UNIT-2

PRECIDITIVE MODELLING & ANALYSIS

Regression analysis, correlation analysis, Rank coorelation, multiple corelation, least square, curve fitting, goodness of fit.

CORELATION :-

The variables are said to be corelated if there is a change in one variable corresponding for a change in another variable.

Types of corelation

Corelation is classified in several ways the important types are :-

* Positive & Negative
* Simple & partial & multiple.
* Linear & Non linear

Positive & Negative Corelation

Corelation is positive (direct) if the variables deviate in the same direction i.e if they increase or decrease together.

Eg:- Height & weight

Corelation is negative (inverse) if the variable deviate in the opposite direction i.e if one variable is increasing and other is decreasing & vice versa.

Eg:- production & price.

Simple, partial & multiple

The distinction b/w simple, partial & multiple is based on the number of variables involved.

Eg:- Rice production, rainfall & fertility are partially corelated.

Linear & Non-linear corelation (curvilinear)

If the amount of change in one variable tends to near a constant ratio to the amount of change in other variable. Then the corelation is said to linear otherwise non-linear.

Eg:- If we double the amount of rainfall
Production of rice & wheat

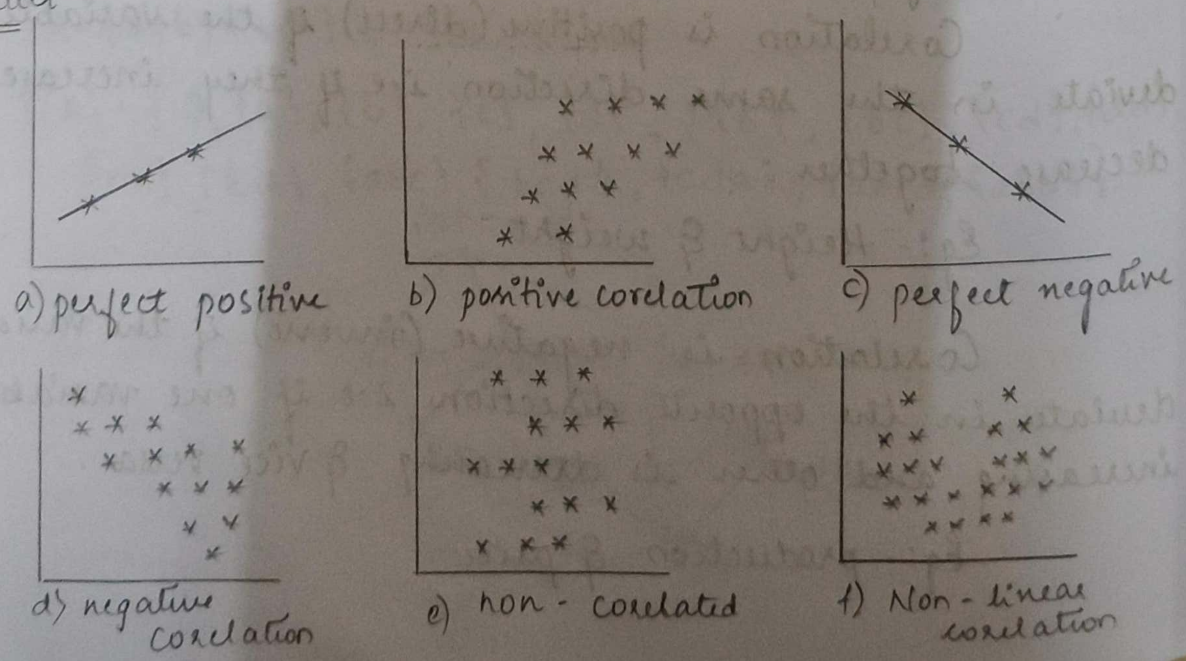
Note:- If there is no variable do not show any related variation they are said to non corelation.

Eg:- value of rupee & atmosphere are non corelated.

Measurements of corelation

1) Scatter diagram.
2) Karl Pearson's (Product Moment) coeff of corelation
3) spermans coefficient of rank correlation.

Scatter



Karl Pearson's coeff of correlation (product moment)

b/w 2 variable (x & y)

$$r = \frac{\text{Covariance}(x,y)}{\sqrt{\text{Variance}(x) \times \text{Variance}(y)}} = \frac{\text{Covariance}(x,y)}{\sigma_x \times \sigma_y}$$

i.e.) $$r = \frac{\frac{1}{n}\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\left[\frac{1}{n}\sum(x-x)^2\right] \times \left[\frac{1}{n}\sum(y-\bar{y})^2\right]}} = \frac{\frac{1}{n}\sum xy}{\sqrt{\frac{1}{n}\sum x^2 \times \frac{1}{n}\sum y^2}}$$

ii) for computational purpose the above formula is used in a simplified form.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\left[n\sum x^2 - (\sum x)^2\right] \times \left[n\sum y^2 - (\sum y)^2\right]}}$$ for raw data.

also $$r = \frac{N\sum fxy - (\sum fx)(\sum fy)}{\sqrt{\left[N\sum f x^2 - (\sum fx)^2\right] \times \left[N\sum fy^2 - (\sum fy)^2\right]}}$$ for tabulated data.

n = no of elements
N = sum of frequency.

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

Problems

1) Calculate the product moment coefficient of correlation b/w the following marks (out of 10). In computer science and mathematics of 5 students.

Student	1	2	3	4	5
Comp Sc	4	7	8	3	4
Maths	5	8	6	3	5

WKT =

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\left[n\sum x^2 - (\sum x)^2\right] \times \left[n\sum y^2 - (\sum y)^2\right]}}$$

Let x & y be the comp Sc & maths respectively.

$n=5$

x	y	x^2	y^2	xy
4	5	16	25	20
7	8	49	64	56
8	6	64	36	18
3	3	9	9	9
4	5	16	25	20
$\Sigma=$ 26	27	154	159	153

$$r = \frac{5\Sigma(153) - (26)(27)}{\sqrt{[5(154)-(26)^2] \times [5(159)-(27)^2]}}$$

$$= \frac{765-702}{\sqrt{[770-676] \times [795-729]}}$$

$$= \frac{63}{\sqrt{94 \times 66}}$$

$$= \frac{63}{\sqrt{6204}}$$

$$= \frac{63}{78.7} = \underline{\underline{0.8}}$$

There is a high positive corelation b/w marks in comp sc & maths of the students.

2) The height & weight of 5 students are given below

Height in cm(x)	160	161	162	163	164
weight in kgs(y)	50	53	54	56	57

Find the coeff of corelation.

WKT $r = \dfrac{\frac{1}{n}\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\left[\frac{1}{n}\sum(x-\bar{x})^2\right] \times \left[\frac{1}{n}\sum(y-\bar{y})^2\right]}}$

Let x & y be the height & weight respectively.

x	y	$(x-\bar{x})$	$(x-\bar{x})^2$	$(y-\bar{y})$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
160	50	-2	4	-4	16	8
161	53	-1	1	-1	1	1
162	54	0	0	0	0	0
163	56	1	1	2	4	2
164	57	2	4	3	9	6
Σ 810	270		10		30	17

$\bar{x} = \frac{\Sigma x}{n} = \frac{810}{5} = 162$

$\bar{y} = \frac{\Sigma y}{n} = \frac{270}{5} = 54$

$r = \dfrac{\frac{1}{5}[17]}{\sqrt{\frac{1}{5}(10) \times \frac{1}{5}(30)}}$

$= \dfrac{\frac{17}{5}}{\sqrt{2 \times 6}}$

$= \dfrac{\frac{17}{5}}{\sqrt{12}} = \dfrac{17}{5\sqrt{12}} = \dfrac{17}{17.32} = 0.981$

$= \underline{0.98}$

There is a high positive corelation b/w height & weight.

The following are the height of 8 persons & one of their wives. find the coeff of corelation b/w heights of husband & wives. 0.96.

Husband (cm)	164	176	178	184	175	167	173	180
Wife (cm	168	174	175	181	173	166	173	179

WKT =

$$r = \frac{\frac{1}{n}\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\left[\frac{1}{n}\sum(x-\bar{x})^2\right]\times\left[\frac{1}{n}\sum(y-\bar{y})^2\right]}}$$

Let x & y be the heights respectively.

x	y	$(x-\bar{x})$	$(x-\bar{x})^2$	$(y-\bar{y})$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
164	168	11	121	6	36	66
176	174	+1	1	0	0	0
178	175	+3	9	+1	1	3
184	181	9	81	7	49	~~44~~ 63
175	173	0	0	-1	1	0
167	166	-8	64	-8	64	~~-5~~ 64
173	173	-2	4	-1	1	+2
180	179	5	25	5	25	25
1397	1389		305		177	224

$$\bar{x} = \frac{\sum x}{n} = \frac{1397}{8} = 174.6 = 175$$

$$\bar{y} = \frac{\sum y}{n} = \frac{1389}{8} = 173.6 = 174$$

$$r = \frac{\frac{1}{8}\times 224}{\sqrt{\left(\frac{1}{8}\times 305\right)\left(\frac{1}{8}\times 177\right)}}$$

$$= \cancel{\frac{224}{8}} \qquad r = \frac{\frac{1}{8}(224)}{\frac{1}{8}\left(\sqrt{305\times 177}\right)}$$

$$= \frac{224}{\sqrt{53985}}$$

$$= \frac{224}{232.34}$$

$$= \underline{\underline{0.96}}$$

28/1/25

1) Calculate the coeff of corelation b/w x & y using following data.

x	218	220	236	225	220	227	228
y	12.3	12.7	12.0	12.2	12.7	12.1	12.0

WKT

$$r = \frac{\frac{1}{n}\sum xy}{\sqrt{\left[\frac{1}{n}\sum x^2\right] - \left[\frac{1}{n}\sum y^2\right]}}$$, $x = x-\bar{x}$ $y = y-\bar{y}$, $n=7$

x	y	$x = x-\bar{x}$	$y = y-\bar{y}$	x^2	y^2	xy
218	12.3	-7	0.3	49	0.09	-2.1
220	12.7	-5	0.7	25	0.49	-3.5
236	12.0	11	0	121	0	0
225	12.2	0	0.2	0	0.04	0
220	12.7	-5	0.7	25	0.49	-3.5
227	12.1	2	0.1	4	0.01	0.2
228	12.0	3	0	9	0	0
1574	86	-1	2	233	1.12	-8.9

$$r = \frac{\frac{1}{7}(-8.9)}{\sqrt{\frac{1}{7}(233) - \frac{1}{7}(1.12)}}$$

$$r = \frac{\frac{1}{7}(-8.9)}{\frac{1}{7}\left(\sqrt{233-1.12}\right)}$$

$$= \frac{-8.9}{\sqrt{231.88}}$$

$$= -\frac{8.9}{15.227} = \underline{-0.55}$$

Rank Correlation

$$\rho = 1 - \left[\frac{6\Sigma d^2}{n^3 - n}\right]$$

where $d = R_1 - R_2$

1) Find the rank corelation coeff for the following data

x	55	56	58	59
y	35	39	38	42

x	y	R_1 (x)	R_2 (y)	d	d^2
55	35	4	4	0	0
56	39	3	2	1	1
58	38	2	3	-1	1
59	42	1	1	0	0
					2

$$\rho = 1 - \left[\frac{6\Sigma d^2}{n^3 - n}\right]$$

$$= 1 - \left[\frac{6(2)}{4^3 - 4}\right]$$

$$= 1 - \frac{\cancel{12}^{1}}{\cancel{60}_{5}}$$

$$= 1 - \frac{1}{5}$$

$$= 4/5 = \underline{\underline{0.8}}$$

2) The following are the marks of 8 students in statistics & mathematics. find the coeff of rank corelation.

marks in stats	25	43	27	35	54	61	37	45
marks in maths	35	47	20	37	63	54	28	40

x	y	R_1	R_2	d	d^2
25	35	8	6	2	4
43	47	4	3	1	1
27	20	7	8	-1	1
35	37	6	5	1	1
54	63	2	1	-1	1
61	54	1	2	2	4
37	28	5	7	1	1

$n = 8$

$$\rho = 1 - \left[\frac{6\Sigma d^2}{n^3 - n}\right]$$

$$= 1 - \left[\frac{6(14)}{512 - 8}\right]$$

$$= 1 - \frac{84}{504}$$

$$= 0.83$$

Rank Correlation

$$\rho = 1 - \left[\frac{6 \Sigma d^2}{n^3 - n}\right]$$

where $d = R_1 - R_2$

1) Find the rank corelation coeff for the following data

x	55	56	58	59
y	35	39	38	42

x	y	R_1 (x)	R_2 (y)	d	d^2
55	35	4	4	0	0
56	39	3	2	1	1
58	38	2	3	-1	1
59	42	1	1	0	0
					2

$$\rho = 1 - \left[\frac{6 \Sigma d^2}{n^3 - n}\right]$$

$$= 1 - \left[\frac{6(2)}{4^3 - 4}\right]$$

$$= 1 - \frac{\cancel{12}\,1}{\cancel{60}\,5}$$

$$= 1 - \frac{1}{5}$$

$$= 4/5 = \underline{\underline{0.8}}$$

2) The following are the marks of 8 students in statistics & mathematics. find the coeff of rank corelation.

marks in stats	25	43	27	35	54	61	37	45
marks in maths	35	47	20	37	63	54	28	40

x	y	R_1	R_2	d	d^2
25	35	8	6	2	4
43	47	4	3	1	1
27	20	7	8	-1	1
35	37	6	5	1	1
54	63	2	1	1	1
61	54	1	2	-1	1
37	28	5	7	2	4
				1	1

$n = 8$

$$\rho = 1 - \left[\frac{6 \Sigma d^2}{n^3 - n}\right]$$

$$= 1 - \left[\frac{6(14)}{512 - 8}\right]$$

$$= 1 - \frac{84}{504}$$

$$= 0.83$$

3) Find the rank of corelation coeff for height (cms) & weight (kgs) of 6 solider in Indian army.

