Page No.:
Date: / /

09/08/19
Friday

CORRELATION COEFFICIENT

Correlation coefficient is used to measure the degree of relationship between two variables.

Correlation coefficient was developed by Karl Pearson.

It is denoted by 'r'.

It always lies between -1 to $+1$.

When $r = -1$ to 0 it is negative correlation.

When $r = 0$, then there exists No correlation.

When $r = 0$ to $+1$, then it is Positive correlation.

When $r = +1$ then it is Perfect Positive correlation.

When $r = -1$ then it is Perfect Negative correlation.

Positive Correlation:

If there is an increase or decrease of one variable will simultaneously affect the increase or decrease of other variable, then the two variables are said to be Positively correlated.

When both variables are deviating in the same direction, it is called Positive correlation.

example: Income and expenditure
Height and Weight.

Negative correlation :

If there is an increase or decrease of one variable it will simultaneously affect decrease or increase of the other variable, it is negatively correlated.

When both variables are deviating in opposite direction, it is called Negative correlation.

example : Price and Demand
Pressure and volume of

No correlation :

When the variables remain idle or static or no connection, then it is called as No correlation.

example : Knowledge and Height. (No relationship).

13/08/19
Tuesday

Formula :

Find the correlation coefficient from the following data.

X	Y
1	5
2	4
3	3
4	2
5	1

(i)

X	Y	XY	X^2	Y^2
1	5	5	1	25
2	4	8	4	16
3	3	9	9	9
4	2	8	16	4
5	1	5	25	1

Page No.:
Date: / /

15	15	35	55	55

(ii) Find $\bar{x}$ & $\bar{y}$ (meaning of x and y)

$$\bar{x} = \frac{\Sigma x}{n} = \frac{15}{5} = \underline{\underline{3}}$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{15}{5} = 3$$

(iii) Find co-variance of x, y

$$cov(x,y) = \frac{\Sigma xy}{n} - (\bar{x} \times \bar{y})$$

$$= \frac{35}{5} - (3 \times 3)$$

$$= 7 - 9$$

$$= \underline{\underline{-2}}$$

(iv) Find σx (standard deviation of x)

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

$$= \sqrt{\frac{55}{5} - (3)^2}$$

$$= \sqrt{11 - 9}$$

$$= \sqrt{2}$$

(v) Find σy

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

$$= \sqrt{\frac{55}{5} - (3)^2}$$

Page No.: ____
Date: / /

$= \sqrt{11-9}$

$= \sqrt{2}$

(vi) Find 'r'

$$r = \frac{\text{cov}(x,y)}{\sigma x \times \sigma y}$$

$$= \frac{-2}{\sqrt{2} \times \sqrt{2}}$$

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

$$\underline{\underline{r = -1}}$$

2)

x	y
5	7
7	8
12	12
16	24
24	36

(i)

x	y	xy	x^2	y^2
5	7	35	25	49
7	8	56	49	64
12	12	144	144	144
16	24	384	256	576
24	36	864	576	1296
64	87	1483	1050	2129

(ii) Find $\bar{x}$ & $\bar{y}$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{64}{5} = \underline{12.8}$$

Page No.: Date: / /

$$\bar{Y} = \frac{\Sigma y}{n} = \frac{87}{5} = \underline{\underline{17.4}}$$

(iii) $$cov(x,y) = \frac{\Sigma xy}{n} - (\bar{x} * \bar{y})$$

$$= \frac{1483}{5} - (12.8 \times 17.4)$$

$$= 296.6 - 222.72$$

$$= \underline{\underline{73.88}}$$

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

$$= \sqrt{\frac{1050}{5} - (12.8)^2}$$

$$= \sqrt{210 - 163.84}$$

$$= \sqrt{46.16} = \underline{\underline{6.794}}$$

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

$$= \sqrt{\frac{2129}{5} - (17.4)^2}$$

$$= \sqrt{425.8 - 302.76}$$

$$= \sqrt{123.04} = \underline{\underline{11.092}}$$

(vi) $$r = \frac{cov(x,y)}{\sigma x \times \sigma y}$$

$$= \frac{73.88}{6.794 \times 11.092}$$

$$= \frac{73.88}{75.359} = \underline{\underline{0.98}}$$

① Match.

Page No.: Date: / /

14.08.19
Wednesday

Rank Correlation

Rank correlation is used to measure intensity between two sets of Ranking data or categorial data.

