4.3	4.5	5.9	5.6	6.1	5.2	3.8	2.1
Pollution removed mg/m^3 (y)	12.6	12.1	11.6	11.8	11.4	11.8	13.2	14.1

find the regression line of y and x.

x	y	x^2	y^2	xy
4.3	12.6	18.49	158.76	54.18
4.5	12.1	20.25	146.41	54.45
5.9	11.6	34.81	134.56	68.44
5.6	11.8	66.08	139.24	66.08
6.1	11.4	37.21	129.96	69.54
5.2	11.8	27.04	139.24	61.36
3.8	13.2	14.44	174.24	50.16
2.1	14.1	4.41	198.81	~~20~~ 29.61
$\boxed{n=8}$	$\boxed{\Sigma y = 98.6}$	$\boxed{\Sigma x^2 = 188.9}$	$\boxed{\Sigma y^2 = 1221.22}$	$\boxed{\Sigma xy = 453.82}$

$\boxed{\Sigma x = 37.5}$

$y = ax + b$

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

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

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

$$98.6 = 8a + b(37.5)$$

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

$$453.82 = a(37.5) + b(\cancel{22.73})\ (188.01)$$

By solving this eqn. we get,

$\boxed{a = 15.49}$ $\boxed{, b = -0.675}$

~~y = ax -~~

$$y = a + bx$$

$$\boxed{y = (15.49) + (-0.675)x}$$

$0 \le P \le 1$

UNIT 3

~~Prof~~

Probability & statistics

$\Omega / S = \{HH, HT, TH, TT\}$

$n = 4$

$2^n = 2^4$

possibility subset $: 2^4$

Event Head is on first coin

$A = \{HT, HH\}$ $\left[\subseteq S\right]$ It is called event

$A \subseteq S$

event is subset of sample space

$$\text{Probability} = \frac{n(A)}{n(S)} = \frac{2}{4}$$

$P(\phi) = 0$

$P(S) = 1$

$P(A) = \underline{n(A)}$

Mutually Exclusive events

$A \cap B = \phi$

Exhaustive events

$A \cup B = S$

Random variable :

exp : Tossing two coins.

$$S/\Omega = \{HH, HT, \cancel{} HT, TT\}.$$

$$X = \text{no. of heads}$$

$X(HH) = 2$

$X(HT) = X(TH) = 1$

$X(TT) = 0$

$X = \{0, 1, 2\} \Longrightarrow$ This X is called as Random Variable.

$f : \underset{\text{mapping}}{\Omega \longrightarrow X}$

This is called mapping of real word with num.

Saturday

Maths Online Lecture

$$\left.\begin{matrix} 25 - 29 \\ 30 - 34 \\ 35 - 39 \\ 40 - 44 \end{matrix}\right\} \text{disjoint classes.}$$

$$\left.\begin{matrix} 24.5 - 29.5 \\ 29.5 - 34.5 \\ 34.5 - 39.5 \\ 39.5 - 44.5 \end{matrix}\right\}$$

* **Correlation** : ~~exter~~ degree of relⁿ betⁿ 2 variables

Probability $0 \leq P \leq 1$ is 0

* **Random Variable** :-

X	0	1	2	3

Continuous random variable
Discrete random variable

Mathematics

Probability and Statistics

Find out mean, median and mode.

wages	15-25	25-35	35-45	45-55	55-65
No. of workers	4	6	12	5	3

* Mean :-

wages (x)	No. of workers (F)	midvalue (M)	(FM)
15-25	4	20	80
25-35	6	30	180
35-45	12	40	480
45-55	5	50	250
55-65	3	60	180
	$\Sigma f = 30$	Σ	$\Sigma fm = 1170$

$$mean = \frac{\Sigma fm}{\Sigma f} = \frac{\cancel{1170}^{39}}{\cancel{30}_{1}} = 39$$

$\boxed{mean = 39}$

* Median

No. of wages (x)	No. of worker (F)	cumulative frequency (CF)	Position of median size of $(N/2)^{th}$ item $= (30/2)^{th} = 15^{th}$
15-25	4	4	
25-35	6	10 → cf	Here we cannot see 15 so we got 22 bcz it is more than 15
$\boxed{35-45}$ - u	$\boxed{12}$ f	22	
45-55	5	27	
55-65	3	30	
	$N = \Sigma f = 30$	Σ	

l

$$\cancel{\text{median} = l + \left[\frac{\frac{N}{2} - cf}{f}\right] \times i}$$

median = ~~l_2 + (~~

$$\boxed{l_1 + (l_2 - l_1)\left(\frac{\frac{N}{2} - c.f}{f}\right)}$$

$l_2 = 45$, $l_1 = 35$, $\frac{N}{2} = 15$, $c.f = 10$, $f = 12$

$$35 + (45 - 35)\left(\frac{15 - 10}{12}\right)$$

$$= 35 + 10\left(\frac{5}{12}\right)$$

* mode :

wages (x)	No. of workers (f)
15-25	4
25-35	6 $\rightarrow f_0$
$l_1 \leftarrow$ 35-45 $\rightarrow l_2$	12 $\rightarrow$ Highest frequenc (f_1)
45-55	5 $\rightarrow f_2$
55-65	3

$$\text{mode} = l_1 + (l_2 - l_1)\left[\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right]$$

$$= 35 + (45 - 35)\left[\frac{12 - 6}{2(12) - 6 - 5}\right]$$

$$= 35 + 10\left[\frac{6}{26 - 11}\right]$$

$$= 35 + 10\left[\frac{6}{15}\right]$$

$$= 45\left[\frac{2}{5}\right]$$

$$\boxed{mode = 18}$$

* **Standard Deviation :**

Individual series . $SD = \sqrt{\frac{\Sigma x^2}{N} - \left[\frac{\Sigma x}{N}\right]^2}$

Discrete series : $SD = \sqrt{\frac{\Sigma Fx^2}{N} - \left[\frac{\Sigma Fx}{N}\right]^2}$

Continuos series $SD = \sqrt{\frac{\Sigma Fm^2}{N} - \left[\frac{\Sigma FM}{N}\right]^2}$
(class interval , frequency)

C.I	0-10	10-20	20-30	30-40
F	3	1	2	4

Calculate standard deviation

C.I.	F	mid value	Fm	Fm^2
0 - 10	3	5	15	$3\times(5)^2 = 75$
10 - 20	1	15	15	$1\times(15)^2 = 225$
20 - 30	2	25	50	$2\times(25)^2 = 1250$
30 - 40	4	35	140	$4\times(35)^2 = 4960$
	$\cancel{\Sigma f}$ N = 10		$\Sigma fm = 220$	$\Sigma fm^2 = 6450$

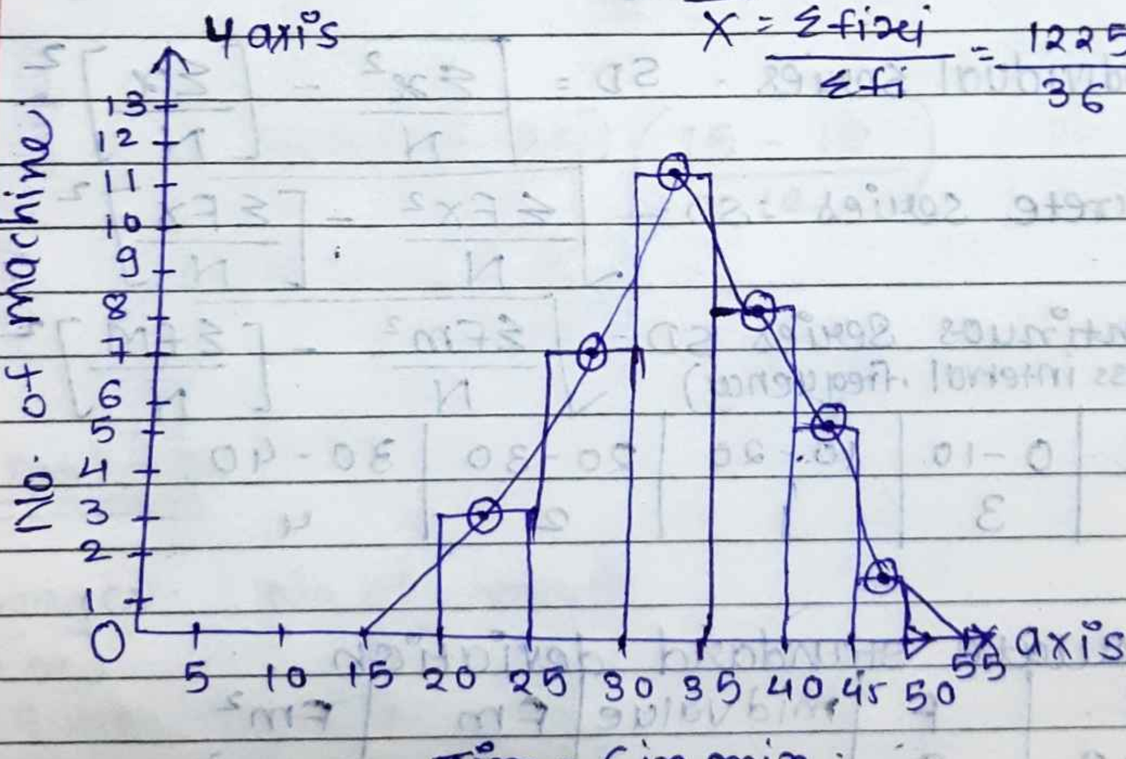
$$SD = \sqrt{\frac{\Sigma fm^2}{N} - \left[\frac{\Sigma fm}{N}\right]^2} = \sqrt{\frac{6450}{10} - \left[\frac{220}{10}\right]^2} = \boxed{12.68}$$

* Histogram :

Draw histogram and frequency polygon for following data

Time (in min)	No. of machine f_i	x_i/m	$f_i x_i$
20 - 25	3	22.5	67.5
25 - 30	7	27.5	192.5
30 - 35	11	32.5	357.5
35 - 40	8	37.5	300
40 - 45	5	42.5	212.5
45 - 50	2	47.5	95

$$\overline{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{1225}{36} = 34.027$$



* Stem and leaf diagram plot:

Q.1] 15, 27, 8, 17, 13, 22, 24, 25, 13, 36, 32, 32, 32, 28, 43, 7.

write down ~~unit~~ Tens place | unit place | make them in asc. order

Stem (S)	Leaf (L)
~~0~~	7 8
1	3 3 5 7
2	2 4 5 7 8
3	2 2 2 6
4	3

7, 8
13, 13, 15, 17
22, 24, 25, 27, 28
32, 32, 32, 36
43

Q.2] 1.2, 2.3, ~~2.3~~, 1.5, 2.4, 3.6, 1.8, 2.7, 3.2, 4.1, 2.9, 4.5, 7.6, 5.8, 9.3, 10.6, 12.4, 10.9.