Height	165	167	166	170	169	172
Weight	61	60	63.5	63	61.5	64

x	y	R_1	R_2	d	d^2
165	61	6	5	1	1
167	60	4	6	-2	4
166	63.5	5	2	3	9
170	63	2	3	-1	1
169	61.5	3	4	-1	1
172	64	1	1	0	0
					16

$$\rho = 1 - \left[\frac{6\Sigma d^2}{n^3 - n}\right]$$

$$= 1 - \left[\frac{6(16)}{6^3 - 6}\right] = 1 - \frac{96}{216 - 6} = 1 - \frac{96}{210}$$

$$= \underline{\underline{0.542}}$$

4) Find the rank corelation coeff for the group of 6 persons in their examination marks & IQ's.

Exam marks	70	60	80	90	10	20
I.Q's	110	100	140	120	80	90

x	y	R_1	R_2	d	d^2
70	110	3	3	0	0
60	100	4	4	0	0
80	140	2	1	1	1
90	120	1	2	-1	1
10	80	6	6	0	0
20	90	5	5	0	0
					2

$$\rho = 1 - \left[\frac{6\Sigma d^2}{n^3 - n}\right]$$

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

$$= 1 - \left[\frac{12}{216 - 6}\right] = 1 - \left[\frac{12}{210}\right]$$

$$= 1 - 0.057$$

$$= 0.943$$

$$= \underline{0.97}$$

28/1/25

Repeated Ranks (Data with tie)

$$\rho = 1 - \left[\frac{6\Sigma d^2 + C.F}{n^3 - n}\right]$$

$$C.F = \left(\frac{m_1^3 - m_1}{12}\right) + \left(\frac{m_2^3 - m_2}{12}\right) + \left(\frac{m_3^3 - m_3}{12}\right) \ldots\ldots$$

1) The following are the marks in statistics (x) & marks in maths (y) of 10 students in I BCA examination. Find the coefficient of rank corelation.

x	43	96	74	38	35	43	22	56	35	80
y	30	94	84	38	30	18	30	41	48	95

x	y	R_1	R_2	d	d^2
43	30	5.5	8	-2.5	6.25
96	94	1	2	-1	1
74	84	3	3	0	0
38	38	7	6	1	1
35	30	8.5	8	0.5	0.25
43	18	5.5	10	-4.5	~~8625~~ 20.2
22	30	10	8	2	4
56	41	4	5	-1	1
35	48	8.5	4	4.5	20.2
80	95	2	1	1	1

$= 54.9 \Rightarrow \underline{55}$

43 → 5 & 6 rank repeating

$m_1 = \frac{5+6}{2} = \frac{11}{2} = 5.5 \quad m_1 \to 2$

35 → 8 & 9

$m_2 = \frac{8+9}{2} = \frac{17}{2} = 8.5 \quad m_2 \to 2$

$\frac{7+8+9}{3} = \frac{24}{3} = 8 \quad m_3 = 3$

$$C.F = \frac{(m_1^3 - m_1)}{12} + \left(\frac{m_2^3 - m_1}{12}\right) + \left(\frac{m_3^3 - m_1}{12}\right)$$

$$= \left(\frac{2^3-2}{12}\right) + \left(\frac{2^3-2}{12}\right) + \left(\frac{3^3-2}{12}\right)$$

$$= \frac{8-2}{12} + \frac{8-2}{12} + \frac{27-3}{12}$$

$$= \frac{6}{12} + \frac{6}{12} + \frac{24}{12}$$

$$= 0.5 + 0.5 + 2 = \underline{\underline{3}}$$

$$\rho = 1 - \left[\frac{6\Sigma d^2 + C.F}{n^3 - n}\right]$$

$$= 1 - \left[\frac{6(55) + 3}{10^3 - 10}\right]$$

$$= 1 - \left[\frac{330 + 3}{1000 - 10}\right]$$

$$= 1 - \left[\frac{333}{990}\right] \Rightarrow \underline{\underline{0.66}}$$

2) The following data of marks obtained by 10 students in physics & chemistry. Find the coffic of rank corelation

Physics (P)	35	56	50	65	44	38	44	50	15	26
Chemistry (C)	50	35	70	25	35	58	75	60	55	35

x	y	R_1	R_2	d	d^2
35	50	8	6	2	4
56	35	2	8	-6	36
50	70	3.5	2	1.5	2.25
65	25	1	10	-9	81
44	35	5.5	8	-2.5	6.25
38	58	7	4	3	9
44	75	5.5	1	4.5	20.2
50	60	3.5	3	0.5	1
15	55	10	5	5	25
26	35	9	8	1	1
					185

$\frac{3+4}{2} = \frac{7}{2} = 3.5 \Rightarrow 2$

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

$\frac{7+8+9}{3} = \frac{24}{3} = 8 \Rightarrow 3$

$C.F = \underline{\underline{3}}$

$$\rho = 1 - \left[\frac{6\,\Sigma d^2 + C.F}{n^3 - n}\right]$$

$$= 1 - \left[\frac{6(185)+3}{10^3-10}\right]$$

$$= 1 - \left[\frac{1110+3}{1000-10}\right] \Rightarrow 1 - \left[\frac{1113}{990}\right] \Rightarrow \underline{\underline{-0.124}}$$

3) Find the rank corelation coeff for following indices of supply & price of an article.

Supply	124	100	105	112	102	93	99	115	123	104	99	113
Price	80	100	102	91	100	111	109	100	89	104	111	102

x	y	R_1	R_2	d	d^2
124	80	1	15	-14	196
100	100	12	10	2	4
105	102	7	7.5	0.5	1
112	91	6	13	-7	49
102	100	10	10	0	0
93	111	15	3	12	144
99	109	13.5	5	8.5	72.2
115	100	4	10	-6	36
123	89	2	14	-12	144
104	104	8	6	2	4
99	111	13.5	3	10.5	110.2
113	102	5	7.5	-25	6.25
121	98	3	12	-9	81
103	111	9	3	6	36
101	123	11	1	10	100
					983

99

$\frac{12+13}{2} = \frac{25}{2} = 12.5 \to 2$

111

$\frac{2+3+4}{2} = \frac{9}{3} = 3 \to 3$

102

$\frac{7+8}{2} = 7.5 \to 2$

100

$\frac{9+10+11}{3} = 10 \to 3$

$$C.F = \left(\frac{2^2-2}{12}\right) + \left(\frac{3^2-3}{12}\right) + \left(\frac{2^2-2}{12}\right) + \left(\frac{3^2-3}{12}\right)$$

$= 0.5 + 0.5 + 2 + 2$

$= 5$

$$\rho = 1 - \left[\frac{6\Sigma d^2 + C.F}{n^3 - n}\right]$$

$$= 1 - \left[\frac{6(983)+5}{15^3-15}\right] = 1 - \left[\frac{5898+5}{3375-15}\right]$$

$$= 1 - \left[\frac{5903}{3360}\right]$$

$= 1 - 1.756$

$= -0.76$

121	103	101
98	111	123

20/1/25

Regression Analysis.

The statistical tool with the help of which it is possible to estimate or predict the unknown values of one variables from the known values of the other (when the variables are corelated) is called as regression.

Important properties of corelation & regression coeff.

1) Direct method

for larger digits i) $r = \dfrac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}}$ where $x = x-\bar{x}$, $y = y-\bar{y}$

raw data ii) $r = \dfrac{n\Sigma xy - \Sigma x\Sigma y}{\sqrt{[n\Sigma x^2-(\Sigma x)^2][n\Sigma y^2-(\Sigma y)^2]}}$

2) Grouped data.

iii) $r = \dfrac{n\Sigma uv - \Sigma u\Sigma v}{\sqrt{[n\Sigma u^2-(\Sigma u)^2][n\Sigma v^2-(\Sigma v)^2]}}$ where $u = x-\bar{x}$, $v = y-\bar{y}$

The lines of regression are $y-\bar{y} = \dfrac{r\sigma_y}{\sigma_x}(x-\bar{x})$ is known as the regression line of y on x.

AND

$x-\bar{x} = \dfrac{r\sigma_x}{\sigma_y}(y-\bar{y})$ is known as x on y.

($\because \sigma$ = standard deviation)

where regression coefficient are.

$$b_{yx} = \frac{r\sigma_y}{\sigma_x} = \frac{\frac{1}{n}\Sigma xy - \bar{x}\bar{y}}{\frac{\Sigma x^2}{n} - (\bar{x})^2}$$

&

$$b_{xy} = \frac{r\sigma_x}{\sigma_y} = \frac{\frac{1}{n}\Sigma xy - \bar{x}\bar{y}}{\frac{\Sigma y^2}{n} - (\bar{y})^2}$$

$$\therefore \sqrt{b_{yx}\, b_{xy}} = \sqrt{\left(r\frac{\sigma_y}{\sigma_x}\right)\left(r\frac{\sigma_x}{\sigma_y}\right)} = \sqrt{r^2} = r$$

WKT The corelation coeff is the geometric mean of regression coeff.

The angle θ b/w the 2 lines of regression is

$$\tan\theta = \frac{r^2-1}{r}\left(\frac{\sigma_x \sigma_y}{\sigma_x^2+\sigma_y^2}\right)$$

Problems

1) Find the corelation coeff & the equation of regression lines for the following data.

x	5	2	1	3	4
y	7	5	2	3	8

$\Rightarrow$ We have

$$r = \frac{n\Sigma uv - \Sigma u \Sigma v}{\sqrt{[n\Sigma u^2 - (\Sigma u)^2] * [n\Sigma v^2 - (\Sigma v)^2]}}$$

x	y	u ($x-\bar{x}$)	v ($y-\bar{y}$)	u^2	v^2	uv
5	7	2	2	4	4	4
2	5	-1	0	1	0	0
1	2	-2	-3	4	9	6
3	3	0	-2	0	4	0
4	8	1	3	1	9	3
15	25	0	0	10	26	13

$\bar{x} = \frac{15}{5} = 3$

$\bar{y} = \frac{25}{5} = 5$

$$r = \frac{n\Sigma uv - \Sigma u \Sigma v}{\sqrt{[n\Sigma u^2 - (\Sigma u)^2] * [n\Sigma v^2 - (\Sigma v)^2]}}$$

$$= \frac{5(13) - 0\times 0}{\sqrt{5(10) - (\cancel{10})^2 \times 5(26) - (\cancel{26})^2}} = \frac{65}{\sqrt{50 - \cancel{100} \times 130 - \cancel{676}}}$$

$$= \cancel{\frac{65}{\sqrt{50 \times -546}}} \qquad = \frac{65}{\sqrt{6500}}$$

$$= \cancel{\frac{65}{\sqrt{27300}}} \qquad = \frac{65}{80.6}$$

$$= \cancel{\frac{65}{165.2}} = \cancel{0.393} \qquad = \underline{\underline{0.80}}$$

The regression line y on x in

$$y - \bar{y} = \frac{r\sigma y}{\sigma x}(x - \bar{x})$$

$$b_{yx} = \frac{r\sigma y}{\sigma x} = \frac{1/n \, \Sigma uv - (\bar{u})(\bar{v})}{\cancel{\sqrt{\;}}\, 1/n \, \Sigma u^2 - (\bar{u})^2}$$

$$= \frac{1/5 \,(13) - (0)(0)}{1/5 \,(10) - 0}$$

$$= \frac{2.6}{2} = \underline{\underline{1.3}}$$

i.e) $y - 5 = 1.3(x - 3)$

$y - 5 = 1.3x - 3(1.3)$

$y = 1.3 - 3(1.3) + 5$

$y = \underline{\underline{1.3x + 1.1}}$ is the regression line of y on x.

The regression line x on y in

$$x - \bar{x} = \frac{r\sigma x}{\sigma y}(y - \bar{y})$$

$$b_{xy} = \frac{r\sigma x}{\sigma y} = \frac{1/n \, \Sigma uv - (\bar{u})(\bar{v})}{1/n \, \Sigma u^2 - (\bar{u})^2}$$

$$= \frac{1/5 \,(13) - (0)(0)}{1/5 \,(26) - 0} = \underline{\underline{0.5}}$$

$\therefore (x-\bar{x}) = \frac{r\sigma x}{\sigma y}(y-\bar{y})$

$x-3 = 0.5(y-5)$

$x-3 = 0.5y - 5(0.5)$

$x = 0.5y - 2.5 + 3$

$x = 0.5y + 0.5$ is the regression line of x on y.