It was found by Spearman.

ρ (row) it is used to measure

Formula :

$$\rho = 1 - \left(\frac{6 \Sigma d_i^2}{n^3 - n}\right)$$

d_i = difference

1) X Y ~~R(X)~~ ~~R(Y)~~ ~~d_i~~ ~~R(X) - R(Y)~~ ~~d_i^2~~

X	Y
55	84
50	96
35	45
42	63
26	44

X	Y	R(X)	R(Y)	d_i = R(X) - R(Y)	d_i^2
55	84	1	2	-1	1
50	96	2	1	1	1
35	45	4	4	0	0
42	63	3	3	0	0
26	44	5	5	0	0
					2

Page No.: Date: / /

$$\rho = 1 - \left(\frac{6\Sigma di^2}{n^3 - n}\right)$$

$$= 1 - \left(\frac{6 \times 2}{5^3 \times 5}\right)$$

$$= 1 - \left(\frac{12}{125 - 5}\right)$$

$$= 1 - \left(\frac{12}{120}\right)$$

$$= 1 - 0.10$$

$$= \underline{\underline{0.90}}$$

$$\boxed{\rho = 0.90}$$

Repeated Rank Correlation

X	Y
55	72
44	84
55	84
72	84
86	70

X	Y	R(X)	R(Y)	di (R(X)−R(Y))	di^2
55	72	3.5	4	−0.5	0.25
44	84	5	2	3	9.00
55	84	3.5	2	1.5	2.25
72	84	2	2	0	0
86	70	1	5	−4	16.00
				Σ	27.50

① Correction Factor.

Page No.:
Date: / /

$$\begin{array}{llllll} & & & 3.5 & 3.5 & \\ X: & 86 & 72 & 55 & 55 & 44 \\ & 1 & 2 & \multicolumn{2}{c}{\frac{3+4}{2}} & 5 \end{array}$$

$$\begin{array}{llllll} & 2 & 2 & 2 & & \\ Y: & 84 & 84 & 84 & 72 & 70 \\ & \multicolumn{3}{c}{\frac{1+2+3}{3}} & 4 & 5 \end{array}$$

Formula:

$$\rho = 1 - \left[\frac{6\left[\Sigma d_i^2 + CF_1 + CF_2 + CF_3 \ldots\right]}{n^3 - n} \right]$$

CF = Correction factor

$$CF = \frac{m^3 - m}{12}$$

where, M = The no. of times value repeated.

$\underline{X}$ 55 repeated 2 times

$$CF_1 = \frac{m^3 - m}{12}$$

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

$$= \frac{8 - 2}{12} = \frac{6}{12} = \underline{\underline{0.5}}$$

$\underline{Y}$ 84 repeated 3 times

$$CF_2 = \frac{m^3 - m}{12}$$

$$= \frac{3^3 - 3}{12}$$

$$= \frac{27 - 3}{12} = \frac{24}{12} = \underline{\underline{2}}$$

$2 \rightarrow 0.5$
$3 \rightarrow 2$
$4 \rightarrow 5$

Page No.: ____
Date: / /

$$\rho = 1 - \left[\frac{6\left(\Sigma di^2 + CF_1 + CF_2\right)}{n^3 - n}\right]$$

$$= 1 - \left[\frac{6 \times (27.50 + 0.5 + 2)}{5^3 - 5}\right]$$

$$= 1 - \left[\frac{6 \times 30}{125 - 5}\right]$$

$$= 1 - \left[\frac{180}{120}\right]$$

$$= 1 - 1.5$$

$$= \underline{-0.5}$$

$$\boxed{\rho = -0.5}$$

16/08/19
Friday

Q) Find the Rank correlation coefficient from the following data :