Stem (S)	Leaf	
1	2 5 8	1.2, 1.5, 1.8
2	3 4 7 9	2.3, 2.4, 2.7, 2.9
3	2 6	3.2, 3.6
4	1 5	4.1, 4.5
5	8	5.8
6		
7	6	7.6
8		
9	3	9.3
10	6, 9	10.6, 10.9
11		
12	4	12.4.

143, 160, 154, 159, 172, 165, 162, 171
146 165, 176 145 165 182 195, 186, 160
158, 167, 172.

1] Display the data in a stem and leaf plot
2] Find the range of height ? 43 cm
3] What is the mode height ? 165 cm
4] What is the median height ? 165 cm.

Stem	leaf
~~14~~	~~3 5 6~~
~~15~~	~~4 8 9~~
~~16~~	~~0 2 5 7 9~~
~~17~~	~~1,2 5 6 7~~
~~18~~	~~2 6~~

~~143, 145, 146~~

	stem	Leaf
	14	3 5 6
	15	4 8 9
165 ←	16	0 0 2 $\boxed{5}$ $\boxed{5}$ 5 7 → most frequent (10th, 11th)
mode	17	1 2 2 5 6
	18	2 6

2) Find range of height.

$$\text{Range} = \max - \min$$
$$= 186 - 143$$
$$= \boxed{43}\ \text{cm}$$

3) mode height

no which is most frequent = 5, $\boxed{165 = \text{mode}}$

4] median :-

Here we have 20 students.
So avg. will be 10^{th} and 11^{th} element.

$$\frac{10^{th} + 11^{th}\text{ ele}}{2} = \frac{5+5}{2} = \frac{10}{2} = 5$$

5] what % of stu. have height greater than 175 cm?

We have 3 stu. greater than 175
i.e. 176, 182, 186

$$\frac{3}{20} \times 100 = 15\%$$

Ans :- 15% stu. have height greater than 175 cm.

* Ogives :-

Find less than
more than ogives for following data

marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of stu.	15	25	60	20	10	8	2	10

Ogives :

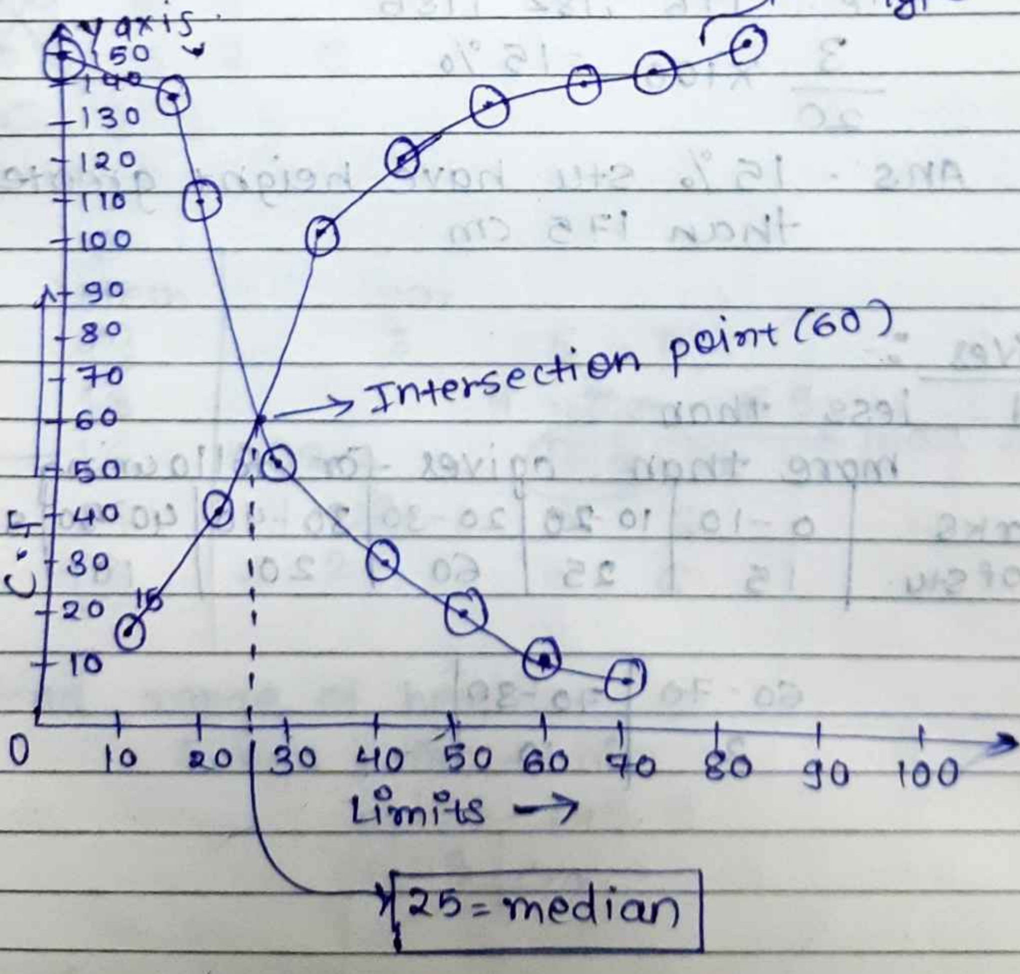
marks C.I.	No. of stu. (F)	less than Type c.f		More than Type	
0-10	15	less than 10	15	more than 0	150
10-20	25	less than 20	40	-11- 10	135
20-30	60	-11- 30	100	-11- 20	110
30-40	20	-11- 40	120	-11- 30	50
40-50	10	-11- 50	130	-11- 40	30
50-60	8	-11- 60	138	-11- 50	20
60-70	2	-11- 70	140	-11- 60	12
70-80	10	-11- 80	150	-11- 70	10

150-15

135-25

same for rest

N=150



For Grouped Data

Arithmatic mean, ~~Geo~~ Geometric mean, Harmonic mean.

find AM, GM, HM

marks (C.I)	0-10	10-20	20-30	30-40
No. of Stu (F)	3	1	2	4

Ans:-

* Arithmatic mean

C.I	No. of stu (F_i)	mid value (x_i)	$f_i x_i$
0-10	3	5	15
10-20	1	15	15
20-30	2	25	50
30-40	4	35	140

$N = 10$ $\quad$ $\Sigma fx = 220$

$$\boxed{\text{Arithmatic mean} = \frac{\Sigma fx}{N}} = \frac{220}{10} = 22$$

* Geometric mean:

C.I	(F)	mid value (x_i)	log x	F.log X
0-10	3	5	0.6989	2.0967
10-20	1	15	1.1760	1.1760
20-30	2	25	1.3979	2.7958
30-40	4	35	1.5440	6.176

$N = 10$ $\quad$ $\Sigma F \log X = 12.2445$

$$\boxed{\text{Geometric mean} = \text{Antilog}\left[\frac{\Sigma f.\log X}{N}\right]}$$

$$= \text{Antilog}\left[\frac{12.2445}{10}\right]$$

$$= \text{Antilog}\left[1.22445\right]$$

$$= 16.7667$$

* Harmonic mean :

C.I.	No. of stu. (f)	mid value (x)	1/x	f(1/x)
0-10	3	5	0.2	0.6
10-20	1	15	0.0666	0.0666
20-30	2	25	0.04	0.08
30-40	4	35	0.0285	0.114
	N=10			$\Sigma f(1/x) = 0.8606$

$$\boxed{\text{Harmonic mean} = \frac{N}{\Sigma f(1/x)}}$$

$$= \frac{10}{0.8606}$$

$$= 11.6198$$

For Ungrouped Data.