2) Calculate the correlation coeff of the following data of 10 students marks in subject maths & economics

Student	A	B	C	D	E	F	G	H	I	J
maths	62	90	39	78	98	65	82	25	75	36
economics	58	86	47	84	91	53	62	60	68	51

x	y	u	v	u^2	v^2	uv
62	58	-3	-8	9	64	24
90	86	25	20	625	400	500
39	47	-26	-19	676	361	494
78	84	+13	18	169	324	234
98	91	33	25	1089	625	825
65	53	0	-13	0	169	0
82	62	17	-4	289	16	-68
25	60	-40	-6	1600	36	240
75	68	10	2	100	4	20
36	51	-29	-15	841	225	435
650	660	00	0	5398	2224	2704

$\bar{x} = \frac{\Sigma x}{n} = \frac{650}{10} = 65$ $\quad \bar{y} = \frac{\Sigma y}{n} = \frac{660}{10} = 66$

$= \frac{10(2704) - (0)(0)}{\sqrt{10(5398) - 0^2 - 10(2224) - 0^2}}$

$= \frac{27040}{\sqrt{53980 \times 22240}}$

$= \frac{27040}{\sqrt{1200515}} = \frac{27040}{34648.4} = 0.78$

3) Given the lines of regression as $8x - 10y + 66 = 0$, $40x - 18y = 214$. Variance of x is 9. Determine the mean of x series & y series & the coefficient of correlation b/w x & y.

Gvn $8\bar{x} - 10\bar{y} = -66 \rightarrow (1) \times 40$

$40\bar{x} - 18\bar{y} = 214 \rightarrow (2) \times 8$

$320\bar{x} - 400\bar{y} = -2640$

$320\bar{x} - 144\bar{y} = 1712$

(-) (+) (-)

$-256\bar{y} = -4352$

$\bar{y} = \frac{\cancel{-4352}^{17}}{\cancel{-256}}$

$\bar{y} = 17$

Substitute $\bar{y} = 17$ in (1)

$8\bar{x} - 10(17) = -66$

$8x - 170 = -66$

$8x = -66 + 170$

$8x = 104$

$x = \frac{\cancel{104}^{13}}{8} \Rightarrow x = 13$

x & $y \Rightarrow 13$ & 17

$\Rightarrow 10y = 8x + 66$

y on x $\quad y = \frac{8}{10}x + \frac{66}{10}$

$\Rightarrow y = 0.8x + 66$

$b_{yx} = 0.8$

$\Rightarrow 40x = 214 + 18y$

x on y $\quad \Rightarrow x = \frac{18}{40}y + \frac{214}{40}$

$x = 0.45y + 5.35$

$b_{xy} = 0.45$

$r = \sqrt{b_{yx} \times b_{xy}} = \sqrt{0.8 \times 0.45}$

$= \sqrt{0.36}$

$r = 0.6$

Gvn $\sigma x^2 = 9$

$\sigma x = \sqrt{9}$

$\sigma x = \underline{3}$

were $b_{yx} = r\frac{\sigma y}{\sigma x}$

$0.8 = (0.6)\frac{\sigma y}{3}$

$2.4 = (0.6)\sigma y$

$\sigma y = \frac{2.4}{0.6} = 4$

$\sigma y = 4$

4) Find the corelation coeff of series x & y whose regression lines are $3x + 2y = 26$, $6x + y = 31$. Hence determine the mean of x & y series.

$3\bar{x} + 2\bar{y} = 26 \times 6$

$6\bar{x} + \bar{y} = 31 \times 3$

$18\bar{x} + 12\bar{y} = 156$

$18\bar{x} + 3\bar{y} = 93$

(-) (-) (-)

$9\bar{y} = 63$

$\bar{y} = \frac{63}{9}$

$\bar{y} = 7$

Substitute in ①

$3\bar{x} + 2(7) = 26$

$3\bar{x} = 14 = 26$

$3\bar{x} = 26 - 14$

$3\bar{x} = 12$

$\bar{x} = \frac{12}{3}$

$\bar{x} = 4$

$\Rightarrow$ y on x $\Rightarrow 2y = 26 - 3x$

$\Rightarrow y = \frac{-3x}{2} + \frac{26}{2}$

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

$\rightarrow$ x on y $\Rightarrow 6x = 31 - y$

$x = \frac{-1}{6}y + \frac{31}{6}$

$b_{xy} = \frac{-1}{6}$

$r = \sqrt{b_{yx} \times b_{xy}}$

$= \sqrt{\frac{-3}{2} * \frac{-1}{6}}$

$= \sqrt{-1.5 * -0.1}$

$= \pm 0.5$ $\Rightarrow r = -0.5$ since the reg coeff is negative.

5) For the regression lines $4x-5y+33=0$ & $20x-9y=107$.
Find i) mean values of x & y
ii) The corelation b/w x & y
iii) The variance of y given the variance of x is 9.

$4x-5y+33=0 \rightarrow ①$
$20x-9y=107 \rightarrow ②$

$4x-5y=-33 \quad \times 20$
$20x-9y=107 \quad \times 4$

$\cancel{80x}-100y=-660$
$\cancel{80x}-36y=428$
$(-) \quad (+) \quad (-)$

$-64y=-1088$

$y=\frac{\cancel{-1088}}{\cancel{-64}}$

$y=17$

$y=17$ substitute in ①

$4x-5(17)=-33$
$4x-85=-33$
$4x=-33+85$
$4x=52$
$x=\frac{52}{4}$
$x=13$

$\Rightarrow y$ on $x \Rightarrow +5y=4x+33$
$y=\frac{4}{5}x+\frac{33}{5}$
$y=0.8x+6.6$
$b_{yx}=0.8x$

$\Rightarrow x$ on $y \Rightarrow 20x=107+9y$
$x=\frac{9}{20}y+\frac{107}{20} \Rightarrow 0.45$ $+0.35$
$b_{xy}=0.45$

$r = \sqrt{b_{yx} \times b_{xy}} = \sqrt{0.8 \times 0.45}$

$= \sqrt{0.36}$

$r = 0.6$

$\sigma_x^2 = 9$

$\sigma_x = \sqrt{9}$

$\sigma_x = 3$

we've $b_{yx} = \frac{r \sigma_y}{\sigma_x}$

$0.8 = \frac{0.6 (\sigma_y)}{3}$

$2.4 = (0.6) \sigma_y$

$\sigma_y = \frac{2.4}{0.6}$

$\sigma_y = 4$

6) Given $\bar{x} = 5$, $\bar{y} = 11$, $n = 9$ $\Sigma(x_i - \bar{x})^2 = 60$, $\Sigma(y_i - \bar{y})^2 = 70$
$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 62$

In the usual notation, find

i) coeff of corelation

ii) Regression lines

iii) Also estimate the value of y when $x = 10$ & value of x when $y = 10$.

We have

i) $b_{yx} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{62}{60}$

$= 1.03$

$b_{xy} = \frac{\Sigma(x_i - \bar{x})(y_i - y)}{\Sigma(y_i - \bar{y})^2} = \frac{62}{70} = 0.88$

$r = \sqrt{b_{yx} \times b_{xy}} = \sqrt{1.03 \times 0.88} = 0.95$

ii) The reg lines y on x

$(y - \bar{y}) = b_{yx}(x - \bar{x})$

$y - 11 = (1.03)(x - 5)$

$\Rightarrow y = 1.03x - 5(1.03) + 11$

$y = 1.03x + 5.85$

iii) put $x=10$

$$y = 1.03x + 5.85$$

$$y = 1.03(10) + 5.85$$

$$y = \underline{16.15}$$

ii) The reg line on x & y

$$x - \bar{x} = b_{xy}(y - \bar{y})$$

$$\Rightarrow x - 5 = (0.88)(y - 11)$$

$$x = 0.88y - 11(0.88) + 5$$

$$x = \underline{0.88y - 4.68}$$

put $y = 10$

$$x = 0.88(10) - 4.68$$

$$= 8.8 - 4.68$$

$$x = \underline{4.12}$$

7) The height and weight of 6 students are as follows obtained the 2 regression lines the expected height of a person who's height is 60 kg.

x	153	157	168	160	170	163
y	48	50	50	49	54	53

we've $b_{yx} = \dfrac{1/n\, \Sigma uv - (\bar{u})(\bar{v})}{1/n\, \Sigma u^2 - (\Sigma u)^2} = \dfrac{n\Sigma uv - \Sigma(u)\Sigma(v)}{n\Sigma u^2 - (\Sigma u)^2}$

$b_{xy} = \dfrac{1/n\, \Sigma uv - (\bar{u})(\bar{v})}{1/n(\Sigma v^2) - (\Sigma v)^2} = \dfrac{n\Sigma uv - (\Sigma u)(\Sigma v)}{n\Sigma v^2 - (\Sigma v)^2}$

x	y	$u=x-\bar{x}$	$v=y-\bar{y}$	u^2	v^2	uv
153	48	-9	-3	81	9	27
157	50	-5	-1	25	1	5
168	50	6	-1	36	1	-6
160	49	2	-2	4	4	4
170	54	8	3	64	9	24
163	53	+1	2	1	4	2
971	304			211	28	56

$\bar{x} = \frac{\Sigma x}{n} = \frac{971}{6} = 161.83 \approx 162$

$\bar{y} = \frac{\Sigma y}{n} = \frac{304}{6} = 50.6 \approx 51$

$b_{yx} = \frac{n\Sigma uv - (\Sigma u)(\Sigma v)}{n\Sigma u^2 - (\Sigma u)^2} = \frac{6(56) - (-1)(-2)}{6(211) - (-1)^2}$

$= \frac{336-2}{1266-1} = \frac{334}{1265} = 0.264$

$b_{xy} = \frac{n\Sigma uv - (\Sigma u)(\Sigma v)}{n\Sigma v^2 - (\Sigma v)^2} = \frac{6(56) - (-1)(-2)}{6(28) - (-2)^2}$

$= \frac{334}{164} = 2.0$

The regression line y on x.

$(y - \bar{y}) = b_{yx}(x - \bar{x})$

$y = 50.6 = (0.26)(x - 161.83)$

$y = (0.26)x - (0.26)(-161.83) + 50.6$

$y = 0.26x + 0.26(161.83) + 50.6$

$y = 0.26x + 92.67$

The regression line x on y.

$(x-\bar{x}) = b_{xy}(y-\bar{y})$

$x - 161.83 = 2.03(y - 50.6)$

$x = 2.03y - (2.03)(50.6) + 161.83$

$x = 2.03y + 59.12$

here height (x), weight (y)

Find x when $y = 60$

Consider the regression line x on y

$x = 2.03y + 59.12$.

when $y = 60$

$x = 2.03(60) + 59.12$

$x = 180.9$ cms

10/2/25

Fitting a straight line

$y = a + bx$ / $Y = A + BX$

The normal equation are :-

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

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

(or)

$y = ax + b$

Normal eqn

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

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

Power curve

$y = ax^b$

$\log y = \log(a x^b)$

$= \log a = \log x^b$

$= \log a = b \log x.$

put $\log y = y$, $\log x = x$, $\log a = A$ $\Rightarrow y = A + bx$

Normal eqn

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

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

parabola

$y = a + bx + cx^2$

Normal eqn

$\Sigma y = na + b\Sigma x + c\Sigma x^2$

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

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

(or)

$y = ax^2 + bx + c$

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

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

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

Problems

1) Fit a straight line to the following data.

x	6	7	7	8	8	8	9	9	10
y	5	5	4	5	4	3	4	3	3

Soln

The curve of best fit are $y = a + bx$.

The normal equation are

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

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

x	y	x^2	xy
6	5	36	30
7	5	49	35
7	4	49	28
8	5	64	40
8	4	64	32
8	3	64	24
9	4	81	36
9	3	81	27
10	3	100	30
72	36	588	282

$n = 9$

The normal eqn becomes.

$36 = 9a + 72b$

$282 = 72a + 588b$

$\Rightarrow 9a + 72b = 36 \quad \times 8$

$72a + 588b = 282 \quad \times 1$

$\cancel{72}a + 576b = 288$

$\cancel{72}a + 588b = 282$

$(-) \quad (-) \quad (-)$

$-12b = 6$

$b = \frac{\cancel{6}}{-\cancel{12}\,2} = \frac{1}{-2}$

$= \underline{-0.5}$

substitute in $9a + 72b = 36$

$= 9a + 72(-0.5) = 36$

$= 9a + (-36) = 36$

$= 9$

$72a + 588b = 282$

$72a + 588(0.5) = 282$

$72a + (-294) = 282$

$72a - 294 = 282$

$72a = 282 + 294$

$72 = 576$

$a = \frac{576}{72}$

$\boxed{a = 8}$

$\boxed{y = 8 - 0.5x}$

Q. Fit a straight line to the following data

x	1911	1922	1931	1941	1951
y	15	23	28	32	39

Let $X = \frac{x - 1931}{10}$

The eqn of line becomes $y = a + bX$

The normal eqn are

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

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

x	y	$X = \frac{x-1931}{10}$	x^2	xy.
1911	15	-2	4	-30
1921	23	-1	1	-23
1931	28	0	0	0
1941	32	1	1	32
1951	39	2	4	78
	137	0	10	57.