X	Y
55	36
44	44
36	55
22	20
90	32
100	29
75	75

Page No.:
Date: / /

X	Y	R(X)	R(Y)	$d_i = R(X) - R(Y)$	d_i^2
55	36	4	4	0	0
44	44	5	3	2	4
36	55	6	2	4	16
22	20	7	7	0	0
90	32	2	5	-3	9
100	29	1	6	-5	25
75	75	3	1	2	4
					$\cancel{46}$ 58

$$\rho = 1 - \left(\frac{6 \Sigma d_i^2}{n^3 - n}\right)$$

$$= 1 - \left(\frac{6 \times \cancel{46}\,58}{7^3 - 7}\right)$$

$$= 1 - \left(\frac{\cancel{276}\,348}{343 - 7}\right)$$

$$= 1 - \left(\frac{\cancel{276}\,348}{336}\right)$$

$$= 1 - \cancel{0.821}\ 1.036$$

$= \cancel{0.179}$ ~~(or)~~ $\cancel{0.18}$

$\rho = \cancel{0.179}\ \cancel{0.18}$

$= -0.036$

$$\boxed{\rho = -0.036}$$

Page No.:
Date: / /

Q) Find Repeated Rank correlation from the following data :

X	Y
55	96
42	96
36	96
55	84
72	100
84	27

X	Y	R(X)	R(Y)	$di = R(X) - R(Y)$	di^2
55	96	3.5	3	0.5	0.25
42	96	5	3	2	4
36	96	6	3	3	9
55	84	3.5	5	-1.5	2.25
72	100	2	1	1	1
84	27	1	6	-5	25
					41.5

X - 55 repeated 2 times.

84	72	55	55	42	36
		3.5	3.5		
1	2	$\frac{3+4}{2}$		5	6

Y - 96 repeated 3 times

100	96	96	96	84	27
	3	3	3		
1	$\frac{2+3+4}{3}$			5	6

calculation of CF

X 55 repeated 2 times.

$CF_1 = \frac{m^3 - m}{12}$

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

$$= \frac{8-2}{12}$$

$$= \underline{\underline{0.5}}$$

Y - 96 repeated 3 times

$$CF_2 = \frac{m_2^3 - m}{12}$$

$$= \frac{3^3 - 3}{12}$$

$$= \frac{27-3}{12}$$

$$= \underline{\underline{2}}$$

$$\rho = 1 - \left[\frac{6[\Sigma di^2 + CF_1 + CF_2]}{n^3 - n}\right]$$

$$= 1 - \left[\frac{6(41.5 + 0.5 + 2)}{6^3 - 6}\right]$$

$$= 1 - \left[\frac{6 \times 44}{216 - 6}\right]$$

$$= 1 - \left[\frac{264}{210}\right]$$

$$= 1 - 1.26$$

$$= -\underline{\underline{0.26}}$$

$$\boxed{\rho = \underline{\underline{-0.26}}}$$

variable → operator
$y = 2x + 5$
equal constant

13

Page No.: Date: / /

Left side variable - Dependent variable
Right side variable - Independent variable.

Regression Lines :

Regression lines are the lines which are used to estimate the value of dependent variable by giving the value of independent variable.

There are two regression lines:

(i) X on Y.
(ii) Y on X.

X on Y

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

Y on X

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

Q) Find the regression lines from the following data:

X	Y	Y
10	8	12
15	12	30
14	20	20
6	36	10
5	14	25

$r = 0.40$

X	Y	XY	X^2	Y^2
10	8	80	100	64
15	12	180	225	144
14	20	280	196	400
6	36	216	36	1296
5	14	70	25	196
50	90	826	582	2100

Page No.:
Date: / /

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

$$= \frac{50}{5}$$

$$= 10$$

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

$$= \frac{90}{5}$$

$$= 18$$

$$cov(x,y) = \frac{\Sigma xy}{n} - (\bar{x} \times \bar{y})$$

$$= \frac{826}{5} - (10 \times 18)$$

$$= 165.20 - 180$$

$$= -14.8$$

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

$$= \sqrt{\frac{582}{5} - (10)^2}$$

$$= \sqrt{116.40 - 100}$$

$$= \sqrt{16.4}$$

$$= 4.05$$

Page No.:
Date: / /

$$\sigma_y = \sqrt{\frac{\Sigma y^2}{n} - (\bar{y})^2}$$

$$= \sqrt{\frac{2100}{5} - (18)^2}$$

$$= \sqrt{420 - 324}$$

$$= \sqrt{96}$$

$$= \underline{\underline{9.80}}$$

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

$$= \frac{-14.8}{4.05 \times 9.80}$$

$$= \frac{-14.8}{39.69}$$

$$= \underline{\underline{-0.37}}$$

$$\boxed{r = -0.37}$$

Regression lines

X on Y

$$X - \bar{X} = r \times \frac{\sigma_x}{\sigma_y} (Y - \bar{Y})$$

$$X - 10 = -0.37 \times \frac{4.05}{9.80} (Y - 18)$$

$$X - 10 = -0.37 \times 0.413 (Y - 18)$$

$$X - 10 = -0.153 (Y - 18)$$

Page No.:
Date: / /

$$X - 10 = -0.153y + 2.75 + 10$$

$$\cancel{X - 10 = 2.597y}$$

$$\cancel{X = 2.597y + 10}$$

$$X = \underline{-0.153y + 12.75}$$

Y on X

$$y - \bar{y} = r \times \frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