Q.1] Arithmatic mean

X	2	4	8	16
F	2	3	3	2

Ans :-

X	F	Fx
2	2	4
4	3	12
8	3	24
16	2	32
	N=10	$\Sigma fx = 72$

$$\text{Ari. mean } \bar{x} = \frac{\Sigma fx}{N}$$

$$\bar{x} = \frac{72}{10} = 7.2$$

$$\boxed{AM = 7.2}$$

Geometric mean

x	F	log x	F. log x
2	2	0.301	0.602
4	3	0.602	1.806
8	3	0.903	2.709
16	2	1.204	2.408

$N = 10$ $\qquad \Sigma f \cdot \log x = 7.525$

$$GM = \text{Antilog}\left[\frac{1}{N} \cdot \Sigma f \log x\right]$$

$$= \text{Antilog}\left[\frac{1}{10} \cdot 7.525\right]$$

$$= \text{Al}\left[0.7525\right]$$

$$\boxed{GM = 5.655}$$

* Harmonic Mean

x	F	f/x
2	2	1
4	3	3/4
8	3	3/8
16	2	1/8

$N = 10$ $\qquad N = 9/4$

$$\boxed{HM = \frac{N}{(\Sigma f/x)}} = \frac{10}{9/4} = \frac{40}{9} = 4.44$$

$$\boxed{HM = 4.44}$$

4.4
9) 40
 - 36
 040

Measure of Dispersion :

Dispersion measure the extent to which items vary from some central value (mean, median mode)

measure of Dispersion

Absolute meas
1. Range
2. Quartile Deviation
3. Mean Deviation
4. Standard deviation
5. Variance

Relative mea
1. Coefficient of range
2. Coefficient of Quartile dev
3. -11- mean deviation
4. -11- SD
5. -11- Variance

(1) Range = Largest obs (L) - Smallest obs (S)

(2) Coefficient of Range $= \frac{L-S}{L+S} \times 100$

* Individual series :-

13, 20, 7, 15, 29, 35

S = 7   L = 35

Range = L - S
= 35 - 7
= 28

Coefficient of Range $= \frac{L-S}{L+S} \times 100$

$= \frac{35-7}{35+7} \times 100$

= 66.67 %

* Discrete

X	F
5	3
10	7
15	5
20	12
25	9

L = 25   S = 5

Range = 25 - 5
= 20

Co-efficient of Range $= \frac{L-S}{L+S} \times 100$

$= \frac{25-5}{25+5} \times 100$

$= 66.67\%$

* Continuose Data

C.I.	f
(S) 20-25	2
25-30	5
30-35	9
35-40	4
40-45 (L)	7

Range = L - S
= 45 - 20
= 25

Co-efficient of Range $= \frac{L-S}{L+S} \times 100$

$= \frac{45-20}{45+20} \times 100$

$= 38.46\%$

Inclusive series

20-24	1
25-29	2
30-34	3
35-39	4
40-44	5

Exclusive

19.5-24.5	1
24.5-29.5	2
29.5-34.5	3
34.5-39.5	4
39.5-44.5	5

* Quartile Deviation :-

1. $QD = \dfrac{Q_3 - Q_1}{2}$

2. Co-efficient of $QD = \dfrac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

$Q_1 \rightarrow$ first Quartile (lower Quartile)
$Q_2 \rightarrow$ Second Quartile.
$Q_3 \rightarrow$ Third Quartile (upper Quartile).

* Individual series :

2, 7, 9, 14, 1
1, 2, 7, 9, 14

median

median makes equal 2 parts.

Q_1 Q_2 Q_3

Quartile makes equal 4 part.

5, 9, 20, ~~23~~, 35, 40, 3, 10
3, 5, 9, 10, 20, 35, 40 $n = 7$

i) $Q_1 = \left(\dfrac{n+1}{4}\right)^{th}$ term $= \left(\dfrac{7+1}{4}\right)^{th}$ term
$= 2^{nd}$ Term $= 5$

$\boxed{Q_1 = 5}$

ii) $Q_3 = 3\left(\dfrac{n+1}{4}\right)^{th}$ term $= 3\left(\dfrac{7+1}{4}\right)^{th}$ term
$= 6^{th}$ term $= 35$

$\boxed{Q_3 = 35}$

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

$= \frac{35-5}{2}$

$= \frac{30}{2}$

$\boxed{QD = 15}$

Coefficient of QD $= \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

$= \frac{35-5}{35+5} \times 100 = \frac{30}{40} \times 100$

$\boxed{\text{coefficient QD} = 75\%}$

If I get $\left(\frac{n+1}{4}\right)^{th}$ term as

$(2.25)^{th}$ term:

$= 2^{nd}$ term $+ (0.25)(3^{rd} - 2^{nd})$

$=$

* Discrete series :-

x	f
5	2
9	2
2	1
10	5
15	3

Arrange

x	f	cf
2	1	1
5	2	3
9	2	5
10	5	10
15	3	13
	$N = \Sigma f = 13$	

1st term

2nd term

$Q_1 = \left(\frac{N+1}{4}\right)^{th} = \left(\frac{13+1}{4}\right)^{th} = (3.5)^{th}$ term

(3.5^{th})

$= 3^{rd}$ term $+ 0.5\,(4^{th} - 3^{rd}$ term$)$

$= 5 + 0.5\,\cancel{(10-9)}\,(9-5)$

$= 3 + 0.5\,\cancel{(1)}$

$\boxed{\cancel{Q_1 = 3.5}}$ $\boxed{Q_1 = 7}$

$$Q_3 = 3\left(\frac{N+1}{4}\right)^{th} = 3\left(\frac{13+1}{4}\right)^{th} \text{ term}$$

$$= (10.5)^{th} \text{ term}$$

$$= 10^{th} \text{ term} + (0.5)(11^{th} - 10^{th} \text{ term})$$

$$= 10 + (0.5)(15-10)$$

$$\boxed{Q_3 = 12.5}$$

$$QD = \frac{Q_3 - Q_1}{2} = \cancel{\frac{7-12.5}{2}} \; \frac{12.5-7}{2} = 2.75$$

$$\text{Coefficient of } QD = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

$$= \cancel{\frac{7-12.5}{7+12.5} \times 100}$$

$$= \frac{12.5-7}{12.5+7} \times 100 = 28.2\%.$$

* continuose series.

class	f
0 - 10	3
10 - 20	9
20 - 30	12
30 - 40	11
40 - 50	5
50 - 60	6

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

$$\text{coefficient } QD = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

$$Q_1 = L + \left(\frac{\frac{N}{4} - c.f}{f}\right) \times h$$

L = lower limit

N → Σf

CF → Upper Cf

f → freq^n

h → class interval = 0-10
= 10

$$Q_3 = L + \left(\frac{\frac{3N}{4} - \Sigma c.f}{f}\right) \times h.$$

class	f	cf
0 - 10	3	3
10 - 20	9	12
20 - 30	12	24
30 - 40	11	35
40 - 50	5	40
50 - 60	6	46

$\rightarrow$ cf for Q_1

for Q_1 bcz $\frac{N}{4} = \frac{46}{4} = \boxed{11.5}$

$\rightarrow$ for Q_3 bcz $\frac{3N}{4} = 34.5$

$N = \Sigma f = 46$

For Q_1 quartile $\frac{N}{4}$ take class interval of $N/4 = \frac{10^{th}}{4}$

Q_2, quartile $\left(\frac{3N}{4}\right)$ take class interval of $\left(\frac{3N}{4}\right)^{th}$ term.

$$Q_1 = L + \left(\frac{\left(\frac{N}{4} - c.f\right)}{f}\right) \times h$$

$$= 10 + \left(\frac{11.5 - 3}{9}\right) \times 10$$

$$= 10 + \frac{8.5}{9} \quad \boxed{Q_1 = 19.44}$$

$$Q_3 = L + \left(\frac{\frac{3N}{4} - c.f}{f}\right) \times h$$

$$= 30 + \left(\frac{34.5 - 24}{11}\right) \times 10$$

$$\boxed{Q_3 = 39.54}$$

$$QD = \frac{Q_3 - Q_2}{2} = \frac{39.54 - 19.44}{2} = 10.05$$

Coefficient QD $= \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100 = 34.07\%$

Standard Deviation :

Individual Series

$$SD = \sigma = \sqrt{\frac{(x_i - \overline{x})^2}{N}} \quad \overline{x} = \text{mean} \quad \text{ungrouped}$$

class mark

$$\sigma = \sqrt{\frac{\Sigma f_i (x_i - \overline{x})^2}{n}} \quad \text{grouped data}$$

$$SD = \sqrt{\frac{\Sigma f_i x_i^2}{n} - \overline{x}^2}$$

$$\frac{\Sigma (x_i - \overline{x})}{N} \qquad \cancel{\sqrt{\frac{\Sigma f_i (x_i - \overline{x})}{N}}}$$

f_i	x_i	$f x_i$	$x_i - \overline{x}$
11	5		
15	15		
25	25		
12	35		
7	45		

$$\boxed{\overline{x} = \frac{\Sigma f x_i}{\Sigma f}} = \frac{1640}{70} =$$

$$\overline{x} = \frac{\Sigma f x_i}{\Sigma f} = 23.42$$

SD and variance. eg.

C.I	f_i	x_i	f_ix_i	x_i^2	$f_ix_i^2$
0-10	11	5	55	25	275
10-20	15	15	225	225	3375
20-30	25	25	625	625	15625
30-40	12	35	420	1225	14700
40-50	7	45	315	2025	14175

$\Sigma f_i = 70$ $\Sigma f_ix_i = 1640$ $\Sigma f_ix_i^2 = 48150$

$$\bar{X} = \frac{\Sigma f_ix_i}{N} = \frac{1640}{70} = 23.4285$$

$$SD = \sqrt{\frac{\Sigma f_ix_i^2}{N} - (\bar{X})^2}$$

$$= \sqrt{\frac{48150}{70} - (23.42)^2} = 11.80$$

$$\boxed{\begin{aligned} Var &= (SD)^2 \\ &= (11.80)^2 \\ &= 139.24 \end{aligned}}$$

$$\boxed{\text{Coefficient of SD} = \frac{\sigma}{\bar{X}}} = \frac{11.80}{\frac{\Sigma fx}{N}} = \frac{11.80}{23.42} = 0.5038$$

$$\boxed{\text{Coefficient of variance} = \frac{SD}{\bar{X}} \times 100} = \frac{11.80}{23.42} \times 100$$

$$= 50.38$$

Mean Deviation

marks	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60	60 - 70
No. of stu	8	6	12	5	2	7

Ans :-

marks	No. of stud (f)	mid point (x)	fx	$\lvert x-\bar{x}\rvert$	$f(x-\bar{x})$
10 - 20	8	15	120	22	176
20 - 30	6	25	150	12	72
30 - 40	12	35	420	2	24
40 - 50	5	45	225	8	40
50 - 60	2	55	110	18	36
60 - 70	7	65	455	28	196
	$N: \Sigma f = 40$	Σx	$\Sigma f_i x_i = 1480$		$\Sigma f(x-\bar{x}) = 544$

$$\bar{x} = \frac{\Sigma f_i x_i}{N} = \frac{1480}{40} = 37$$

$$\boxed{\text{Mean deviation} = \cancel{\Sigma f} \frac{\Sigma f(x-\bar{x})}{N}} = \frac{544}{40} = 13.6$$

$$\boxed{\text{Coefficient of MD} = \frac{MD}{\bar{x}} = \frac{13.6}{37}} = 0.367$$

Box-Whisker Plot : Test / 5 N. Summary.