$\Sigma y = na + \Sigma bx \Rightarrow 137 = 5a + b(0)$

$= a = \frac{137}{5}$

$= 27.4$

$$\Sigma xy = a\Sigma x + b\Sigma x^2$$
$$57 = a(0) + 10b$$
$$b = \frac{57}{10} = \underline{\underline{5.7}}$$
$$\boxed{y = 27.4 + 5.7x}$$
$$y = 27.4 + 5.7\left[\frac{x - 1931}{10}\right]$$
$$= 27.4 + \frac{5.7}{10}(x - 1931)$$
$$y = 27.4 + 0.57x - 1931 \times 0.57$$
$$y = 27.4 + 0.57x - 1100.67$$
$$y = \underline{\underline{0.57x - 1073.27}}$$

3) Fit a curve of the form $y = ax^b$ to the following data.

x	3	4	5	6	7
y	126	42	20	6	3

we've $y = ax^b$
$$\log y = \log a + b \log x$$
Let $\log x = X$, $\log y = Y$, $\log a = A$
$$\Rightarrow Y = A + bX$$
The normal eqn
$$\Sigma Y = nA + b\Sigma X$$
$$\Sigma XY = A\Sigma X + b\Sigma X^2$$

x	y	X	Y	X^2	XY
3	126	0.4771	2.1004	0.2276	1.0021
4	42	0.6021	1.6232	0.3625	0.9773
5	20	0.6989	1.3010	0.4884	0.9092
6	6	0.7781	0.7782	0.6054	0.6055
7	3	0.8451	0.4771	0.7141	0.4031
		3.4013	6.2799	2.3982	3.8972

$\Sigma Y = nA + b\Sigma X$

$6.2799 = 5A + 3.4013b$

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

$3.8972 = 3.4013A + 2.3982b$ solving eqn

$A = 4.7254$

Antilog = shift + log(

$B = -4.4386$

$Y = A + BX$

Antilog (A) = a = Antilog (42754)

$= 18853.84$

$y = (18853.84)x^{(-4.4386)}$

A) Find the exponential eqn fit of $y = ae^{bx}$.

x	1	2	3	4	5	6
y	1.6	4.5	13.8	40.2	125	300

We've $y = ae^{bx}$

$\log y = \log a + \log e^{bx}$ $[\because \log e = 1]$

$= \log a + bx(\log e)$

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

Let $Y = \log y$ & $A = \log a$

$\Rightarrow Y = A + bx$

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

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

x	y	Y	X^2	XY	X
1	1.6	0.2041	0	0	0
2	4.5	0.6532	0.09060	0.19661	0.3010
3	13.8	1.1399	0.22764	0.54386	0.47712
4	40.2	1.6042	0.3624	0.9657	0.6020
5	125	2.096	0.4884	1.4648	0.6989
6	300	2.4771	0.621180	1.95232	0.78815
		8.1745	1.7902	5.1232	3.6553

$\Sigma Y = nA + b\Sigma X$

$8.1745 = 6A + 3.6553b$

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

$5.1232 = 3.6553A + 1.7902b$

$X = 1.56214$

$Y = -0.32784$

$Y = A + BX$

$\text{Antilog}(A) = a = \text{Antilog}(1.56214)$

$= 36.4871$

$y = (36.4871)x^{(-0.32784)}$

Fit a parabola for the following data.

x	1	2	3	4	5	6	7	8	9
y	2	6	7	8	10	11	11	10	9

Let $X = x - 5$ $\quad Y = y - 10$

$$y = ax^2 + bx + c$$

$$\Rightarrow Y = aX^2 + bX + C$$

The normal eqn are.

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

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

$$\Sigma X^2 Y = a\Sigma X^4 + b\Sigma X^3 + C\Sigma X^2$$

x	y	$X = x-5$	$Y = y-10$	X^2	X^3	X^4	XY	X^2Y
1	2	-4	-8	16	-64	256	32	-128
2	6	-3	-4	9	-27	81	12	-36
3	7	-2	-3	4	-8	16	6	-12
4	8	-1	-2	1	-1	1	2	-2
5	10	0	0	0	0	0	0	0
6	11	1	1	1	1	1	1	1
7	11	2	1	4	8	16	2	4
8	10	3	0	9	27	81	0	0
9	9	4.	-1	16	64	256	-4	-16
		$\Sigma X = 0$	$\Sigma Y = -16$	$\Sigma X^2 = 60$	0	708	51	-189

① $-16 = a(60) + b(0) + 9(c)$

$$60a + 9c = -16$$

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

$51 = a(0) + b(60+) 0(c)$

$60b = 51$

③ $+788a + 0b + 60c = -189$

$a = -0.27$

$b = 0.85$

$c = 0.004$

$\Rightarrow y = aX^2 + bX + c$

$y - 10 = -0.27(x-5)^2 + 0.85(x-5) + 0.004$

$y = -0.27(x-5)^2 + 0.85x - 5(0.85) + 0.004 + 10$

$y = -0.27(x-5)^2 + 0.85x + 5.754$

6) Fit a parbola for the following data.

x	1911	1921	1931	1941	1951
y	15	23	28	32	39

Let $X = \left[\frac{x - 1931}{10}\right]$

$y = ax^2 + bx + c$

The normal eqn are.

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

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

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

x	y	$X = \frac{x-1931}{10}$	X^2	X^3	X^4	XY	X^2Y
1911	15	-2	4	-08	16		
1921	23	-1	1	-1	1		
1931	28	0	0	0	0		
1941	32	1	1	1	1		
1951	39	2	4	8	16		

17/2/25

UNIT - 3

Probability & Joint probability

Random variable:-

Let S be a sample space associated with a random experiment, a function $X: S \to R$ which assigns to each element $x \in S$ the one and only real number is called as random variable.

Eg:- In a experiment of tossing 2 fair coins we have sample space

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

suppose each of 4 sample points are arranged then finds the no of heads on X Then.

S	TT	TH	HT	HH
X - No of head	0	1	1	2

P(getting 0 head) = $\frac{1}{4}$

P(getting 1 head) = $\frac{2}{4} = \frac{1}{2}$

P(getting 2 head) = $\frac{1}{4}$

$$\Sigma P(x) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$$

Discrete random variable:-

A variable X which takes a countable no of values $x_1, x_2 \cdots x_n$ with probabilities $P_1, P_2 \cdots P_n$, is a discrete random variables then the distribution of the function $F: R \to R$ is the discrete random variable can be defined as $F(x) = P(X \leq x) = \sum_{x_i \leq x} p_i$ is also called as probability mass function (pmf) of the discrete random variable X.

Here (i) $P(x_i) \geq 0$ for all x &

(ii) $\sum p(x_i) = 1$

Eg:- A random experiment consists of 3 tosses of a fair coin. Let X is a discrete random variable which assigns to get the heads. find the probability mass function and distribution.

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

$P(\text{getting 0 head}) = \frac{1}{8}$

$P(\text{getting 1 head} = \frac{3}{8}$

$P(\text{getting 2 head}) = \frac{3}{8}$

$P(\text{3 heads}) = \frac{1}{8}$

$$\sum P = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = \underline{\underline{1}}$$

$$F(x) = \begin{cases} 0, \frac{1}{8} & , \text{if } x < 0 \\ \frac{1}{8} + \frac{3}{8} = \frac{4}{8} = \frac{1}{2} & , \text{if } 0 \leq x < 1 \\ \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8} & , \text{if } 1 \leq x < 2 \\ \frac{7}{8} + \frac{1}{8} = \frac{8}{8} = 1 & , \text{if } 2 \leq x < 3 \\ & , \text{if } x \geq 3 \end{cases}$$

<u>Syllabus</u>

Random variable & probability <u>dist</u>

Introduction & its property, random variable, its types, DRV, CRV and its distribution, 2-D random variable, joint probability func, marginal density func, some special probability <u>dis</u> – binomial, poisson, uniform, exponential & normal dist

Continuous Random variable

A random variable X is said to CRV if it can take any value in an interval which may be finite or infinite.

Let X be a CRV taking the values in the interval $(-\infty, \infty)$. Let $f(x)$ be a function such that

(i) $f(x)$ is integrable on $(-\infty, \infty)$

(ii) $f(x) \geq 0 \;\forall\; x \in (-\infty, \infty)$

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

$$\text{Mean} = \bar{x} = \int_{-\infty}^{\infty} x\, f(x)\,dx$$

$$\text{Variance} = \sigma^2 = \int_{-\infty}^{\infty} (x-\bar{x})^2\, f(x)\,dx$$

Problems

1) Let X denote the RV which has the following probability func

Value of X : x	0	1	2	3	4	5	6	7
p(x)	0	k	2k	2k	3k	k^2	$2k^2$	$7k^2+k$

(i) find k

(ii) evaluate $p(x<6)$, $p(x>6)$ & $p(0<x<5)$

(iii) if $p(x<k)$ find the max value of k

(iv) determine the dist func of X.

Soln:- Gvn that X has the probability func such that

$\sum_{i=0}^{7} P(x) = 1$

$0 + k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k^2 = 1$

$10k^2 + 9k = 1$

$\Rightarrow 10k^2 + 9k - 1 = 0$

$\cancel{10k^2 + 10k - 1k - 1 = 0}$

$\cancel{10}$

-10 → +10, -1

$\Rightarrow k = 0.1$

-10 → +10, -1

$\frac{+10}{10}$ $\frac{-1}{10}$

$10k^2 + 10k - 1k - 1 = 0$

$10k(k+1) - 1(k+1) = 0$

$(10k - 1)(k+1) = 0$

$10k - 1 = 0$ $\quad k+1 = 0$

$10k = 1$ $\quad k = -1$

$k = \frac{1}{10}$

$k = 0.1$

$\left(k + \frac{10}{10}\right)\left(k - \frac{1}{10}\right) = 0$

$k + \frac{10}{10} = 0$ $\quad k - \frac{1}{10} = 0$

$k = \frac{-10}{10} = -1$ $\quad k = \frac{1}{10} = 0.1$

$k = \frac{1}{10}$

i) $P(x)$ $\quad 0 \quad 1/10 \quad 2/10 \quad 3/10 \quad 1/100 \quad 2/100 \quad 7/100 + 1/10$

ii) $P(x<6) = P(x=0) + P(x=1) + P(x=2) + (P(x=3) + P(x=4) + P(x=5)$

$= 0 + 1/10 + 2/10 + 2/10 + 3/10 + 1/100$

$= 8/10 + 1/100 = 81/100 = 0.81$

$P(x>6) = 1 - P(x<6)$

$= 1 - \frac{81}{100} = \frac{19}{100} = 0.19$

$$p(0<x<5) = p(x=1)+p(x=2)+p(x=3)+p(x=4)$$
$$= \frac{1}{10}+\frac{2}{10}+\frac{2}{10}+\frac{3}{10}$$
$$= \frac{8}{10} = \underline{\underline{0.8}}$$

iii) $p(x<k) = \underline{\underline{\frac{1}{2}}}$

iv) Dist func $F(x)$

x	$P(x)$	$F(x)$	put $k=0.1$
0	0	0	
1	k	$0+k=k$	0.1
2	$2k$	$k+2k=3k$	0.3
3	$2k$	$3k+2k=5k$	0.5
4	$3k$	$5k+3k=8k$	0.8
5	k^2	$8k+k^2=$	0.81
6	$2k^2$	$8k+k^2+2k^2$	0.83
7	$7k^2+k$	$8k+3k^2+7k^2+k$	$\underline{\underline{1}}$

2) Let X be DRV taking the values from $\{1, 2, 3, 4 \dots\}$ let $f(x) = \frac{1}{2^x}$ be the probability density function then find $P(A)$ where $A = \{1, 3, 5, 7 \dots\}$

$f(x) = \frac{1}{2^x}$.

$\Rightarrow$ We've $P(A) = P(x=1)+P(x=3)+P(x=5)+\dots$
$$= \frac{1}{2^1}+\frac{1}{2^3}+\frac{1}{2^5}+\frac{1}{2^7}+\dots$$
$$= \left(\frac{1}{2}\right)^1+\left(\frac{1}{2}\right)^3+\left(\frac{1}{2}\right)^5+\left(\frac{1}{2}\right)^7+\dots$$
$$= n+n^3+n^5+n^7+\dots$$
$\left[\because x+x^3+x^5+x^7\dots = \frac{x}{1-x^2}\right]$

Here $n = \frac{1}{2}$

$\because P(n) = \frac{n}{1-n^2}$

$$= \frac{\frac{1}{2}}{1-\left(\frac{1}{2}\right)^2} = \frac{\frac{1}{2}}{1-\frac{1}{4}} = \frac{\frac{1}{2}}{\frac{4-1}{4}} = \frac{\frac{1}{2}}{\frac{3}{4}}$$

$$= \frac{1}{\cancel{2}} \times \frac{\cancel{4}^2}{3} = \frac{2}{3}$$

3) Find the constant c so that $f(x) = c\left(\frac{2}{3}\right)^x$; $x = 1, 2, 3 \ldots$ satisfies the pdf of a discrete random variable x

Soln If f is a p.d.f then $\sum_x (f(x,)) = 1$

Here we've $f(x) = c\left(\frac{2}{3}\right)^x$.

$$\therefore f(x) = c\left\{\left(\frac{2}{3}\right)^1 + \left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^3 + \ldots\right\} = 1$$

$$= c\left\{\frac{2/3}{1-2/3}\right\} = 1 \qquad \left\{\because x + x^2 + x^3 + \ldots = \frac{x}{1-x}\right\}$$

$$= c\left\{\frac{2/3}{1/3}\right\} = 1$$

$$= 2c = 1$$

$$= \boxed{c = 1/2}$$

4) Find the constant c so that $f(x = \begin{cases} cxe^{-x} & \text{if } 0 < x < \infty \\ 0 & \text{if otherwise} \end{cases}$ satisfies the condn of pdf of the continuous RV.