$$y - 18 = -0.37 \times \frac{9.80}{4.05}(x - 10)$$

$$y - 18 = -0.37 \times 2.42 (x - 10)$$

$$y - 18 = -0.8952 (x - 10)$$

$$y = -0.8952x + 8.952 + 18$$

$$\cancel{y - 18 = 8.05x}$$

$$\cancel{y = 8.06x + 18}$$

$$y = \underline{-0.895x + 26.95}$$

19/08/19
Monday

Q) Find the Regression Lines from the following data:

X	Y
5	6
7	4
12	15
8	5
10	10

Page No.:
Date: / /

X	y	xy	x^2	y^2
5	6	30	25	36
7	4	28	49	16
12	15	180	144	225
8	5	40	64	25
10	10	100	100	100
42	40	378	382	402

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

$$= \frac{42}{5}$$

$$= 8.4$$

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

$$= \frac{40}{5}$$

$$= 8$$

$$cov(x,y) = \frac{\Sigma xy}{n} - (\bar{x} \times \bar{y})$$

$$= \frac{378}{5} - (8.4 \times 8)$$

$$= 75.6 - 67.2$$

$$= 8.4$$

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

$$= \sqrt{\frac{382}{5} - (8.4)^2}$$

Page No.:
Date: / /

$$= \sqrt{76.4 - 70.56}$$

$$= \sqrt{5.84}$$

$$= \underline{2.42}$$

$$\sigma_y = \sqrt{\frac{\Sigma y^2}{n} - (\bar{y})^2}$$

$$= \sqrt{\frac{402}{5} - (8)^2}$$

$$= \sqrt{80.4 - 64}$$

$$= \sqrt{16.4}$$

$$= \underline{4.05}$$

$$r = \frac{\text{cov.}(x, y)}{\sigma x \times \sigma y}$$

$$= \frac{8.4}{2.42 \times 4.05}$$

$$= \frac{8.4}{9.80}$$

$$= \underline{0.86}$$

$$\boxed{r = 0.86}$$

Regression Lines :

X on Y

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

$$x - 8.4 = 0.86 \times \frac{2.42}{4.05} (y - 8)$$

$$x - 8.4 = 0.86 \times 0.597 (y - 8)$$

$$x - 8.4 = 0.5234 (y - 8)$$

$$x = 0.5234 y - 4.16 + 8.4$$

$$\cancel{x - 8.4 = ... - 3.64y}$$

$$\cancel{x = -3.64y + 8.4} \quad x = 0.52y + 4.24$$

Y on X

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

$$y - 8 = 0.86 \times \frac{4.05}{2.42} (x - 8.4)$$

$$y - 8 = 0.86 \times 1.67 (x - 8.4)$$

$$y - 8 = 1.44 (x - 8.4)$$

$$y = 1.44x - 12.10 + 8.$$

$$\cancel{y - 8 = -10.66x}$$

$$\cancel{y = -10.66x + 8.}$$

$$y = 1.44x - 4.10$$

Q) Find the Regression lines from the following data:

X	Y	XY	X^2	Y^2
8	5	40	64	25
12	5	60	144	25
24	5	120	576	25
6	10	60	36	100
10	15	150	100	225
60	40	430	920	400

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

$$= \frac{60}{5}$$

$$= 12$$

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

$$= \frac{40}{5}$$

$$= 8$$

$$\text{cov}(x, y) = \frac{\Sigma XY}{n} - (\bar{X} \times \bar{Y})$$

$$= \frac{430}{5} - (12 \times 8)$$

$$= 86 - 96$$

$$= -10$$

Page No.:
Date: / /

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

$$= \sqrt{\frac{920}{5} - (12)^2}$$

$$= \sqrt{184 - 144}$$

$$= \sqrt{40}$$

$$= \underline{6.32}$$

$$\sigma_y = \sqrt{\frac{\Sigma y^2}{n} - (\bar{y})^2}$$

$$= \sqrt{\frac{400}{5} - (8)^2}$$

$$= \sqrt{80 - 64}$$

$$= \sqrt{16}$$

$$= \underline{4}$$

$$r = \frac{cov(x, y)}{\sigma_x \times \sigma_y}$$

$$= \frac{-10}{6.32 \times 4}$$

$$= \frac{-10}{25.28}$$

$$= \underline{-0.40}$$

Page No.