Q. { 4, 5, 2, 1, 7, 5, 4, 4, 2, 15, 6, 2, 5, 5, 1, 3, 2
8, 3, 7 }

Ans:-
1] Arrange it in increasing order
{ 1, 1, 2, 2, 2, 2, 3, 3, 4, $\boxed{4\ 4}$, 5 5 5 5 6 7 7 8, 15 }
central value

median $= Q_2 = \frac{N}{2} = \frac{20}{2}^{th} = 10^{th}$ and 11

$= \frac{4+4}{2} = 8/2 = 4$

$\boxed{Q_2 = 4}$

$Q_1 = 10^{th}$ No. $= \frac{10}{2} = 5^{th}$ and 6^{th}

$\frac{2+2}{2} = 4/2 = 2$

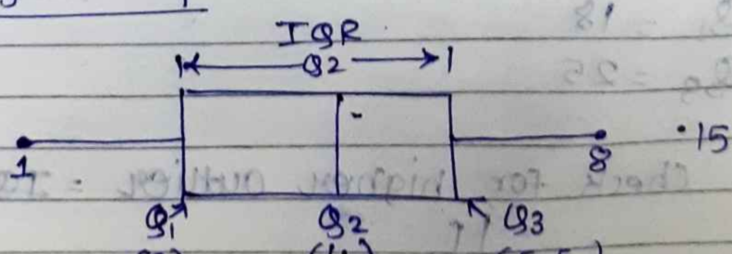
$\boxed{Q_2 = 2}$

$Q_3 = 10^{th}$ NO. $= 10/2 = 5^{th}$ and 6^{th}

$= \frac{5+6}{2} = 11/2 = 5.5$

$\boxed{Q_3 = 5.5}$

Ans:-



$IQR = Q_3 - Q_1 = 5.5 - 2 = 3.5$
lowest Datum $= Q_1 - 1.5(IQR) = -3.25$
Highest Datum $= Q_3 + 1.5\ (IQR) = 10.75$

Q3 / upper quartile

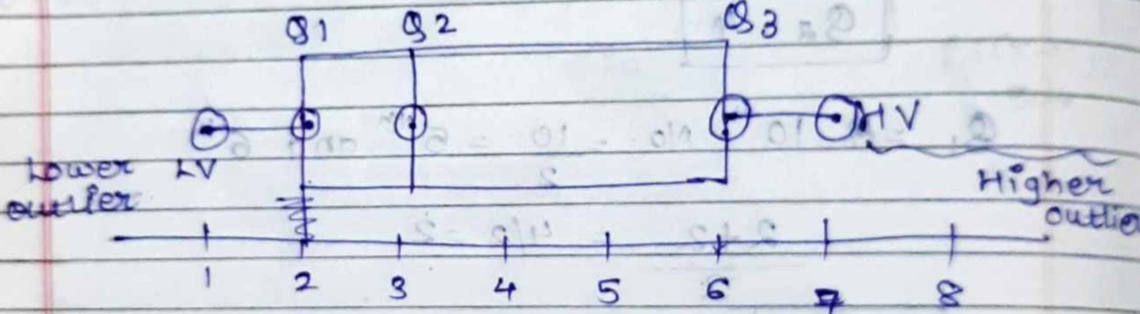
Eg 2

1, 2, 2, 3, 5, 6, 7

↓ Q1 lower Quartile

↓ median Q2

lowest value = 1
Q1 = 2
median / Q2 = 3
upper Quartile / Q3 = 6
Highest value = 7



Data is positively skewed.

Eg 3

Arrange in asc

~~22, 25, 17, 19~~ 17, 17, 18, 19, 20, 22, 23, 25, 33, 64

$Q_2 = \frac{20+22}{2} = 21$

$Q_1 = 18$

$Q_3 = 25$

Check for higher outlier = IQR = $Q_3 - Q_1$
= 25 - 18
= 7

$= Q_3 + [1.5 * IQR]$
$= Q_3 + [1.5 * 7]$
$= 35.5$

So, here any data greater than 35.5 will be higher outlier

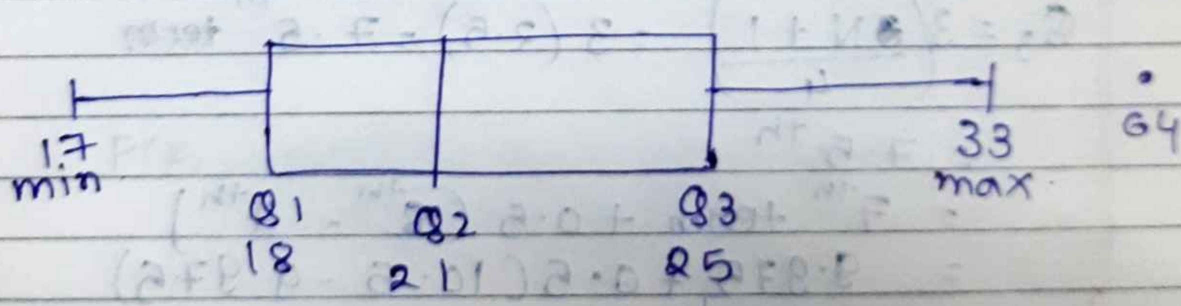
Lower outlier $= Q_1 - [1.5 * IQR]$

$= Q_1 - [1.5 \times 7]$

$= 7.5$

So, any value less than 7.5 will be lower outlier.

So, $Q_1 = 18$	min value $= 17$
$Q_2 = 21$	max value $= 33$
$Q_3 = 25$	higher outlier $= \underline{\underline{64}}$
$IQR = 7$	



Eg.

3.8, 5.8, 5.99, 6.0, 6.7, 7.0, 9.975, ~~10~~ 10.5, 20.0

$N = 9$

$\boxed{Q_2 = 6.7}$

$Q_1 = \left(\frac{N+1}{4}\right) = \left(\frac{9+1}{4}\right) = \frac{10}{4} = 2.5^{th}$ term

2.5^{th}

2^{nd} term $+ 0.5\,(3^{rd} - 2^{nd})$

$= 5.8 + 0.5\,(5.99 - 5.8)$

$\boxed{Q_1 = 5.895}$

$Q_3 = 3\left(\frac{N+1}{4}\right) = 3\,(2.5) = 7.5^{th}$ term.

$= 7.5^{th}$

$= 7^{th}$ term $+ 0.5\,(8^{th} - 7^{th})$

$= 9.975 + 0.5\,(10.5 - 9.975)$

$\boxed{Q_3 = 10.2375}$

$IQR = Q_3 - Q_1 = 4.3425$

~~lowest~~ Highest outlier -

Q_3 [1.5 * IQR]

$= 10.2375\,(1.5 * 4.3425)$

$= 66.6845$

Lower ~~Quarti~~ outlier

$= Q_1\,[1.5 * IQR]$

$= 5.895\,(1.5 * 4.3425)$

$= 38.3985$

UNIT 4

* Random Experiment:-
  Rolling of dice

* Sample space
  $S = \{1, 2, 3, 4, 5, 6,\}$
  $n(S) = 6$

* Event : Possible outcome : 1 is getting of upper side
  $A = \{1\}$

$$P(A) = \frac{n(A)}{n(S)} = \frac{1}{6}$$

Event 2 : no. is greater than 4.
$A = \{5, 6\}$
$n(A) = 2$

$$P(A) = \frac{n(A)}{n(S)} = \frac{2}{6} = \frac{1}{3}$$

* $\boxed{\text{Algebra of Event}}$

(1) Complimentary event:-
'not' A' Event

(2) $A \cup B$ either A or B or both.

(3) $A \cap B$ only common ele from A and B

* Discrete Random Variable :
RV. which can be only finite no of values in finite obser^n interval $\int$

* Continuose Random
RV which can be infinite no of values. Σ

Probability Mass function :
PMF should be

Q.1]
i) $f(x) \geq 0$
ii) $\Sigma f(x_i) = 1$
iii) $f(x_i) = P(x = x_i)$.

Q.1] A coin is flipped twice and random variable X is the no. of Heads.

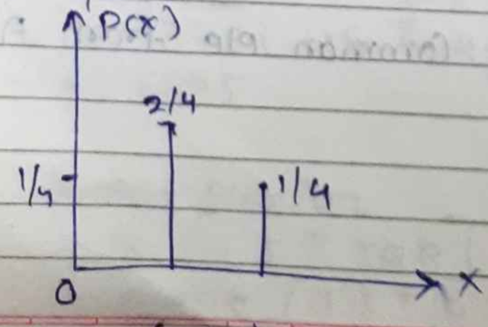
$S = \{HH, HT, TH, TT\}$.