Soln Since the $f(x)$ satisfies p.d.f

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

$$\Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\infty} f(x)\,dx = 1$$

$$\int_0^\infty f(x)\,dx = 1$$

$$\int_0^\infty cxe^{-x}\,dx = 1 \rightarrow ①$$

$$① \rightarrow c\int_0^\infty xe^{-x}\,dx$$

$$\left[\int x^n dx = \frac{x^{n+1}}{n+1} \right.$$

ILATE
I - integration
L - log
A - arithmetic
T - trigon
E - exponential

Integrate by parts

$$\boxed{\int u\,dv = uv - \int v\,du}$$

Differenciate

$u = x$

$du = 1 \cdot dx$

$du = dx$

Integrate

$dv = e^{-x}\,dx$

$\int dv = \int e^{-x}\,dx$

$v = \frac{e^{-x}}{-1}$

$v = -e^{-x}$

$$\int e^{ax}dx = \frac{e^{ax}}{a}$$

$$\int e^{-ax}dx = \frac{e^{-ax}}{-a}$$

$$c\int xe^{-x}dx = x(-e^{-x}) - \int(-e^{-x})\,dx$$

$$= -xe^{-x} + \int e^{-x}\,dx$$

$$= -xe^{-x} + \left(\frac{e^{-x}}{-1}\right)$$

$e^0 = 1$

$e^\infty = \infty$

$e^{-\infty} = 0$

$$c\int_0^\infty xe^{-x}\,dx = \left[-xe^{-x} - e^{-x}\right]_0^\infty$$

upper limit - lower limit

$$= \left[-\infty e^{-\infty} - e^{-\infty}\right] - \left[0e^0 - e^0\right]$$

$$= (0-0) - (-0(1) - (1))$$

$$= -(-1)$$

$$c\int_0^\infty xe^{-x} = 1$$

Substitute ② in ① $\Rightarrow c\left[\int_0^\infty xe^{-x}\,dx\right] = 1$

$$= c(1) = 1$$

$$c = 1$$

Q) Find the constant k such that $f(x) = \begin{cases} kx^2 & \text{if } 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$

is a probability func also find

i) $P(1 < x < 2)$

ii) Find the distribution func.

⇒ we've $f(x)$ in a p.d.f

i.e) $\int_{-\infty}^{\infty} f(x)\, dx = 1$

$$\int_{-\infty}^{0} f(x)dx + \int_{0}^{3} f(x)dx + \int_{3}^{\infty} f(x)dx = 1$$

$$\int_{0}^{3} f(x)\, dx = 1$$

$$\int_{0}^{3} kx^2\, dx = 1$$

$$k\int_{0}^{3} x^2\, dx = 1 \quad \left[\int x^n dx = \frac{x^{n+1}}{n+1}\right]$$

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

(U-L)

$$k\left[\frac{3^3}{3} - \frac{0^3}{3}\right] = 1$$

$$k\left(\frac{\cancel{27}^{\,9}}{\cancel{3}}\right) = 1$$

$$9k = 1$$

$$\boxed{k = \frac{1}{9}}$$

$$f(x) = kx^2 = \frac{1}{9}x^2 \text{ if } 0 < x < 3$$

i) $\int_{1}^{2} f(x)dx = \int_{1}^{2} \frac{1}{9}x^2\, dx$

$$= \frac{1}{9}\int_{1}^{2} x^2\, dx$$

$$= \frac{1}{9}\left(\frac{x^3}{3}\right)_1^2$$

$$= \frac{1}{9}\left(\frac{2^3}{3} - \frac{1^3}{3}\right)$$

$$= \frac{1}{9}\left(\frac{8-1}{3}\right)$$

$$= \frac{1}{9}\left(\frac{7}{3}\right)$$

$$\boxed{P(1 < x < 2) = \frac{7}{27}}$$

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

$$= \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{3} f(x)\,dx + \int_{0}^{\infty} f(x)\,dx$$

$$= \int_{0}^{3} kx^2\,dx$$

$$= \int_{0}^{3} \frac{1}{9} x^2\,dx$$

$$= \frac{1}{9}\left[\frac{x^3}{3}\right]_0^3$$

$$= \frac{1}{9}\left[\frac{3^3}{3} - \frac{0^3}{3}\right] \Rightarrow \frac{1}{9}\left(\frac{\cancel{27}\,9}{\cancel{3}}\right)$$

$$= \frac{9}{9} = 1$$

$$\therefore F(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ \frac{1}{9}x^2 & \text{if } 0 < x < 3 \\ 0 & \text{if } x > 3 \end{cases}$$

6) Verify the function $f(x) = \begin{cases} 6x(1-x), & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$ is a p.d.f or not. Hence determine mean & standard deviation.

$\rightarrow \int_{-\infty}^{\infty} f(x)dx = 1$ is a p.d.f

To prove $\int_{-\infty}^{0} \overset{\Rightarrow 0}{f(x)dx} + \int_{0}^{1} f(x)dx + \int_{1}^{\infty} \overset{\Rightarrow 0}{f(x)dx} = 1$

$$\int_{0}^{1} f(x)dx = \int_{0}^{1} 6x(1-x)\,dx$$

$$= \int_{0}^{1} (6x - 6x^2)\,dx \quad \left[\because \int x^n dx = \frac{x^{n+1}}{n+1}\right]$$

$$= 6\int_{0}^{1} (x - x^2)\,dx$$

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

$$= 6\left[\frac{1^2}{2} - \frac{1^3}{3} - 0 - 0\right]$$

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

$$= 6\left(\frac{3-2}{6}\right)$$

$$= 6\left(\frac{1}{6}\right)$$

$$\int_{0}^{1} f(x)dx = \underline{\underline{1}}$$

Hence $f(x)$ is p.d.f.

ii) $\text{mean}(\bar{x}) = \int_{-\infty}^{\infty} x f(x)\, dx$

$$= \int_{0}^{1} x\left[6x(1-x)\right] dx$$

$$= 6\int_{0}^{1} (x^2 - x^3)\, dx$$

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

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

$= 6\left[\frac{4-3}{12}\right] = \cancel{6}\left(\frac{1}{\cancel{12}_2}\right) = \frac{1}{2}$

ii) Variance $= \int_{-\infty}^{\infty} (x-\bar{x})^2 f(x)\,dx$

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

$\left[(a-b)^2 = a^2 - 2ab + b^2\right]$

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

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

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

$= 6\left[\frac{x^4}{4} - \frac{x^3}{3} + \frac{x^2}{8} - \frac{x^5}{5} + \frac{x^4}{4} - \frac{x^3}{4\times 3}\right]_0^1$

$= 6\left[\frac{1}{4} - \frac{1}{3} + \frac{1}{8} - \frac{1}{5} + \frac{1}{4} - \frac{1}{12}\right]_0^1$

$= 6\left[0.25 - 0.33 + 0.13 - 0.2 + 0.25 - 0.083\right]$

$= 6\left[0.017\right]$

$= 0.102$

24/2/25

When the random variable discrete we have the following theoritical distribution

i) Bernoulli's distribution

ii) Binomial

iii) Poisson's

iv) Hypogeometric

When the random variable is continous whe have the following theoretical distribution.

i) Normal
ii) Uniformal
iii) Gamma

Binomial distribution.

This dist is effective when an even repeating "n" trail of the random experiment.

A probabilty dist which has the following pmf is called as binomial dist.

$$p(x) = nC_x p^x q^{n-x}, \quad x = 0, 1, 2 \dots n$$
$$0 < p < 1$$
$$q = 1 - p$$

Mean = np

Variance = $npq \Rightarrow S.D = \sqrt{npq}$.

Problems

1) The probability of defective item of bolt in a total of 400 is 0.1. Find the mean and S.D of binomial distribution.

$n = 400 \quad p = 0.1$

$q = 1 - p$

$= 1 - 1.0$

$= 0.9$

Mean $= np$

$= 400(0.1)$

$= 40$

$S.D = \sqrt{npq}$

$= \sqrt{400 \times 0.1 \times 0.9} = \sqrt{36} = 6$

2) The probability of a ship being destroyed on a certain voyage is 0.02. The company owns 6 ships. find the probability of

i) Losing one ship
ii) losing at most 2 ships
iii) losing none in a voyage.

→ p = problem of losing a ship = 0.02

$q = 1 - 0.02 = 0.98$

$n = 6$

Binomial dist

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

$= 6C_x \, (0.02)^x (0.98)^{6-x}$

$nC_x = \frac{n!}{x!(n-x)!}$

i) probability of one ship = $(p(x=1))$

$= 6C_1 \, (0.02)^1 \, (0.98)^{6-1}$

$= 6(0.02)(0.98)^5$

$p(x=1) = \underline{\underline{0.1084}}$

$= \frac{6!}{1(6-1)}$

$= \frac{6 \times \cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2}}{\cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2}}$

$= 6$

ii) propability of losing at most (max) 2 ships $(p(x \leq 2))$

$= p(x=0) + p(x=1) + p(x=2)$

$= \left[6C_0 (0.02)^0 (0.98)^{6-0}\right] + \left[6C_1 (0.02)^1 (0.98)^{6-1}\right]$

$+ \left[6C_2 \; (0.02)^2 \; (0.98)^{6-2}\right]$

$\frac{6!}{2!\,6-2!}$

$= \left[1 \times (0.02)^0 (0.98)^6\right] + 0.1084 + \left[15 \times 0.0004 \times 0.9223\right]$

$\frac{6!}{2!\,4!}$

$\frac{\cancel{6}^{3} \times 5 \times \cancel{4} \times \cancel{3} \times \cancel{2} \times 1}{\cancel{2}!\, \cancel{4} \times \cancel{3} \times \cancel{2}}$

$= \left[1 \times (0.02)^0 \times 0.8858\right] + 0.1084 + \left[0.0055\right]$

$= \underline{\underline{0.9997}}$

iii) probability of losing none = $p(x=0)$

$= [6C_0 (0.02)^0 (0.98)^{6-0}]$

$= [1 \times 1 \times 0.8858]$

$= 0.8858$

3) Every one call out of 15 telephone calls b/w 2 pm to 4pm of a week is busy. Find the probability that out of 6 randomly selected telephone number called are

i) exactly 2 calls are busy

ii) not more than 3 calls are busy

iii) atleast 3 of them are busy.

$\Rightarrow p$ = probability of tele calls are busy = $1/15$

$q = 1 - p = 1 - \frac{1}{15} = \frac{14}{15}$

$6C_2 \to 6$ + shift + $nCr(\frac{0}{\cdot})$ + 2.

The binomial dist is

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

$= 6C_x \left(\frac{1}{15}\right)^x \left(\frac{14}{15}\right)^{6-x}$

$\frac{6!}{2!\,6-2!}$

$\frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{2! \times 4!}$

$= \frac{\overset{3}{\cancel{6}} \times 5 \times \cancel{4} \times \cancel{3} \times \cancel{2}}{\cancel{2} \times \cancel{4} \times \cancel{3} \times \cancel{2}}$

$= 15$

i) P(exactly 2) = $p(x=2)$

$= \left(6C_2 \left(\frac{1}{15}\right)^2 \left(\frac{14}{15}\right)^{6-2}\right)$

$= \cancel{15} \cancel{(0.0043)} \cancel{(0.7587)}$

$= \overset{1}{\cancel{15}} \left(\frac{1}{\underset{15}{\cancel{225}}}\right) \left(\frac{38416}{50625}\right)$

$\Rightarrow \frac{1}{15} \times 0.75883$

$= 0.0505$

ii) p(not more than 3 calls are busy)

$= p(x=0) + p(x=1) + p(x=3) + p(x=3)$

$$= \left[6C_0 \left(\frac{1}{15}\right)^0 \left(\frac{14}{15}\right)^{6-0}\right] + \left[6C_1 \left(\frac{1}{15}\right)^1 \left(\frac{14}{15}\right)^{6-1}\right]$$

$$+ \left[6C_2 \left(\frac{1}{15}\right)^2 \left(\frac{14}{15}\right)^{6-2}\right] + \left[6C_3 \left(\frac{1}{15}\right)^3 \left(\frac{14}{15}\right)^{6-3}\right]$$

$= \cancel{(1 \times 1 \times 0.65981)}$

$= 0.66102 + 0.2833 + 0.0506 + 0.0048$

$= \underline{0.99972}$

iii) p(atleast 3 of them are busy) $\Rightarrow p(x \geq 3)$

$= 1 - p(x=3) + p(x=4) + p(x=5) + p(x=4)$

$$1 - \left[6C_3 \left(\frac{1}{15}\right)^3 * \left(\frac{14}{15}\right)^{6-3}\right] + \left[6C_4 \left(\frac{1}{15}\right)^4 * \left(\frac{14}{15}\right)^{6-4}\right] +$$

$$\left[6C_5 \left(\frac{1}{15}\right)^5 * \left(\frac{14}{15}\right)^{6-5}\right] + \left[6C_6 \left(\frac{1}{15}\right)^6 * \left(\frac{14}{15}\right)^{6-6}\right]$$

$\frac{6!}{3!\,6-3!}$

$\frac{\cancel{6} \times 5 \times \cancel{4} \times \cancel{3} \times \cancel{2}}{\cancel{3} \times \cancel{2} \times \cancel{2} \times 3}$

$5 \times 4 = 20$

$1 - [20 \times 0.00021 \times 0.93]$

* 12 coins are tossed 6 times in how many cases 8 heads & 4 tails are obtained.

→ Here $n = 12$ $N = 256$

8 heads & 4 tails.