Date: / /

$$\boxed{r = -0.40}$$

Regression Lines

X on Y

$$x - \bar{x} = r \times \frac{\sigma_x}{\sigma_y} (y - \bar{y})$$

$$x - 12 = -0.40 \times \frac{6.32}{4} (y - 8)$$

$$x - 12 = -0.63 (y - 8)$$

$$x = -0.63y + 5.04 + 12$$

$$x = -0.63y + 17.04$$

Y on X

$$y - \bar{y} = r \times \frac{\sigma_y}{\sigma_x} (x - \bar{x})$$

$$y - 8 = -0.40 \times \frac{4}{6.32} (x - 12)$$

$$y - 8 = -0.25 (x - 12)$$

$$y = -0.25x + 3 + 8$$

$$y = -0.25x + 11$$

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

Page No.:
Date: / /

(Q) Find the Rank correlation coefficient from the following data.

X	Y	R(X)	R(Y)	di (R(x)-R(y))	di^2
7 55	3 80	7	3	4	16
4 72	2 90	3.5	2	1.5	2.25
2 84	1 100	2	1	1	1
8 50	6 64	8	7	1	1
1 ~~90~~ 55	7 64	1	7	-6	36
3 72	8 64	3.5	7	-3.5	12.25
5 64	4 73	5	4	1	1
6 56	5 65	6	5	1	1
					70.5

$$\rho = 1 - \left[\frac{6(\Sigma di^2 + CF_1 + CF_2 \ldots)}{n^3 - n} \right]$$

Calculation of CF

$$CF = \frac{m^3 - m}{12}$$

X

72 repeated 2 times

$$CF_1 = \frac{2^3 - 2}{12}$$

$$= 0.50$$

Y 64 repeated 3 times

$$CF_2 = \frac{3^3 - 3}{12}$$

$$= \underline{\underline{2}}$$

$$\rho = 1 - \left[\frac{6(70.5 + 0.5 + 2)}{8^3 - 8}\right]$$

$$= 1 - \left[\frac{6 \times 73}{504}\right]$$

$$= 1 - \left[\frac{438}{504}\right]$$

$$= 1 - 0.87$$

$$= \underline{\underline{0.13}}$$

$$\boxed{\rho = 0.13}$$

<u>Regression lines</u>

X	Y	XY	X^2	Y^2
55	72	3960	3025	5184
44	84	3696	~~7056~~ 1936	7056
36	55	1980	1296	3025
72	76	5472	5184	5776
84	84	7056	7056	7056
50	27	1350	2500	729
341	398	23514	20997	28826