Random variable $X = \{0, 1, 2\}$

PMF.

X	0	1	2
P(X)	1/4	1/2	1/4

$P(X=2) = 1/8$



Graph

$S = \{HH, HT, TH, TT\}$.

Event of getting 1 Head

$\boxed{\text{pmf}}$

X	Event	Probability mass fn. $f(x) = P(X = x)$
$x = 0$	TT	$f(0) = P(X = 0) = 1/4$
$x = 1$	TH, HT	$f(x = 1) = 2/4 = 1/2$
$X = 2$	HH	$f(x = 2) = 1/4$

$$\text{PMF} = P(x) = \begin{cases} P(X = x_i) ; & X = x_i \\ 0 & x \neq x_i \end{cases}$$

$$P(x) = \begin{cases} P(X = 0) = 1/4 ; & x = 0 \\ P(x = 1) = 1/2 ; & x = 1 \\ P(x = 2) = 1/4 ; & x = 2 \\ 0 ; & x \neq x \end{cases}$$

Karl Pearson coefficient of skewness

Invest	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Frequency	12	18	20	15	10	3	2

C.I	f_i	x_i	$f_i x_i$	x_i^2	$f_i x_i^2$	cf
10-20	12	15	180	225	2700	12
20-30	18	25	450	625	11736	
30-40	20	35	700	1225	24500	
40-50	15	45	675	2025	30375	
50-60	10	55	550	3025	30250	
60-70	3	65	195	4225	12675	
70-80	2	75	150	5625	11250	
$\Sigma f = N$ 80			$\Sigma f_i x_i =$ 2900	$\Sigma x_i^2 =$	$\Sigma f_i x_i^2$ 123486	

$$\text{Coefficient of } S_{kp} = \frac{\text{mean} - \text{mode}}{SD}$$

$$\boxed{\text{mean}} = \frac{\Sigma f_i x_i}{N} = \frac{2900}{80} = 36.25$$

$\boxed{\text{mode}} =$ $f_0 = 18$ $\quad l_1 = 30$

$f_1 = 20$ $\quad l_2 = 40$

$f_2 = 15$

$$l_1 + (l_2 - l_1)\left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right)$$

$$= 30 + (40-30)\left(\frac{20-18}{2(20) - 18 - 15}\right)$$

$$= 40\left(\frac{20-18}{40-18-15}\right) = 11.42$$

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

$$= \sqrt{\frac{123486}{80} - (36.25)^2}$$

$$= 15.14$$

$$\boxed{\text{coefficient of } (S_{kp}) = \frac{\text{mean} - \text{mode}}{SD}}$$

$$= \frac{36.25 - 11.42}{15.14}$$

$$= 1.64$$

* Bowley's coefficient of skewness

$$\cancel{(S_{kB}) = \frac{Q_3 + Q_1 - 2\,\text{Median}}{Q_3 - Q_1}} \quad \cancel{(S_{kB}) = (Q_3 - Q_2)}$$

marks	no. of stu (f)	cf	
0-10	8	8	
10-20	10	18	$\rightarrow Q_1$
20-30	11	29	$\rightarrow Q_2$
30-40	4	33	$\rightarrow Q_3$
40-50	3	36	
	$\Sigma f = 36$		

$$Q_3 = 3\left(\frac{N}{4}\right.$$

$$\boxed{(S_{kB}) = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)}}$$

$N = 36$

To find Q_1

$Q_1 \quad Q_2 \quad Q_3$

$$\frac{N}{4} = \frac{36}{4} = 9$$

$$Q_1 = \frac{l_1 + (l_2 - l_1)\left(\frac{N}{2} - c.f\right)}{f}$$

$l_1 = 10$ $\quad N/2 = 9$ $\quad f = 10$

$l_2 = 20$ $\quad c.f = 8$

$$= \frac{10 + (20-10)(9-8)}{10}$$

$$= \frac{10 + 10(1)}{10}$$

$$\boxed{Q_1 = 2}$$

for Q_2,

$$\frac{N}{2} = \frac{36}{2} = 18$$

$$Q_2 = \frac{l_1 + (l_2 - l_1)\left(\frac{N}{2} - c.f\right)}{f}$$

$l_1 = 20$ $\quad N/2 = 18$ $\quad c.f = 18$

$l_2 = 30$ $\quad f = 11$

$$= \frac{20 + (30-20)(18-18)}{11}$$

$$\boxed{Q_2 = 1.81}$$

For, Q_3 :

$$\frac{3N}{4} = \frac{3(36)}{4} = 27$$

$$Q_3 = \frac{l_1 + (l_2 - l_1)\left(3N/4 - c.f\right)}{f}$$

$l_1 = 30$ $\quad 3N/4 = 27$ $\quad f = 4$.

$l_2 = 40$ $\quad c.f = 29$

$$Q_3 = \frac{30 + (40-30)(27-29)}{4}$$

$$Q_3 = 2.5$$

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

$$= \frac{(2.5 - 1.81) - (1.81 - 2)}{(2.5 - 1.8) + (1.81 - 2)}$$

$$= \frac{0.9}{0.5}$$

$$\boxed{(Sk)_B = 1.8}$$

Regression Analysis :-

x	y	xy	x^2	y^2
10	40	400	100	1600
12	38	456	144	1444
13	43	559	169	1849
12	45	540	144	2025
16	37	592	256	1369
15	43	645	225	1849
$\Sigma x =$ 78	$\Sigma y =$ 246	$\Sigma xy =$ 3192	$\Sigma x^2 =$ 1038	$\Sigma y^2 =$ 10136

(A) Regression of y on x.

$$\Sigma y = a_0 N + a_1 \Sigma x$$
$$\Sigma xy = a_0 \Sigma x + a_1 \Sigma x^2$$

$$246 = a_0 6 + a_1 \cdot 78$$
$$3192 = a_0 78 + a_1 \cdot 1038$$

$$a_0 = 44.25$$
$$a_1 = -0.25$$

$$y = a_0 + a_1 \cdot x$$
$$\boxed{y = 44.25 + (-0.25)x}$$

Random variable

Discrete

$P(X=x_i) = p_i$

i] $p_i \geq 0, \forall i$

ii] $\sum p_i = 1$

Continuiose

$P(X=x_i) = f(x_i)$

- i] $f(x) \geq 0, \forall x$

ii] $\int_{-\infty}^{\infty} f(x)\,dx = 1$

- $\int_a^b f(x)dx = P(a < x < b)$

- $P(a < x \leq b) = P(a \leq x < b)$
$= P(a \leq x \leq b) = P(a < x \leq b)$

Distribution function.

$$F_x(x) = \begin{cases} P(X \leq x) & , x\text{ - discrete} \\ \int_{-\infty}^{x} f(x)\,dx & , X\text{ - conti} \end{cases}$$

$$\text{mean} = \begin{cases} \sum x p_i & , x\text{ - discrete} \\ \int_{-\infty}^{\infty} x f(x)\,dx & , x\text{ - conti} \end{cases}$$

$$\text{variance} = \begin{cases} \sum (x-\bar{x})p_i \\ \int_{-\infty}^{\infty} (x-\bar{x})^2 f(x)\,dx \end{cases}$$

eg:

Q.1) Check whether above is pdf or not? and find mean & variance

→ $\boxed{f(x) = 6x(1-x)}$

Ans:- 1] (i)

$f(x) = 6x - 6x^2$

$= 6(x - x^2) \geq 0$

as $6 > 0$, $x^2 < x$,

(ii) $\int_{-\infty}^{\infty} f(x) \cdot dx = \int_0^1 (6x - 6x^2) dx$

$= \left[\frac{6x^2}{2} - \frac{6x^3}{3}\right]_0^1$

$= [3x^2 - 2x^3]_0^1$

$= 3 - 2$

$= 1$

(iii) mean $= \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx$

$= \int_0^1 x(6x)(1-x) \cdot dx$

$= \int_0^1 (6x^2 - 6x^3) \cdot dx$

$= \left[\frac{6x^3}{3} - \frac{6x^4}{4}\right]_0^1$

$= \left[2x^3 - \frac{3}{2}x^4\right]_0^1$

$= 2 - \frac{3}{2}$

$\boxed{\text{mean} = 1/2}$

(iv) variance

$Var(x) = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x) \cdot dx$ — $\bar{x}$: mean

$= \int_0^1 \left(x - \frac{1}{2}\right)^2 6x(1-x) \cdot dx$

$= \int_0^1 \left(x^2 - x + \frac{1}{4}\right)(6x - 6x^2) dx$

$= \int_0^1 \left[6x^3 - 6x^2 + \frac{3x}{2} - 6x^4 + 6x^3 - \frac{3}{2}x^2\right] dx$

$= \int_0^1 \left[12x^3 - 6x^4 - \frac{15x^2}{2} + \frac{3x}{2}\right] dx$

$= Var(x) = \left[3x^4 - \frac{6}{5}x^5 - \frac{5}{2}x^3 + \frac{3x^2}{4}\right]_0^1$

$= \left[3 - \frac{6}{5} - \frac{5}{2} + \frac{3}{4}\right]$

$= 1/20$

$\boxed{Var(x) = 1/20}$

discrete

mean $= \Sigma x p_i$

$var(x) = \Sigma (x - \bar{x})^2 p_i$

1) A function

$$f(x) = \begin{cases} 0 & , x < 2 \\ (2x+3)k, & 2 \le x \le 4 \\ 0 & , x > 4 \end{cases}$$

is pdf, then find k.