$p = \frac{1}{2}$ $q = \frac{1}{2}$

The binomial distribution is

$$p(x) = N\left[nC_x \, p^x \, q^{n-x}\right]$$

$$p(\text{getting 8 heads}) = 256\left[12C_8\left(\frac{1}{2}\right)^8\left(\frac{1}{2}\right)^{12-8}\right]$$

$$= 256\left[12C_8\left(\tfrac{1}{2}\right)^8\left(\tfrac{1}{2}\right)^4\right]$$

$\frac{12!}{8!\,12-8!}$

$\frac{12\times 11\times 10\times 9\times 8\times 7\times 6\times 5\times 4\times 3\times 2}{8\times 7\times 6\times 5\times 4\times 3\times 2\times 4\times 3\times 2}$

$$= \cancel{256}\left[495 \times \frac{1}{\cancel{256}} \times \frac{1}{16}\right]$$

$$= \frac{495}{16}$$

$$= 30.9375$$

$$= \underline{31}$$

$$p(\text{getting 4 tails}) = 256\left[12C_4\left(\tfrac{1}{2}\right)^4\left(\tfrac{1}{2}\right)^{12-4}\right]$$

$$= \cancel{256}\left[495 \times \frac{1}{16} \times \frac{1}{\cancel{256}}\right]$$

$$= \frac{495}{16}$$

$$= 30.9375$$

$$= \underline{31}$$

5) The no of telephone lines busy after instant of time this is a binomial variant with programming 0.2 if 10 lines are chosen as random find the probability that i) 5 lines are busy
ii) atmost 2 lines are busy.

⇒

$p = 0.2 \quad q = 1 - p = 1 - 0.2 = 0.8$

$n = 10$

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

$= 10C_x (0.2)^x (0.8)^{10-x}.$

i) $p(x=5) = 10C_5 (0.2)^5 (0.8)^{10-5}$

$= 252 \, (0.00032)\,(0.32768)$

$= 0.0264$

$\frac{10!}{5!\,10-5!}$

$= \frac{\cancel{10} \times \cancel{9} \times \cancel{8} \times 7 \times 6 \times \cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2}}{\cancel{2} \times \cancel{3} \times \cancel{4} \times \cancel{5}!\;\cancel{5} \times \cancel{4} \times \cancel{3} \times 1}$

ii) $p(x \le 2) = p(x=0) + p(x=1) + p(x=2)$

$= 10C_0 (0.2)^0 (0.8)^{10-0} + 10C_1 (0.2)^1 (0.8)^{10-1}$

$+ \; 10C_2 (0.2)^2 (0.8)^{10-2}$

$= (1 \times 1 \times 0.10737) + (10 \times 0.2 \times 0.13421) + (45 \times 0.04 \times 0.1677)$

$\frac{10!}{2!\,10-2!}$

$\frac{\overset{5}{\cancel{10}} \times 9 \times \cancel{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2}}{\cancel{2}!\; \cancel{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2}}$

$= 0.10737 + 0.26842 + 0.30186$

$= 0.67765$

6) The mean of the binomial distribution is 6 the standard deviation is $\sqrt{2}$ find n & p.

⇒ Mean = 6

SD = $\sqrt{2}$

$npq = \text{Variance} = (SD)^2 = (\sqrt{2})^2 = 2.$

$$\frac{\text{Variance}}{\text{Mean}} = \frac{npq}{np} = \frac{\cancel{2}}{\cancel{6}_3} = \frac{1}{3}$$

$$\Rightarrow q = \frac{1}{3}$$

$$q = 1 - p$$

$$p = 1 - q \Rightarrow 1 - 1/3 = \underline{\underline{2/3}}$$

$p = 2/3$ & mean $= 6$

$np = \text{mean} = 6$

$np = 6$

$n(2/3) = 6 \Rightarrow n = \cancel{6}^3 (3/\cancel{2})$

$$\underline{\underline{n = 9}}$$

2) In a large consinement of electric lamps 5% as defective a random sample of 8 lamps are taken for inspection. What is the probability that one or more defective.

$\Rightarrow p = 5\% = \frac{5}{100} = 0.05$

$n = 8 \qquad q = 1 - p$

$= 1 - 0.05 = \underline{\underline{0.95}}$

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

$= 8C_x (0.05)^x (0.95)^{n-x}.$

P(one or more defects) $= 1 - P(\text{no of no defective})$

$= 1 - P(x = 0)$

$= 1 - [8C_0 (0.05)^0 (0.95)^{8-0}]$

$= \underline{\underline{0.3366}}$

Poisson distribution

A DRV 'X' is said to have the non integer negative value with probability density func is known as poisson distribution.

$$p(x) = P(X = x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!}, & x = 0, 1, 2, \\ 0, & \text{otherwise} \end{cases}$$

where λ is parameter of dist

Mean $= E(X) = \lambda$

Variance $= \lambda$

1) If X has the poisson dist and $p(X=0) = p(X=1) = k$ show that $k = \frac{1}{e}$

$\Rightarrow$ Since X is poisson variate we have

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

Gvn $p(X=0) = p(X=1) = k$

$$\frac{e^{-\lambda}\lambda^0}{0!} = \frac{e^{-\lambda}\lambda^1}{1!}$$

$$\frac{\cancel{e^{-\lambda}}(1)}{1} = \frac{\cancel{e^{-\lambda}}\lambda}{1}$$

$$\Rightarrow 1 = \lambda \Rightarrow \underline{\underline{\lambda = 1}}$$

Let $p(x=1) = k$

$$\frac{e^{-\lambda}\lambda^1}{1!} = k \quad [\because \lambda = 1]$$

$$= \frac{e^{-\lambda}(1)^1}{1} = k$$

$e^{-\lambda} = k$

$e^{-1} = k$

$k = e^{-1}$

$\underline{\underline{k = \frac{1}{e}}}$

Assuming that 1 in 80 birth in case of twins. Calculate the probability of 2 or more ~~girls~~ births of twins on a day when 30 births occurs use

i) Binomial dist

ii) poisson

→ $p = \frac{1}{80} = 0.0125 \quad n = 30 \quad q = 1 - p = 1 - 0.0125$
$= 0.9875$

i) Bino dist

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

$= 30C_x (0.0125)^x (0.9875)^{30-x}$.

$P(x \geq 2) = 1 - P(x < 2)$

$= 1 - [P(x=0) + P(x=1)]$

$= 1 - \left[30C_0 (0.0125)^0 (0.9875)^{30-0}\right] + \left[30C_1 (0.0125)^1 (0.9875)^{30-1}\right]$

$= 1 - [1 \times 1 \times 0.68566] + [30 \times 0.0125 \times 0.69434]$

$= 1 - [0.68566 + 0.26037]$

$= 1 - 0.94603$

$= 0.05397$

$= 0.054$

ii) Poisson dist

Mean $= \lambda = np \Rightarrow \lambda = 30(0.0125)$

$\lambda = 0.375$

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

$$p(x \geq 2) = 1 - p(x < 2)$$
$$= 1 - [P(x=0) + P(x=1)]$$
$$= 1 - \left\{ \frac{e^{-0.375} (0.375)^0}{0!} + \frac{e^{-0.375} (0.375)^1}{1!} \right\}$$

(shift ln → e^x)

$$= 1 - \left\{ \frac{0.687 \times 1}{0!} + \frac{0.687 \times 0.375}{1!} \right\}$$
$$= 1 - \left\{ \frac{0.6873}{1} + \frac{0.2577}{1} \right\}$$
$$= 1 - 0.945$$
$$= \underline{\underline{0.055}}$$

3) The no of persons joining is in a lisense queue a min has poission distribution with parameter 5.8. find the probability that

i) No one joins the queue in particular min
ii) 2 or more persons joined the queue in the minute.

$\Rightarrow \lambda = 5.8$

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad ; \quad x = 0$$

$$= \frac{e^{-5.8} (5.8)^0}{0!} = 3.02 \times 10^{-3}$$
$$= \underline{\underline{0.00302}}.$$

$$= \frac{330.29 \times 1}{1}$$
$$= \underline{\underline{330.29}} \qquad = \underline{\underline{0.003}}$$

The no of accidents occuring in a city for a day is a poison variate with mean 0.8 find the probability that on a randomly selected day
i) There are no accidents
ii) There are accidents

⇒ Mean $= \lambda = 0.8$

$$P(x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-0.8}(0.8)^x}{x!}$$

i) $P(\text{no accident}) = P(x=0)$

$$= \frac{e^{-0.8}(0.8)^0}{0!}$$

$$P(x=0) = 0.449$$

ii) $P(\text{There are accidents}) = 1 - P(\text{no accident})$

$$= 1 - P(x=0)$$

$$= 1 - 0.449$$

$$= 0.551$$

5) 2% of fuses are manufactured by a firm are expected to be defective. Find the probability that a box containing 200 fuses

$n \to$ value is >100 so poison

i) defective fuses
ii) 3 or more defective fuses.

⇒ Gvn - $p = \frac{2}{100} = 0.2$ & $n = 200$

Mean $= \lambda = np = 200(0.02) = 4$

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!} \Rightarrow \frac{e^{-4} (4)^x}{x}$$

i) P(There are defective) = 1 − P(no defective)

$$= 1 - P(x=0)$$

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

$$= 1 - 0.018$$

$$= \underline{\underline{0.982}}$$

ii) P(3 or more defective) = 1 − P(x < 3)

$$= 1 - [P(x=0) + P(x=1) + P(x=2)]$$

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

$$= 1 - \left[0.018 + 0.072 + \frac{0.288}{2} \right]$$

$$= 1 - [0.018 + 0.072 + 0.144]$$

$$= 1 - 0.234$$

$$= \underline{\underline{0.766}}$$

Exponential Dist

The probability density func for exponential dist is

$$P(x) = \begin{cases} \frac{1}{\sigma} e^{-\frac{x}{\sigma}} & \text{for } x > 0 \\ 0 & \text{for } x < 0 \end{cases}$$

& mean = $\bar{x}$ = σ

variance = σ^2

S.D = σ

The lifetime of a certain kind of battery is a random variable which has exponential dist with mean of 200 hrs find the probability that such battery will last i) atmost 100 hours.
ii) last anywhere b/w 400 to 600 hrs.

$p(x) = \frac{1}{\sigma} e^{-\frac{x}{\sigma}}$ and $\sigma = 200$

$= \frac{1}{200} = e^{-\frac{x}{200}}$

i) $P(\text{last atmost 100 hours}) = \int_0^{100} p(x)\, dx$

$= \int_0^{100} \frac{1}{200} e^{-\frac{x}{200}} dx.$

$= \frac{1}{200} \int_0^{100} e^{-\frac{x}{200}} dx$ $\left[\int e^{-\frac{1}{a}} dx = \frac{e^{-\frac{x}{a}}}{-1/a} \right]$

$= \frac{1}{200} \left[\frac{e^{-\frac{x}{200}}}{-\frac{1}{200}} \right]_{x=0}^{100}$

$= -\left[e^{-\frac{x}{200}} \right]_0^{100} = -\left[e^{-\frac{100}{200}} - e^0 \right]$

$= -\left[e^{-0.5} - 1 \right] = 1 - e^{-0.5} = 0.3935$

ii) $P(400 \leq x \leq 600) = \int_{400}^{600} p(x)\, dx$

$= \frac{1}{200} \int_{400}^{600} e^{-\frac{x}{200}} dx$

$= \frac{1}{\cancel{200}} \left[\frac{e^{-\frac{x}{200}}}{-\frac{1}{\cancel{200}}} \right]_{400}^{600}$

$$= \left[e^{-\frac{x}{200}} \right]_{400}^{600} = \left[-e^{-\frac{600}{200}} - \left(-e^{-\frac{400}{200}} \right) \right]$$

$$= -e^{-3} + e^{-2}$$

$$= e^{-3} - e^{-2}$$

$$= 0.135 - 0.049$$

$$= \underline{0.086}$$

2) The length of a telephone conversation has been found to have an exp dist with 3min mean. What is the probability that a call may last

i) more than 1 min

ii) less than 3 min

$\Rightarrow \sigma = 3$ min

$$P(x) = \frac{1}{\sigma} e^{-\frac{x}{\sigma}}$$

$$= \frac{1}{3} e^{-\frac{x}{3}}$$

i) $P(\text{more than } 1 \text{ min}) = p(x > 1)$

$$= 1 - p(x \le 1)$$

$$= 1 - \int_0^1 p(x)\,dx$$

$$= 1 - \left[\frac{1}{3} \int_0^1 e^{-\frac{x}{3}}\,dx \right]$$

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

$$= 1 + \left[e^{-x/3} \right]_0^1$$

$$= 1 + \left[e^{-1/3} - e^0 \right]$$

$$= \cancel{1} + e^{-1/3} \cancel{-1}$$

$$= e^{-1/3}$$

$$= \underline{0.7165}$$

ii) P(less than 3 min) = P(x < 3)

$$= \int_0^2 p(x)\,dx$$

$$= \frac{1}{3}\int_0^2 e^{-x/3}\,dx.$$

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

$$= \left[-e^{-x/3}\right]_0^2$$

$$= -e^{-2/3} - (-e^{0})$$

$$= -e^{-2/3} + 1$$

$$= 1 - e^{-2/3}$$

$$= 0.4866$$

Normal distribution

A continous random variable x is to follow a normal dist if its probability density func is given by $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, $-\infty < x < \infty$ and $\sigma > 0$

hence variable x is called as normal variate.

mean = $E(x) = \mu$

Variance = σ^2

SD = σ

1. Find the area under the normal curves in each of the cases
i) $t = 0$ & $t = 1.48$
ii) $t = 0.78$ & $t = 0$
iii) $t = -0.46$ & $t = 2.21$
iv) $t = 0.81$ & $t = 1.84$
v) To the left of $t = -0.5$
vi) Right of $t = 1.84$
vii) Right of $t = 0.46$