Page No.:
Date: / /

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

$$= \frac{341}{6}$$

$$= \underline{\underline{56.83}}$$

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

$$= \frac{398}{6}$$

$$= \underline{\underline{66.33}}$$

$$\text{cov}(x,y) = \frac{\Sigma xy}{n} - (\bar{x} \times \bar{y})$$

$$= \frac{23514}{6} - (56.83 \times 66.33)$$

$$= 3919 - 3769.53$$

$$= \underline{\underline{149.47}}$$

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

$$= \sqrt{\frac{20997}{6} - (56.83)^2}$$

$$= \sqrt{3499.5 - 3229.65}$$

$$= \sqrt{269.85}$$

$$= \underline{\underline{16.43}}$$

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

$$= \sqrt{\frac{28826}{6} - (66.33)^2}$$

$$= \sqrt{4804.33 - 4399.67}$$

$$= \sqrt{404.66}$$

$$= \underline{\underline{20.12}}$$

$$r = \frac{\operatorname{cov}(x, y)}{\sigma_x \times \sigma_y}$$

$$= \frac{149.47}{16.43 \times 20.12}$$

$$= \frac{149.47}{330.57}$$

$$= \underline{\underline{0.45}}$$

$$\boxed{r = 0.45}$$

<u>Regression lines</u>

<u>X on Y</u>

$$X - \bar{X} = r \times \frac{\sigma_x}{\sigma_y} (Y - \bar{Y})$$

$$X - 56.83 \times 0.45 \times \frac{16.43}{20.12} (Y - 66.33)$$

$$X - 56.83 = 0.37\ (Y - 66.33)$$

$$X = 0.37Y - 24.54 + 56.83$$

$$X = \underline{0.37Y + 32.29}$$

Y on X

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

$$Y - 66.33 = 0.45 \times \frac{20.12}{16.43}\ (X - 56.83)$$

$$Y - 66.33 = 0.55\ (X - 56.83)$$

$$Y = 0.55x - 31.32 + 66.33$$

$$Y = \underline{0.55x + 35.01}$$

Correlation coefficient

X	Y	XY	X^2	Y^2
10	7	70	100	49
12	8	96	144	64
15	9	135	225	81
14	11	154	196	121
6	10	60	36	100
57	45	515	701	415

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

$$= \frac{57}{5}$$

$$= \underline{11.40}$$

Page No.: Date: / /

$$\bar{Y} = \frac{\Sigma y}{n}$$

$$= \frac{45}{5}$$

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

$$\text{cov}(x, y) = \frac{\Sigma xy}{n} - (\bar{X} \times \bar{Y})$$

$$= \frac{515}{5} - (11.40 \times 9)$$

$$= 103 - 102.6$$

$$= \underline{\underline{0.40}}$$

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

$$= \sqrt{\frac{701}{5} - (11.4)^2}$$

$$= \sqrt{140.2 - 129.96}$$

$$= \sqrt{10.24}$$

$$= \underline{\underline{3.20}}$$

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

$$= \sqrt{\frac{415}{5} - (9)^2}$$

$$= \sqrt{83 - 81}$$

Page No.:
Date: / /

$= \sqrt{2}$

$= 1.41$

$$r = \frac{cov(x, y)}{\sigma_x \times \sigma_y}$$

$$= \frac{0.40}{3.20 \times 1.41}$$

$$= \frac{0.40}{4.51}$$

$= 0.09$

$\boxed{r = 0.09}$

Good

22/08/19
Thursday

Q) From the following regression lines $2x + 3y = 6$ and $x + 2y = 2$ find the following

(i) Mean of x & y

(ii) r.

(iii) If $\sigma_x = 2$ find σ_y

① Regression coefficient of x on y } change
② Regression coefficient of y on x } associated

Snap this
then with someone

① For finding the mean value of x and y solve the given equation and whatever the value comes for x and y be $\bar{x}$ and $\bar{y}$.

$$2x + 3y = 6$$
$$② \times \quad x + 2y = 2$$

$$2x + 3y = 6$$
$$2x + 4y = 4$$
$$-y = 2$$
$$\boxed{y = -2}$$

$$x + 2y = 2$$
$$x + 2(-2) = 2$$
$$x + (-4) = 2$$
$$x = 2 + 4$$
$$\boxed{x = 6}$$

$$\boxed{\bar{x} = 6 \qquad \bar{y} = -2.}$$

② $r = \pm\sqrt{bxy \times byx}$

$$bxy = r \times \frac{\sigma x}{\sigma y}$$
$$byx = r \times \frac{\sigma y}{\sigma x}$$

Page No.:
Date: / /

Substitute -2 in eq. ②

$$2x + 3y = 6.$$

$$2x + 3(-2) = 6$$
$$2x + (-6) = 6$$
$$2x = 6 + 6$$
$$x = \frac{12}{2}$$
$$= 6$$

$$\boxed{x = 6}$$

$\bar{X} = 6$ $\bar{Y} = -2$

② $r = \pm\sqrt{bxy \times byx}$

$bxy = r \times \frac{\sigma x}{\sigma y}$

$byx = r \times \frac{\sigma y}{\sigma x}$

case (i) ① → x on y
② → y on x

$$10x = -5y + 50$$
$$x = \frac{-5}{10}y + \frac{50}{10}$$

$$x = \frac{-1}{2}y + \frac{50}{10}$$

$$bxy = \frac{-1}{2}$$

$3y = -2x + 6$

$y = -\frac{2}{3}x + \frac{6}{3}$

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

$r = \pm\sqrt{bxy \times byx}$

$= \pm\sqrt{\frac{-1}{2} \times \frac{-2}{3}}$

$= \sqrt{0.33}$

$\boxed{r = -0.58}$

(3) $\sigma x = 5 \quad \sigma y = ?$

$bxy = r \times \frac{\sigma x}{\sigma y}$

$\frac{-1}{2} = -0.58 \times \frac{5}{\sigma y}$

$\sigma y = 0.58 \times 5 \times 2$

$= 5.80$

$\boxed{\sigma y = 5.80}$

MBA 90% BCM 25% BIO 98%

why → (reasons)

ANOVA.