Ans:-

By taking integration.

$$\int_{-\infty}^{\infty} f(x) \cdot dx = 1$$

$$\Rightarrow \int_{2}^{4} k(2x+3) dx = 1$$

$$\Rightarrow k\left[\frac{2x^2}{2} + 3x\right]_2^4 = 1$$

$$\Rightarrow k\left[x^2 + 3x\right]_2^4 = 1$$

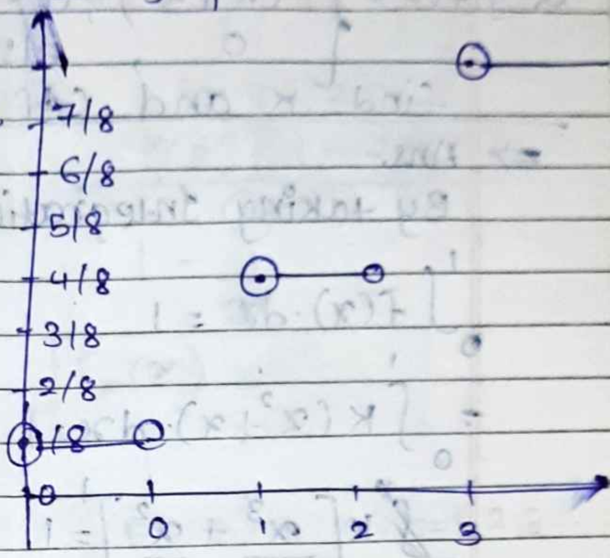
$$\Rightarrow k[(16+12) - (4+6)] = 1$$

$$\Rightarrow k\ [28 - 10] = 1$$

$$\Rightarrow \cancel{k = (} \cdot k(18) = 1$$

$$\Rightarrow \boxed{k = 1/18}$$

cdf graph



Pdf

* Draw cdf (Cummulative).

$X = x$	0	1	2	3	
$P(X = x)$	1/8	3/8	3/8	1/8	⇒ Pdf
$F_X(x)$	1/8	4/8	7/8	8/8	⇒ cdf

X ↓ Random Variable

$F_x(0) = P(X \le 0)$

$F_x(1) = P(X \le 1) = P(X = 0) +$

$P(X = 1)$

Q.2] $f(x) = \begin{cases} k(x^2+x), & 0 < x \le 1 \\ 0, & \text{else} \end{cases}$

find k and cdf

$\Rightarrow$ Ans:-

By taking Integration

$$\int_0^1 f(x)\cdot dx = 1$$

$$= \int_0^1 k(x^2+x)\cdot dx = 1$$

$$= k\left[\frac{x^3}{3} + \frac{x^2}{2}\right]_0^1 = 1$$

$$= k\left[\frac{1}{3} + \frac{1}{2}\right] = 1$$

$$\cancel{= k\left[\frac{5}{6}\right] = 1}$$

$\therefore k =$

$$k\left[\frac{2+3}{6}\right] = 1$$

$\frac{1}{3} + \frac{1}{2}$

$\frac{2+3}{6}$ $\quad k\left(\frac{5}{6}\right) = 1$

$$\boxed{k = \frac{6}{5}}$$

ii] To find cdf

Continuose RV

$$F_X(x) = \int_{-\infty}^{x} f(x)\cdot dx$$

1) $x \in (-\infty, x)$

$$\boxed{F_X(x) = \int_{-\infty}^{x} f(x)\cdot dx = 0}$$

2) $x \in [0, 1]$

$$F_X(x) = \int_{-\infty}^{x} f(x)\cdot dx = \int_0^x \frac{6}{5}(x^2+x)\cdot dx$$

$$\boxed{F_X(x) = \frac{6}{5}\left[\frac{x^3}{3} + \frac{x^2}{2}\right]}$$

3) $x \in [1, \infty]$

$$F_X(x) = \int_{-\infty}^{x} f(x)\cdot dx$$

$$\boxed{F_X(x) = 0}$$

Ans:-

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{6}{5}\left(\frac{x^3}{3} + \frac{x^2}{2}\right), & 0 \le x \le 1 \\ 0, & x > 1 \end{cases}$$

i] $\frac{d}{dx} F_X(x) = f(x)$

$$F_X(x) = \frac{6}{5}\left[\frac{x^3}{3} + \frac{x^2}{2}\right]$$

$$\frac{d}{dx} F_X = \frac{6}{5}\left[x^2 + x\right]$$

8] X - discrete RV with range
$R_X = \{1, 2, 3, \dots\}$ &
Pmf = $P_X(k) = \frac{1}{2^k}$, $k = 1, 2 \dots$

a] find and plot cdf of $f(x)$.
b] find $P(2 < X \leq 5)$.
c] find $P(X \geq 4)$

Ans :- To check Pmf
i] $p_i > 0$, $\frac{1}{2k^2}$ is '+'.

ii] $\sum p_i = \sum P_X(k) = \sum_{k=1}^{\infty} \frac{1}{2^k}$

$= \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \dots$

common diff $= \frac{1}{2}$

$= \frac{a(1-r^n)}{1-r}$ ~~will be~~ Do not consider

$= \frac{a}{1-r} = \frac{1/2}{1-1/2} = 1$

a] find and plot cdf for $f(x)$
$F_X(x) = P(X \leq x)$
$= \sum_{k=1}^{\infty} P_X(k)$

$F_X(x) = \sum \frac{1}{2^k}$

$= \frac{a(1-r^k)}{1-r}$

$= \frac{(1/2)\left[1 - \left(\frac{1}{2}\right)^k\right]}{1 - (1/2)}$

$= \frac{(1/2)[1 - 1/2^k]}{(1/2)}$

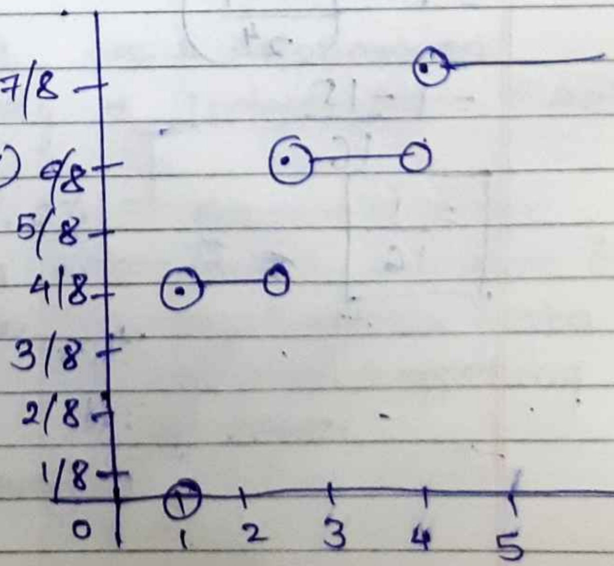
$= 1 - \frac{1}{2^k}$

$\boxed{F_X(x) = \frac{2^k - 1}{2^k}}$

$F_X(k) = \quad , k = 1, 2, 3 \dots$
$k = 1$ $F_X(1) = \frac{2-1}{2^k} = \frac{2-1}{2^1}$
$= \frac{1}{2}$

$k = 2$, $F_X(2) = \frac{2^2 - 1}{2^2} = \frac{3}{4}$

$k = 3$, $F_X(3) = \frac{2^3 - 1}{2^3} = \frac{7}{8}$



cdf

b] find $P(2 < x \le 5)$

$= P(x=3) + P(x=4) + P(x=5)$

$= \frac{7}{8} + \frac{1}{16} + \frac{1}{32}$

$= \frac{7}{32}$

$\boxed{P(a < X < b) = F_X(b) - F_X(a)}$

$P(2 < X \le 5) = F_X(5) - F_X(2)$

$= \left(\frac{2^5 - 1}{2^5}\right) - \left(\frac{2^2-1}{2^2}\right) =$

$\boxed{= \frac{7}{32}}$

c] Find $P(X > 4)$

$= 1 - P(X \le 4)$

$= 1 - F_X(4)$

$= 1 - \left(\frac{2^4 - 1}{2^4}\right)$

$= 1 - \frac{15}{16}$

$\boxed{= \frac{1}{16}}$

Conditional probability

$P(A/B) = \frac{P(A \cap B)}{P(B)}$, $P(B) \neq 0$

$P(B/A) = \frac{P(A \cap B)}{P(A)}$, $P(A) \neq 0$

$P(A \cap B) = P\left(\frac{A}{B}\right) \cdot P(B)$

Baye's Theorem

B_1, B_2, B_n - events

$P(B_i/A) = \frac{P(A/B_i) \cdot P(B_i)}{\sum_{i=1}^{n} P(A/B_i) \cdot P(B_i)}$,

$P(B_i) \neq 0 \; ; \forall i$

Q.1 Box 1 contains 3 white 4 red flowers. Box 2 contains 5 white and 6 red flowers. One flower is chosen as a random from one of the box and it is found to be white. Find the probability that it was drawn from box 1

Ans:-

U_1: flower drawn from box 1
U_2: " box 2.
W: white flower
R: Red flower

$$P(U_1/W) = \frac{P(W/U_1)\cdot P(U_1)}{P(W/U_1)\cdot P(U_1) + P(W/U_2)\cdot P(U_2)}$$

$P(U_1) = P(U_2) = 1/2$

$P(W/U_1) = 3/7\ [3/3+4]$

$P(W/U_2) = 5/11\ [5/5+6]$

By substituting we get:

$$= \frac{(3/7)\cdot(1/2)}{(3/7)(1/2) + (5/11)(1/2)}$$

$$\boxed{= \frac{33}{68}}$$

Q.2] In bolt factory, machine A, B, C manufactures respectively 25%, 35% and 40% of the total. If the output 5%, 4%, 2% are defective bolts a bolt is drawn at random from the product and is found to be defective and what is the probability that is manufactured by machine B.