11/3/25

Pr3 2 times mode 1

shift + dist 2

i) $t=0$ & $t=1.48$

$P(0 \le t \le 1.48)$

$= 0.4306$

ii) $t=0.7$ & $t=0$

$P(0 \le t \le 0.78)$

$= 0.2823$

iii) $t=-0.46$ & $t=2.21$

$= P(-0.46 \le t \le 0) + P(0 \le t \le 2.21)$

$= P(0 \le t \le 0.46) + P(0 \le t \le 2.21)$

$= 0.6636$

iv) $t = 0.81$ & $t = \cancel{184}$ $\Rightarrow P(0.8 \le t \le 1.84)$

$= P(0 \le t \le 1.84) - P(0 \le t \le 0.81)$

$= P(z = 1.84) - P(z = 0.81)$

$= P(1.84) - P(0.81)$

$= 0.4671 - 0.2910$

$= 0.1761$

iv) probability (left of $t = -0.5$)

$= 0.5 - P(0 \le 1 \le 0.5)$

$= 0.5 - P(z = 0.5)$

$= 0.5 - 0.1915$

$= 0.3085$

v) probability (right of $t = 0.48$)

$= 0.5 - P(z = 0.48)$

$= 0.5 - 0.1849$

$= 0.3156$

iii) P(Right of t = 1.46) = 0.5 − P(z = 1.46)

= 0.5 − 0.4279

= 0.0721

In a dist which is exactly normal 31% of the items are under 45 and 8% are over 64. Find the mean and standard deviation of the distribution.

i) under 45 = 31% = $\frac{31}{100}$ = 0.31

$\Rightarrow t_1 = 0.31$ ($\because t_1$ is −ve)

$P(-0.31 \le t \le t_1) = 0.5 - P(0 \le t \le t_1)$

$= 0.5 - 0.31$

$= 0.19$ ($\because t$ is −ve)

$= \underline{\underline{-0.19}}$

ii) over 64 = 8% = $\frac{8}{100}$ = 0.08

$P(0.8 \le t \le t_2) = 0.5 - P(0 \le t \le t_2)$

$= 0.5 - 0.08$

$= \underline{\underline{0.42}}$

$t = \frac{x-\mu}{\sigma}$

$t_1 = \frac{45-\mu}{\sigma}$

$-0.19\sigma = 45 - \mu \rightarrow (1)$

$\mu - 0.19\sigma = 45$

$\boxed{\begin{array}{l}\mu = 50.91 \simeq 51 \\ \sigma = 31.15 \simeq 31\end{array}}$

$t = \frac{x-\mu}{\sigma}$

$t_2 = \frac{64-\mu}{\sigma}$

$0.42\sigma = 64 - \mu \rightarrow (2)$

$\mu + 0.42\sigma = 64$

3) Mean life of electric bulbs manufactured by a firm is 1200 hrs and the stand dev is 200 hrs.

i) In a lot of 10000 bulbs, how many bulbs are expected to have a life of 1050 hrs and more

ii) what is the % of bulbs which are expected to fuse 64 1500 hrs of service.

→ Let X denote the life of bulbs which is a normal variate with $\mu = 1200$ hrs & $\sigma = 200$ hrs.

$$Z = \frac{X - \mu}{\sigma}$$

$$Z = \frac{X - 1200}{200}$$

i) $P(X \geq 1050) = P\left[Z \geq \frac{1050 - 1200}{200}\right]$

$= P(Z \geq -0.75)$

$= P(Z = 0) + P(0 \leq Z \leq 0.75)$

$= 0.5 + 0.2734$

$= 0.7734$

In a lot of 10000 bulbs $= N * P(x \geq 1050)$

$= 10000 \times 0.7734$

$= 7734$

ii) $P(\text{less than } 1500 \text{ hrs}) = P\left(Z < \frac{1500 - 1200}{200}\right)$

$= P(Z < 1.5)$

$= P(-\infty \leq Z \leq 0) + P(Z = 1.5)$

$= 0.5 + 0.4332$

$= 0.9332$

The percentage of bulb expected to be fuse $\Big\} = 100 \times 0.9332$

$= 93.32\%$

Height of students is normally dist with mean 165 cm & standard dev 5 cm find the probability that the height of the student is

i) more than 177 cm

ii) less than 162 cm.

The normal variate $z = \frac{x-\mu}{\sigma}$

$$z = \frac{x-165}{5}$$

i) $P(\text{more than } 177\text{cms}) = P(x > 177)$

$$= P\left(z - \frac{177-165}{5}\right)$$

$$= P(z > 2.4)$$

$$= P(0 \le z \le \infty) - P(z = 2.4)$$

$$= 0.5 - 0.4918$$

$$= 0.0082$$

ii) $P(\text{more than } 162\text{cm}) = P(x < 162)$

$$= P\left(z < \frac{162-165}{5}\right)$$

$$= P(z < -0.6)$$

$$= P(-\infty \le z \le 0) - P(0 < z < -0.6)$$

$$= 0.5 - P(z = -0.6)$$

$$= 0.5 - 0.2258$$

$$= 0.2742$$

5) X is normal variate with mean 42 and standard deviation 4 and find the probability that a value taken by x is

i) less than 50
ii) greater than 50
iii) less than 40
iv) greater than 40
v) b/w 43 & 46
vi) b/w 40 & 44
v) b/w 37 & 41 0.2957

⇒ Gvn $\mu = 42$ & $\sigma = 4$

∴ The normal variate $z = \frac{x - \mu}{\sigma}$

$\Rightarrow z = \frac{x - 42}{4}$

i) P(less than 50) = $p(x \leq 50)$

$= p(z < \frac{50 - 42}{4})$

$= P(z < 2)$

$= p(-\infty < z < 0) + P(0 < z < 2)$

$= 0.5 + P(z = 2)$

$= 0.5 + 2.4772.$

$= 0.9772$

ii) p(greater than 50) = $p(x \geq 50)$

$= p(z > \frac{50 - 42}{4})$

$= p(z > 2)$

$= p(0 < z < \infty) - P(0 < z < 2)$

$= 0.5 - P(z = 2)$

$= 0.5 - 0.4772.$

$= 0.0228$

iii) P(less than 40) $= P(x \le 40)$

$= P(z < \frac{40-42}{4})$

$= P(z < -0.5)$

$= P(-\infty < z < 0) - P(0 < z < -0.5)$

$= 0.5 - P(z = -0.5)$

$= 0.5 - 0.1915$

$= 0.3085$

iv) P(greater than 40) $= P(x \ge 40)$

$= P(z > \frac{40-42}{4})$

$= P(z > -0.5)$

$= \cancel{P(-\infty < z} \; P(-0.5 < z < 0) + P(0 < z < \infty)$

$= P(0 < z < -0.5) + P(0 < z < \infty)$

$= 0.1915 + 0.5$

$= 0.6915$

v) P(b/w 43 & 46) $= P(43 \le x \le 46)$

$= P(\frac{43-42}{4} \le x \le \frac{46-42}{4})$

$= P(0.25 \le x \le 1)$

$= P(0 \le z \le 1) - P(0 \le z \le 0.25)$

$= P(0.3413 - 0.0987)$

$= 0.2426$

vi) P(b/w 40 & 44) $= P(40 \le x \le 44)$

$= P(\frac{40-42}{4} \le x \le \frac{44-42}{4})$

$= P(-0.5 \le x \le 0.5)$

$= P(0 \le z \le -0.5)$

$= P(-0.5 \le z \le 0) + P(0 \le z \le 0.5)$

$= 2\, P(0 \le z \le 0.5)$

$= 2\,(P(x = 0.5)$

$= 2(0.1915)$

$= 0.383$

viii) P(b/w 37 & 41) $= P(37 \le x \le 41)$

$= P\left(\frac{37-42}{4} \le x \le \frac{41-42}{4}\right)$

$= P(-1.25 \le x \le -0.25)$

$= P(0 \le x \le 1.25) - P(0 \le x \le -0.25)$

$= 0.3944 - 0.0987$

$= 0.\underline{2957}$

6/3/25

Joint probability function.

The function $p(x(s), y(s))$ is defined by

(small) $p_i = P(X = x_i \cap Y = y_i) = P(x_i, y_i)$

X \ Y	y_1	y_2	....	y_m	Total
x_1	P_{11}	P_{12}	....	P_{1m}	$P_{1.}$
x_2	P_{21}	P_{22}	....	P_{2m}	$P_{2.}$
⋮	⋮				P_3
⋮	⋮				⋮
x_n	P_{n1}	P_{n2}	....	P_{nm}	$P_{m.}$
Total	$P_{1.}$	$P_{2.}$		$P_{.m}$	

$$\sum_{1}^{n} \sum_{i=1}^{m} P(x_i\ y_i) = 1$$

1) For the following bivariate probability distri of X & Y find.

i) $P(x \le 1, Y = 2)$

ii) $P(x \le 1)$

iii) $P(Y = 3)$

iv) $P(Y \le 3)$

v) $P(x < 3, Y < 4)$

X\Y	1	2	3	4	5	6
0	0	0	1/32	2/32	2/32	3/32
1	1/16	1/16	1/8	1/8	1/8	1/8
2	1/32	1/32	1/64	1/64	0	2/64

Soln

X/Y	1	2	3	4	5	6	P(x)
0	0	0	1/32	2/32	2/32	3/32	8/32
1	1/16	1/16	1/8	1/8	1/6	1/8	10/16
2	1/32	1/32	1/64	1/64	0	2/64	8/64
p(y)	3/32	3/32	11/64	13/64	6/32	16/64	1

i) $P(x \le 1, y = 2) = P(x = 0, y = 2) + P(x = 1, y = 2)$

$= 0 + \frac{1}{16} = \frac{1}{16}$

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

$= \frac{8}{32} + \frac{10}{16} = \frac{7}{8}$

iii) $P(Y = 3) = \frac{11}{64}$

3) Given the following bivariate probability dist obtain.

i) marginal dist of X & Y

ii) condn dist of X given Y = 2.

$y \backslash x$	-1	0	1
0	1/15	2/15	1/15
1	3/15	2/15	1/15
2	2/15	1/15	2/15

⇒ soln

$y \backslash x$	-1	0	1	$P(y)$
0	1/15	2/15	1/15	4/15
1	3/15	2/15	1/15	6/15
2	2/15	1/15	2/15	5/15
$P(x)$	6/15	5/15	4/15	1

i) marginal dist. of x

$P(x=-1) = 6/15$

$P(x=0) = 5/15$

$P(x=1) = 4/15$

marginal dist of y.

$P(y=0) = 4/15$

$P(y=1) = 6/15$

$P(y=2) = 5/15$

ii) condn probability.

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

$$P(x=x/y=2) = \frac{P(x-x) \cap P(y=2)}{P(y=2)}$$

$$= \frac{P(x=-1) \cap P(Y=2)}{P(Y=2)} = \frac{\frac{2}{\cancel{15}}}{\frac{5}{\cancel{15}}} = \frac{2}{5}$$

Given the following probability dist

Y \ X	0	1	2
0	0	1/27	2/27
1	2/27	3/27	4/27
2	4/27	5/27	6/27

i) find the marginal dist of X & Y

ii) find the condⁿ dist of X given Y=1.

→

Y \ X	0	1	2	P(y)
0	0	1/27	2/27	3/27
1	2/27	3/27	4/27	9/27
2	4/27	5/27	6/27	15/27
P(x)	6/27	9/27	12/27	1

i) marginal distribution of X.

$P(x=0) = 6/27$

$P(x=1) = 9/27$

$P(x=2) = 12/27$

marginal distribution of Y.

$P(y=0) = 3/27$

$P(y=1) = 9/27$

$P(y=2) = 15/27$

ii) Condⁿ probability

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

$$P(X=x/Y=y) = \frac{P(x=x) \cap P(y=1)}{P(y=1)}$$

$$= \frac{P(x=0) \cap P(y=1)}{P(y=1)} = \frac{\frac{2}{\cancel{27}}}{\frac{9}{\cancel{27}}} = \frac{2}{9}$$

Unit - 4

Sampling theory

Population :-
A group or collection of objects is called as population or universe.

Sample :-
A finite subset of a population and the no of objects in a sample is called as sample size.

The inherent or unavoidable error in any such approximation is known as sample error.

The constants in the population
Eg:- mean, SD etc are called as parameters.

On the other hand if we find the mean, SD of the sample then it is called as statistics.

parameter is a statistical constant of the popln and statistics is the function of sample values.

The mean of different samples in a popln tabulated in the form of frequency dist then the resulting dist is a sampling dist

standard deviation of sampling dist is called as standard error (SE)

Testing of hypothesis

- Hypothesis is a statement about the values of the parameter

- Testing of hypothesis is a procedure for deciding whether to accept or reject the hypothesis

The hypothesis which is being tested for possible rejection is null hypothesis.

Denoted by 'H_0'

If the null hypothesis is found to be false with another hypothesis which contradicts the null hypothesis is known as alternative hypo

Denoted by 'H_1'

19/3/25

Critical region

The set of samples which lead to the rejection of null hypothesis is called as critical region.

Type 1 and Type 2 errors.

Actual Sample	Decision	Error.
H_0 is true	Accept	correct dec
H_0 is true	Reject	Type I error
H_0 is not true	Accept	Type II error
H_0 is not true	Reject	Correct decision

Error of a first type (Type - I error) is taking a wrong decision to reject when the null hypothesis is actually true.