Page No.: Date: / /

$$byx = r \times \frac{\sigma_y}{\sigma_x}$$

$$\frac{-2}{3} = \frac{-0.58 \times \sigma_y}{5}$$

$$\frac{10}{3 \times 0.58} = \sigma_y$$

$$\sigma_y = \frac{10}{1.74}$$

$$\boxed{\sigma_y = 5.8}$$

ANOVA : Analysis of variance. Identify the variation.

Analysis of variance is used to compare the means of more than 2 groups. (min. 3 groups.)

ANOVA is developed by prof. R.A. Fisher. It is widely used in Agronomical data. → agriculture.

There are 2 types of ANOVA.

(i) One way Anova - difference by 1 factor
(ii) Two way Anova - difference by 2 factors

→ Only one factor.
→ There are two factors.

eg:- Row & column - two way
Only Row - One way.
or column.

Type of crop.

① Sum of total in the first row.

Summation Σ − +

Product Π − ×

①

$R_1 = 56$	A	8	12	16	5	15
$R_2 = 52$	B	10	15	18	9	–
$R_3 = 50$	C	12	8	15	5	10
$G = 158$						

Fit one way Anova table from the following data:

① Find row total and Grand total.

② Find the CF. (Correction Factor).

$$CF = \frac{G^2}{N \text{ (No. of observations)}}$$

$$= \frac{(158)^2}{14}$$

$$= 1783.14$$

③ Find TSS (Total Sum of Square)

$$TSS = \Sigma\Sigma x_{ij}^2 - CF$$

$$= 8^2 + 12^2 + 16^2 + 5^2 + 15^2 + 10^2 + 15^2 + 18^2 + 9^2 + 12^2 + 8^2 + 15^2 + 5^2 + 10^2 - 1783.14$$

$$= 64 + 144 + 256 + 25 + 225 + 100 + 225 + 324 + 81 + 144 + 64 + 225 + 25 + 100 - 1783.14$$

$$= 2002 - 1783.14$$

$$TSS = 218.86$$

① i^{th} row total

28/08/19
Wednesday

④ $SSR = \sum \frac{R_i^2}{n_i} - CF$

① Sum of square due to Row
i = is used to denote a no. of row.
n_i = no of values in 1^{st} row.

$i = \left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3} - CF\right]$

$= \left[\frac{56^2}{5} + \frac{52^2}{4} + \frac{50^2}{5} - 1783.14\right]$

$= \left[\frac{3136}{5} + \frac{2704}{4} + \frac{2500}{5} - 1783.14\right]$

$= [627.2 + 676 + 500 - 1783.14]$

$= 1803.2 - 1783.14$

$\boxed{SSR = 20.06}$

⑤ $SSE = TSS - SSR$
sum of square due to error.

$SSE = 218.86 - 20.06$
$= 198.80$

$\boxed{SSE = 198.80}$

① ratio must be either 1 or more than 1

One Way Anova Table :

~~Source of variation :~~ change into positive

Source of variation if problem is there	Degree's Of Freedom	Sum of Square
able to solve Row	R - 1 ⇒ 3 - 1 = 2	SSR = 20.06
able to solve Error	N - R = 14 - 3 = 11	SSE = 198.80
Total	N - 1 = 14 - 1 = 13	TSS = 218.86

R - 1 = No. of rows
N - 1 = Total no. of observations.

Degree's Of freedom : defines the no. of independent combination made by researcher.