Ans:-

$$P(B/D) = \frac{P(D/B)\cdot P(B)}{P(D/B)P(B) + P(D/A)\cdot P(A) + P(D/C)\cdot P(C)}$$

$P(D/B) = 4/100 = 0.04$ (D → Defective, B → Bolt)

$P(B) = 0.35$ i.e. $35/100$

$P(D/A) = 0.05$ i.e. $5/100$

$P(A) = 0.25$ i.e. $25/100$

$P(D/C) = 0.02$ i.e. $2/100$

$P(C) = 0.40$ i.e. $40/100$

$$P(B/D) = \frac{0.04 \times 0.35}{(0.04)(0.35) + (0.05)(0.25) + (0.02)(0.4)}$$

$= P(B/D) = 0.41$

~~* Multiplication~~

* Independent Events

↓

⇒ Happening of one doesn't influence in any way the probability of happening of other.

* Multiplication law of probability

The probability of two independent events will happen is the product of probability of their happening respectively.

$$\boxed{P(AB) = P(A) \cdot P(B)}$$

* An article manufactured by Company consist of 2 parts A and B. In the process of manufacture of part A. 9 (9/100) out of 100 are to be defective. similarly 5 out of 100 (5/100) are likely to be defective in the manufacture of part B. calculate the probability that assembled article won't be defective.

Ans:- Given A and B are independent

P (Non-defective) = P (Part. A non defective) × P (part B. non-defective).

$$= \left(\frac{91}{100}\right)\left(\frac{95}{100}\right)$$

$$\boxed{= 0.8645}$$

Q2:- Three groups of children Contains respectively, [3 girls, 1 boy], [2 girls, 2 boys], [1 girls, 3 boys]. One child is selected at a random from each group find the chance of selecting 1 girl and 2 boys.

Ans :-

$$\overset{g}{(3/4)} \times \overset{b}{(2/4)} \times \overset{b}{(3/4)} +$$

$$\overset{b}{(1/4)} \times \overset{g}{(2/4)} \times \overset{b}{(3/4)} +$$

$$\overset{b}{(1/4)} \times \overset{b}{(2/4)} \times \overset{g}{(1/4)}$$

$= 13/32$

Q.3 There are 6 positive and 8 negative numbers. 4 no are chosen at random without replacement and multiplied what is the probability that product is positived number.

	6 positive	8 Negative
I	4	0
II	2	2
III	0	4.

$$p = \frac{{}^6C_4 \cdot {}^8C_0 + {}^6C_2 \cdot {}^8C_2 + {}^6C_0 \cdot {}^8C_4}{{}^{14}C_4}$$

$$= \boxed{\frac{505}{1001}}$$

* Diffent

There are ${}^nC_x$ ways to get x many success from n trials

$P(x) = {}^nC_x (p^x)(q^{n-x})$

n = no of trials

p = Prob of success

q = Prob of failure

x = no of success

Q.1] Find the probability of getting 4 heads in 6 tosses of fair coin.

$x = 4$; $n = 6$; $p = 1/2$, $q = 1/2$

$P(x=4) = {}^nC_x \, p^x \cdot q^{n-x}$

$= {}^6C_4 (1/2)^4 (1/2)^{6-4}$

$= \left(\frac{1}{2}\right)^6 \cdot {}^6C_4 =$

$= \left(\frac{1}{2}\right)^6 \cdot 15$

$\boxed{= \frac{15}{64}}$

01/04/2024

* Binomial Distribution

$E(x) = np$

$SD \ (\sigma) = \sqrt{npq}$

$var(x) = npq$

Q. A factory turns out an article by mass production method from the past expe. it is found that 20 articles on an avg. are rejected out of every batch of 100. find the mean & variance of the rejected article.

Ans:-

$P = \frac{20}{100} = 0.2$

$E(x) = np = 100 \times 0.2 = 20$

$var(x) = npq = 100 \times 0.2 \times 0.8 = 16$

Q. 30000 test were performed in Corona panda. out of which 18000 are '+'. Today we are doing 10 test what is the probability that atmost 8 patient will be Corona +.?

$P(X \le 8) = 1 - P(X=9)$
$- P(X=10)$
$P(X=9) = {}^{10}C_9 (p)^9 (q)^1$
$= 10 (0.6)^9 (0.4)^1$
$P(X=10) = {}^{10}C_{10} (p)^{10} \cdot (q)^0$
$= 1 (0.6)^{10}$

$P(X \le 8) =$

Q. Probability of a man hitting a target is 1/3.
a] If he fires 5 times, what is probablity of his hitting the target atleast twice.
b] How many times must he fire so that the prob of hitting his target atleast once is more than 90%

Ans:
→ a] $P(X \ge 2) = 1 - P(X<2)$
$= 1 - P(X=0) - P(X=1)$
$= 1 - {}^5C_0 p^0 q^5 - ({}^5C_1 \cdot p^1 q^4)$
$= 1 - \left(\frac{2}{3}\right)^5 \left(5\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^4\right)$

$= 0.33$

b] $P(X \ge 1) = 0.9$
$1 - P(X<1) = 0.9$
$1 - P(X=0) = 0.9$
$1 - {}^nC_0 p^0 q^n = 0.9$
$1 - \left(\frac{2}{3}\right)^n = 0.9$

$\left(\frac{2}{3}\right)^n = 0.1$

$\log\left(\frac{2}{3}\right)^n = \log(0.1)$

$n \log(0.6) = \log(0.1)$
$n = \frac{\log(0.1)}{\log(0.6)}$

* Poison Distribution

$P(X=x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$

$\lambda = \text{mean} = np$
$x = \text{no. of success.}$

$P(X=4) = \frac{e^{-2} (2)^4}{4!}$

$\lambda = np$
$\frac{\lambda}{n} = p$

$E(X) = var(X) = \lambda = np$

Q. Find out a car hire from has two car which hires out day by day the no of demands for a car on each day is distributed as poison median with mean is 1.5 calculate proportion of days on which the car is used or some demand is refused.

Ans:-

$$P(X=x) = \frac{e^{-\lambda} \cdot \lambda^{x}}{x!}$$

$\lambda = 1.5$ $\qquad = \frac{e^{-1.5}(1.5)^{x}}{x!}$

x = no. of cars

(i) Neither car is used

$$P(X=0) = \frac{e^{-1.5}(1.5)^{0}}{0!}$$

$$= 0.2231$$

(ii) Some demand is refused

$$P(X>2) = 1 - P(X\leq 2)$$

$$= 1 - P(X=0)\ P(X=1) - P(X=2)$$

$$= 0.1912$$

Q. A hospital which holds receives an avg. of 4 emergency calls in 10 min interval what is prob. that there are 2 emergency calls and there are 3 emergency calls in interval of min ~~are 3~~

(i) There are atmost 2 calls.

$$P(X\leq 2) = P(X=0) + P(X=1) + P(X=2)$$

$$= \left(\frac{e^{-4}(4)^{0}}{0!}\right) + \left(\frac{e^{-4}(4)^{1}}{1!}\right) + \left(\frac{e^{-4}(4)^{2}}{2!}\right)$$

$$= 0.238$$

(ii) $P(X=3) = \frac{e^{-4}(4)^{3}}{3!}$

$$= 0.195$$

Q. Find out a car hire from has two car which hires out day by day the no of demands for a car on each day is distributed as poison median with mean is 1.5 calculate proportion of days on which the car is used or some demand is refused.

Ans:-

$$P(X=x) = \frac{e^{-\lambda} \cdot \lambda^{x}}{x!}$$

$\lambda = 1.5$

$$= \frac{e^{-1.5}(1.5)^{x}}{x!}$$

x = no. of cars

(i) Neither car is used

$$P(X=0) = \frac{e^{-1.5}(1.5)^{0}}{0!}$$

$$= 0.2231$$

(ii) Some demand is refused

$$P(X>2) = 1 - P(X \le 2)$$

$$= 1 - P(X=0)\ P(X=1) - P(X=2)$$

$$= 0.1912$$

Q. A hospital which holds receives an avg. of 4 emergency calls in 10 min interval what is prob. that there are 2 emergency calls and there are 3 emergency calls in interval of min ~~are 2~~

(i) There are atmost 2 calls.

$$P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$$

$$= \left(\frac{e^{-4}(4)^{0}}{0!}\right) + \left(\frac{e^{-4}(4)^{1}}{1!}\right) + \left(\frac{e^{-4}(4)^{2}}{2!}\right)$$

$$= 0.238$$

(ii) $P(X=3) = \frac{e^{-4}(4)^{3}}{3!}$

$$= 0.195$$

3/04/2024

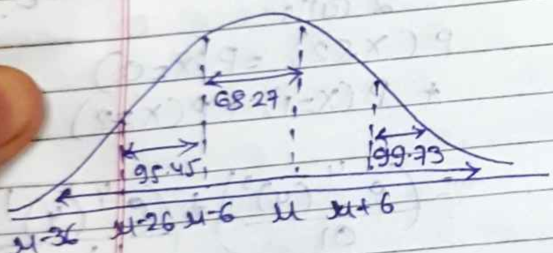
* ~~Testimonial~~ normal
Standard ~~module~~
distribution

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$-\infty < x < \infty$

$-\infty < \mu < \infty$

$\sigma^2 > 0$



Q.1] Find the probability of random variable following N.D will take on a value bet 0.87 and 1.28.