Error of a second type (Type 2 error) is taking a wrong decision to accept when the null hypothesis is actually not true.

Level of significance
is the probability of rejection of null hypothesis when it is actually true usually the level of significance is fixed at 0.05 & 0.01 in other words 5% or 1%.

Denoted by α.

If $\alpha = 0.05$ (L.S)
then the critical value k is 1.96

If $\alpha = 0.01$ (LS)
k is 2.58

1 tail & 2 tail test.

While testing H_0 if the critical region is considered at one tail of the sample distribution then the test is a one tail test.

If the C.R is considered at both the tails of the sampling dist then the test is 2 tail test.

Test for mean.

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$\bar{x}$ = sample mean
μ = population mean
σ = S.D
n = no of object.

Test for equality mean.

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

If $|z|_{cal} > k \Rightarrow H_0$ is rejected

& $|z|_{cal} \leq k \Rightarrow H_0$ is accepted.

Test for proportion

$$z = \frac{\bar{x} - P}{\sqrt{\frac{PQ}{n}}}$$, $\bar{x}$ = mean
P = proportion
P, Q - probabilistic value.
n - no of objects.

Test for equality of proportion

$$z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$ where $p_1 = \frac{x_1}{n_1}$

$p_2 = \frac{x_2}{n_2}$.

Problems

1) The mean & S.D of boys of a college are 47 kgs and 3.1 kgs respectively.

The mean & S.D of weight of girls of a college are 45 kgs and 2.8 kgs resp. from the college 16 boys & 9 girls are randomly selected.

i) Find the mean & S.D of 16 selected boys.
ii) Find the mean & S.D mean weight of 9 selected girls.
iii) Find the mean & S.D of difference of mean weight of selected boys & girls.

$\Rightarrow$ $\mu_1 = 47$ $\sigma_1 = 3.1$

$\mu_2 = 45$ $\sigma_2 = 2.8$

$n_1 = 16$ $n_2 = 9$

i) Mean $\bar{x}_1 = \mu_1 = 47$

$$S.D = \frac{\sigma_1}{\sqrt{n_1}} = \frac{3.1}{\sqrt{16}} = 0.775$$

ii) Mean $\bar{x}_2 = \mu_2 = 45$

$$S.D = \frac{\sigma_1}{\sqrt{n_2}} = \frac{45}{\sqrt{9}} = 0.933$$

iii) Mean of $(\bar{x}_1 - \bar{x}_2) = \mu_1 - \mu_2$

$= 47 - 45 = 2$ kgs

$$S.D \text{ of } (\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$= \sqrt{\frac{(3.1)^2}{16} + \frac{(2.8)^2}{9}}$$

$$= \sqrt{\frac{9.61}{16} + \frac{7.84}{9}}$$

$$= \sqrt{0.6004 + 0.8711}$$

$$= \sqrt{1.4711}$$

$= 1.212$ kgs

2) A coin is tossed 144 times and a person gets 80 heads can we say that the coined is unbiased.

$\Rightarrow$ H_0: The coin is unbiased

$n = 144$, P(getting head) $= \frac{1}{2}$

$\Rightarrow P = \frac{1}{2}$ & $Q = 1 - P$

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

X = no of success = no of heads 80.

$$Z = \frac{X - np}{\sqrt{n(PQ)}}$$

$$= \frac{80 - (144)(1/2)}{\sqrt{144(1/2)(1/2)}}$$

$z = 1.33$

$|z| = 1.33 < k = 1.96$ at 5% (L.S)

$\therefore$ Ho is accepted

Hence the coin is unbiased.

20/3/25

3) In a city, A 32% of voters voted for party X, in a city B, 29% of voters voted for party X.

i) Among 70 randomly selected voters from city A, if P_1 is the proportion of votes find the standard error of P_1

ii) Among 60 randomly selected votes from city B, if P_2 is the proportion of votes find the standard error.

iii) find the standard error of $P_1 - P_2$.

City A

$P_1 = 32\% = \frac{32}{100} = 0.32$

$Q_1 = 1 - P_1 = 1 - 0.32 = 0.68$

$n_1 = 70$

$P_1)$ $SE = \sqrt{\frac{P_1 Q_1}{n_1}}$

$= \sqrt{\frac{(0.32)(0.68)}{70}}$

$= \sqrt{\frac{0.2176}{70}}$

$= \sqrt{0.0031}$

$= 0.05567$

City B

$P_2 = 29\% = \frac{29}{100} = 0.29$

$Q_2 = 1 - P_2 = 1 - 0.29 = 0.71$

$n_2 = 60$

$SE(P_2) = \sqrt{\frac{P_2 Q_2}{n_2}}$

$= \sqrt{\frac{(0.29)(0.71)}{60}}$

$= \sqrt{\frac{0.2059}{60}}$

$= \sqrt{0.00343}$

$= 0.0585$

iii) $SE(P_1 - P_2) = \sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}$

$= \sqrt{\frac{(0.32)(0.68)}{70} + \frac{(0.29)(0.71)}{60}}$

$= \sqrt{0.0031 + 0.00343}$

$= \sqrt{0.00653}$

$= \underline{0.0808}$

4) The proportion of women in a city is 0.48 %. among 64 randomly selected people of the society let P_1 will be proportion of women, in another selection of 86 people let P_2 be the proportion of women.

i) standard error of (P_1)

ii) $SE(P_2)$

iii) SE of diff $(P_1 - P_2)$

i) $P_1 = 0.48$

$n_1 = 64$

$Q_1 = 1 - P_1 = 1 - 0.48$
$= 0.52$

$SE(P_1) = \sqrt{\frac{P_1 Q_1}{n_1}}$

$= \sqrt{\frac{(0.48)(0.52)}{64}}$

$= \sqrt{\frac{0.2496}{64}}$

$= \sqrt{0.0039}$

$= \underline{0.0624}$

ii) $P_2 = 0.48$

$n_2 = 86$

$Q_2 = 1 - P_2 = 1 - 0.48$
$= 0.52$

$SE(P_2) = \sqrt{\frac{P_2 Q_2}{n_2}}$

$= \sqrt{\frac{(0.48)(0.52)}{86}}$

$= \sqrt{\frac{0.2496}{86}}$

$= \sqrt{0.0029}$

$= \underline{0.0538}$

iii) $SE(P_1 - P_2) = \sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}$

$= \sqrt{\frac{(0.48)(0.52)}{64} + \frac{(0.48)(0.52)}{86}}$

$= \sqrt{0.0039 + 0.0029}$

$= \sqrt{0.0068}$

$= \underline{\underline{0.0824}}$

5) Its required to Test the hypothesis that an average height of punjabis are 180 cm, for this a random sample containing of 50 punjabis are considered the mean and S.D of heights of the these are found to be 178.9 cm and 3.3 based on these data would you conclude?
(use 5% level of significance)

$\Rightarrow$ ~~X = no of success = no of heads = 80~~

$\mu = 180 \text{ cm}$ $\quad n = 50$

$\bar{x} = 178.9 \text{ cm}$ $\quad \sigma = 3.3 \text{ cm}$

H_0: Avg hight of punjabis are 180 cm.

$$|Z| = \frac{|\bar{x} - \mu|}{\sigma/\sqrt{n}}$$

$$= \frac{|178.9 - 180|}{\frac{3.3}{\sqrt{50}}}$$

$$|z|_{cal} = \frac{-1.1}{\frac{3.3}{\sqrt{50}}} = \frac{-1.1}{\frac{3.3}{7.07}} = \frac{-1.1}{0.466}$$

$$|z|_{cal} = 2.36$$

At 5% L.S, the critical value $k = 1.96$

$|z|_{cal} = 2.36 > 1.96$

$\Rightarrow |z| > k$

$\therefore H_0$ is rejected.

Hence the avg height of punjabi are not 180cm.

6) A machine is designed to fill, bottels of 200ml of medicine a sample of 100 bottles are measured had a mean content of 201.3ml & if the SD of fillings known to be 5ml test whether the machine is functioning property (use 5% L.S)

$\Rightarrow$ $\mu = 200ml$ $\quad n = 100$

$\bar{x} = 201.3ml$ $\quad \sigma = 5ml$.

H_0: The machine functioning property

(The machine on an.avg, fills 200ml)

$$|z| = \frac{|\bar{x} - \mu|}{\sigma/\sqrt{n}} = \frac{|201.3 - 200|}{5/\sqrt{100}}$$

$$|z|_{cal} = 2.6$$

At 5% L.S, the critical value $k = 1.96$

$$|z|_{cal} = 2.6 > k = 1.96$$

$$|z| > k$$

$\therefore$ H_0 is rejected

Hence the machine is not functioning property.

7) It is known that IQ of boys has standard deviation 10 and IQ of girls has standard deviati 12. Mean IQ of 200 randomly selected boys is 99 and mean IQ of 300 randomly selected girls is 97. Can it be concluded that on an average boys and girls have the same IQ. (Use 1% L.S)

$\Rightarrow$ Here

$n_1 = 200$ $\quad \bar{x}_1 = 99$ $\quad \sigma_1 = 10$

$n_2 = 300$ $\quad \bar{x}_2 = 97$ $\quad \sigma_2 = 12$.

H_0: The boys & girls have same IQ.

$$|z| = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{|99-97|}{\sqrt{\frac{10^2}{200} + \frac{12^2}{300}}}$$

$$= \frac{2}{\sqrt{\frac{100}{200} + \frac{144}{300}}}$$

$$= \frac{2}{\sqrt{0.5+0.48}} = \frac{2}{\sqrt{0.98}} = \frac{2}{0.9899}$$

$$= 2.02$$

$|z|_{cal} = 2.02 < k = 2.58$.

$\Rightarrow |z|_{cal} < k$

$\therefore H_0$ is accepted.

Hence the boys & girls have the same IQ.

8) A random sample of 500 apples was taken from a large consinement of apples and 65 were found to be bad find the standard error of the population of bad once. In a sample this size. Hence deduce the % percentage of bad apples in the consinement almost certainly lies b/w 8.5 & 17.5.

$\Rightarrow$

$n = 500$, Let x = no of bad apples = 65.

p = proportion of bad ones $= \frac{65}{500} = 0.13$

$q = 1 - p \Rightarrow 1 - 0.13$
$= 0.87$

$$S.E = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.13)(0.87)}{500}}$$

$$= \sqrt{\frac{0.1131}{500}}$$

$$= \sqrt{0.0002262}$$

$$= 0.015$$

Limit of bad apples.

$$\boxed{p \pm k(S.E)}$$

$\Rightarrow 0.13 \pm \overset{(2.58)}{3}(0.015)$

$= 0.13 + 3(0.015)$ & $0.13 - 3(0.015)$

$= 0.175$ & 0.085

$= 17.5 \times 100$ & 8.5×100

$\therefore$ The limits lies b/w 8.5 & 17.5

9) A sample of 900 days taken from a metalogical record of a certain district and 100 of them are found to be foggy. what are the possible limits of % of foggy days in the district.

$\Rightarrow$ $n = 900$

x = no of foggy days = 100

$$p = \frac{100}{900} = \frac{1}{9}$$

$$q = 1 - p = 1 - \frac{1}{9} = \frac{8}{9}$$

$$SE = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(1/9)(8/9)}{900}}$$

$$= \underline{0.0105}$$

The propability limits of the foggy days

$$= p \pm k(SE)$$

$$= \left(\frac{1}{9}\right) \pm 3(0.0105)$$

$\Rightarrow 0.11 + 3(0.0105)$ & $0.111 - 3(0.0105)$

$0.1111 + 0.0315$ & $0.1111 - 0.0315$

$\underline{0.1426}$ & $\underline{0.0796}$

The percentage of limit are b/w

7.96 & $\underline{14.26}$

10) 500 articles from a factory are examined and found to be 2% defective. 800 similar articles from a 2nd factory are found to have only 1.5% defectiveness. can it be reasonable to conclude that the products of first factory are inferior to those of the second factory.

$\Rightarrow n_1 = 500 \quad P_1 = 2\% = \frac{2}{100} = 0.02$

$n_2 = 800 \quad P_2 = 1.5\% = \frac{1.5}{100} = 0.015$

$H_0 = P_1 - P_2$ (proportion of 2 factory are same)

$$|z| = \frac{|p_1 - P_2|}{\sqrt{pq(1/n_1 + 1/n_2)}} \quad \text{where } p = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}.$$

$$p = \frac{500(0.02) + 800(0.015)}{500+800}$$

$$= \frac{10 + 12}{1300} = \frac{22}{1300} = 0.0169 \cong 0.017$$

$$q = 1 - p = 1 - 0.017 = 0.983$$

$$|z| = \frac{(0.02 - 0.015)}{\sqrt{(0.017)(0.983)\left(\frac{1}{500} + \frac{1}{800}\right)}}$$

$$= \frac{0.005}{\sqrt{0.016711(0.002 + 0.00125)}}$$

$$= \frac{0.005}{\sqrt{0.016711(0.00325)}}$$

$$= \frac{0.005}{\sqrt{0.0000543}}$$

$$= \frac{0.005}{0.0073}$$

$$= 0.6849$$

$|z|_{cal} = 0.6849 < k = 1.96$ (5% L.S)

$|z| < k \Rightarrow$

H_0 is accepted.

Conclusion:-

The difference of proportion is not significant at 5% of L.S. Hence the factories are producing the similar product which do not differ in quality. Thus the one is not inferior to other.