0.87 and 1.28

$Z \sim N(0,1)$

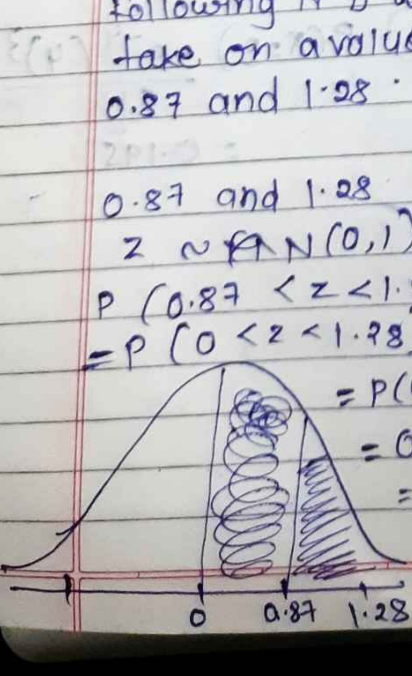
$P(0.87 < z < 1.28)$

$= P(0 < z < 1.28)$

$= P(0 < z < 0.87)$

$= 0.3997 - 0.3078$

$= 0.0919$



6/4/24

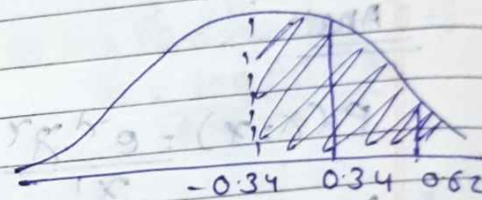
Q. Find probability that a random var having SD will take a val. betn -0.34 and 0.62

$P(-0.34 < z < 0.62)$

$\Rightarrow P(0 < z < 0.34) + P(0 < z < 0.62)$

$= 0.1331 + 0.2324$

$= 0.3655$



Q.2) For normal variate with mean 2.5 & SD 3.5 find the probability that

i) $P(2 \le X \le 4.5)$

ii) $P(-1.5 \le X \le 5.1)$

$\Rightarrow z = \frac{X-\mu}{\sigma} = \frac{X-2.5}{3.5}$

i) at $X=2$, $z = \frac{2-2.5}{3.5}$

$= -0.14$

6/4/2

at $x = 4.5$, $z = \frac{4.5 - 2.5}{3.5}$

$= \underline{0.57}$

$P(2 \le x \le 4.5) =$

$= P(-0.14 \le z \le 0.57)$

$= P(0 \le z \le 0.14) +$

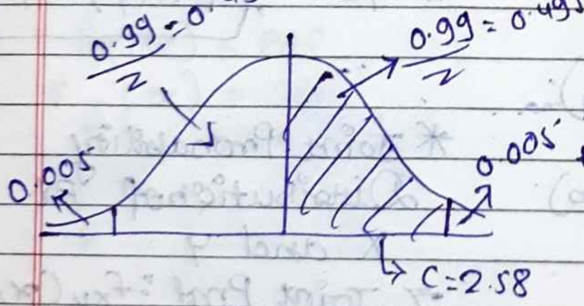
$P(0 \le z \le 0.57)$

$= 0.0557 + 0.2157$

$= \underline{\underline{0.2714}}$

Q.2) If $z \sim N(0,1)$. Find C such that $P(|z| > C) = 0.01$

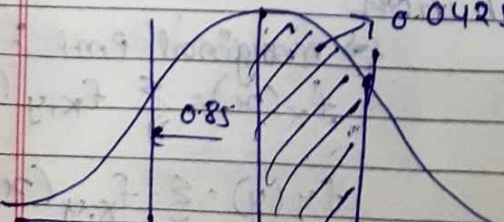
$\frac{0.99}{2} = 0.495$



$P(2.58) = 0.495$

$1 - 0.01 = 0.99$

Q. If $z \sim N(0,1)$ find c such that $P(|z| > c) = 0.85$



$P(1.44) = 0.425$, $\boxed{C = 1.44}$

* Two dimensional random variable

→ Experiment : Tossing a coin twice

$\Omega = \{HH, HT, TH, TT\}$

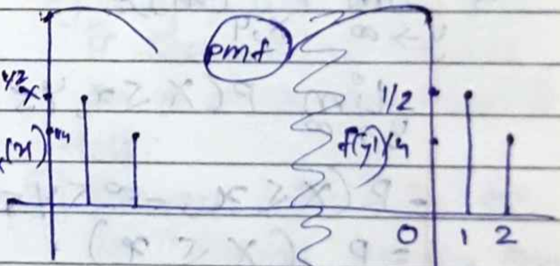
X : No. of heads

Y : No. of tails

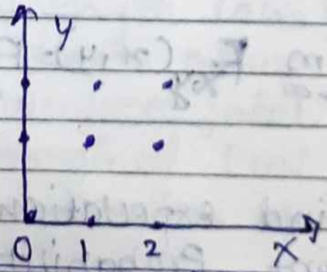
$Y(HH) = 0$, $Y(HT) =$

$Y(TH) = 1$, $Y(TT) = 2$

$X = \{0, 1, 2\}$, $Y = \{0, 1, 2\}$



Both graphs are diff



$\boxed{\text{Joint Probability distribution}}$

$y \backslash x$	0	1	2
0	$P(x=0, y=0) = 0$	$P(x=1, y=0) = 0$	$P(x=2, y=0) = 1/4$
1	$P(x=0, y=1) = 0$	$P(x=1, y=1) = 1/2$	$P(x=2, y=1)$ 0
2	$P(x=0, y=2) = 1/4$	$P(x=1, y=2) = 0$	$P(x=2, y=2) = 0$
Σ	$P(x=0)$	$P(x=1)$	$P(x=2)$

Date
Page

Joint pmf

$$\sum_x \sum_y P(X=x, Y=y) = 1$$

$$\sum_x P(X=x, Y=y) = P(Y=y)$$

$$\sum_y P(X=x, Y=y) = P(X=x)$$

Q. $F_{X,Y}(x,y) = P(X \le x, Y \le y)$

$f_X(x) = \sum_y (f_{x,y}(x,y)$

$f_Y(y) = \sum_x f_{X,Y}(x,y)$

where $f(xy) = (xy)$ is joint pmf

Q. $\lim_{y \to \infty} F_{X,Y}(x,y) =$

$\lim_{y \to \infty} P(X \le x; Y \le y)$

$= P(X \le x, -\infty < Y \le \infty)$

$= P(X \le x)$

$= F_X(x)$

$$\boxed{\lim_{x \to \infty} F_{X,Y}(x,y) = F_Y(y)}$$

* find expectation of previous joint probability distribution
find $E(XY)$

$= E(XY) = \sum_{x=0}^{2} \sum_{y=0}^{2} (xy) P(X=x, Y=y)$

$= \sum_{y=0}^{2} 0 + \cancel{P}\, y P(X=1, Y=y)$

$+ 2y P(X=2; Y=y)$

$E(XY) = 0 + 1 \cdot P(X=1, Y=1)$

$+ 2(1) P(X=2; Y=1)$

$+ (2) P(X=1, Y=2) +$

$2(2) P(X=2, Y=2)$

$= (1/2) + 2(0) + 2(0) + 4(0)$

$= 1/2$, $\boxed{E(XY) = 1/2}$

* Joint Probability Distribution of RV X and Y.

→ Joint Pmf : $f_{XY}(x,y) = P(X=x, Y=y)$

→ Joint Distribution f^n = $F_{X,Y}(x,y) = P(X \le x, Y \le y)$

→ marginal pmf :

$f_X(x) = \sum_y f_{X,Y}(x,y)$

$f_Y(y) = \sum_x f_{X,Y}(x,y)$

marginal Pdf's :

$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \cdot dy$.

$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \cdot dx$.

~~A function is~~ called as,

$\boxed{\text{Joint Pdf :}}$

A function $f_{X,Y}(x,y)$ is called joint Pdf of Continuose bivariant random variable (x,y) if

$F_{X,Y}(x,y) = \iint_{(x,y) \in A} f_{X,Y}(x,y) \cdot dx\, dy$

$f_X(x) = \text{Pdf}$

$F_X(x) = \text{cdf}$

$\frac{\partial F}{\partial x} = f(x)$

$F = \int f(x) \cdot dx$.

* Consider

joint

$P(6 \text{ and red}) = \frac{1}{26}$.

$P(6 \mid \text{red}) = \frac{2}{26} = \frac{1}{13}$.

conditional

Q.1] Consider an experiment tossing a coin 3 times and Record the sequence of heads and tails X be the no. of heads Y = winnings earned in a single play of a game with following rules, based on the outcomes of the probability experiement

1] Player will win \$1 if first 'H' occurs on first toss.

2] Player will win \$2 if ~~Second 'H'~~ Heads' ~~on o~~ occur on Second toss.

3] Player will win \$3 if 'Heads' occurs on third toss.

4] Player loses '\$1' if no Heads occur

Calculate joint pmf, marginal pmf and expectation of $E(xy)$

X = no. of heads

Y = winning earned.

$X = \{0, 1, 2, 3\}$

$Y = \{-1, 1, 2, 3\}$

$S = \{HHH, HTH, HHT, THT, TTH, HTT, THH, TTT\}$

~~Pmf~~ Joint Pmf as X, Y; $f_{XY}(x,y)$

Pmf Y \ X	0	1	2	3
-1	1/8	0	0	0
1	0	1/8	2/8	1/8
2	0	1/8	1/8	0
3	0	1/8	0	0

marginal Pmfs:

X	$f_X(x)$	Y	$f_Y(y)$
0	1/8	-1	1/8
1	3/8	1	4/8
2	3/8	2	2/8
3	1/8	3	1/8
	= 1		= 1

Expectation $E[XY]$