A + B + C = 30
10 10 ↓
No of Independent combinations — This will be found

Source of variance	Degree's of freedom	Sum of Square	Mean sum of square $= \frac{SS}{df}$	F Ratio ①
Row	R - 1 = 3 - 1 = 2	SSR = 20.06	MSSR = $\frac{SSR}{R-1} = \frac{20.06}{2}$ = 10.03	$\frac{MSSR}{MSSE} = \frac{10.03}{18.07}$ = 0.56
Error	N - R = 14 - 3 = 11	SSE = 198.80	MSSE = $\frac{SSE}{N-R} = \frac{198.80}{11}$ = 18.07	⇒ $\frac{MSSE}{MSSR} = \frac{18.07}{10.03}$
Total	N - 1 = 14 - 1 = 13	TSS = 218.86	*	= 1.80

(1) No. of rows
(2) No of columns
(3) If value Comes less than 1 then numerator to denominator & vice versa i.e., $\frac{MSSE}{MSSR}$

30/8/19
Friday

Two Way Anova :

S.V	DF	SS	MSS	F ratio
Row	(R−1)	SSR	MSSR	F Row = $\frac{MSSR}{MSSE}$
Column	(C−1)	SSC	MSSC	F column = $\frac{MSSC}{MSSE}$
Error	(R−1)(C−1)	SSE	MSSE	
Total	RC−1 or N−1	TSS	*	

Q)

STATES

Salesman	I	II	III	IV
A	15	10	12	8
B	14	15	16	5
C	12	10	10	8

(1) Find row total, column total & Grand total

Row total

A = 45 B = 50 C = 40.

= 135

Column total

I = 41 II = 35 III = 38 IV = 21.

= 135

Grand Total = 135

Page No.:
Date: / /

① rows & columns.

j = column
i = row

② Find the CF

$$CF = \frac{G^2}{N}$$
$$= \frac{(135)^2}{12}$$
$$= \frac{18225}{12}$$
$$= \underline{1518.75}$$

③ Find the TSS

$$TSS = \sum\sum x_{ij}^2 - CF$$
$$= 15^2 + 10^2 + 12^2 + 8^2 + 14^2 + 15^2 + 16^2 + 5^2 + 12^2 + 10^2 + 10^2 + 8^2.$$
$$= 225 + 100 + 144 + 64 + 196 + 225 + 256 + 25 + 144 + 100 + 100 + 64 - 1518.75$$
$$= 1643 - 1518.75$$
$$= \underline{124.25}$$

④ ~~SS~~ Find SSR

$$SSR = \sum \frac{R_i^2}{n_i} - CF$$
$$= \left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3}\right] - CF$$
$$= \left[\frac{45^2}{4} + \frac{50^2}{4} + \frac{40^2}{4}\right] - 1518.75$$
$$= \cancel{\left[\frac{2025 + 2500 + 1600}{4}\right] - 1518.75}$$

$$= \left[\frac{\cancel{1531.25}}{\cancel{12}}\right] - \cancel{1518.75}$$

$$= [506.25 + 625 + 400] - 1518.75$$

$$= 1531.25 - 1518.75$$

$$= \underline{\underline{12.5}}$$

(5) Find SSC

$$SSC = \sum \frac{c_j^2}{n_j} - CF$$

$$= \left[\frac{c_1^2}{n_1} + \frac{c_2^2}{n_2} + \frac{c_3^2}{n_3} + \frac{c_4^2}{n_4}\right] - CF$$

$$= \left[\frac{41^2 + 35^2 + 38^2 + 21^2}{3}\right] - 1518.75$$

$$= \left[\frac{1681 + 1225 + 1444 + 441}{3}\right] - 1518.75$$

$$= \left[\frac{4791}{3}\right] - 1518.75$$

$$= 1597 - 1518.75$$

$$= \underline{\underline{78.25}}$$

(6) Find SSE

$$SSE = TSS - SSR - SSC$$

$$= 124.25 - 12.50 - 78.25$$

$$= \underline{\underline{33.50}}$$

(1) No difference of opinion

(2) (Population) Test → O sample
← affecting the population.

40

S.V	D.F	S.S	M.S.S	F ratio
Row	$(R-1)$ $3-1=2$	12.5	$MSSR = \frac{SSR}{R-1}$ $= \frac{12.5}{2} = 6.25$	$F_{row} = \frac{MSSR}{MSSE}$ $= \frac{6.25}{5.59} = 1.12$
column	$(C-1)$ $4-1=3$	78.25	$MSSC = \frac{SSC}{C-1}$ $= \frac{78.25}{3} = 26.08$	$F_{colm} = \frac{MSSC}{MSSE}$ $= \frac{26.08}{5.59} = 4.67$
Error	$(R-1)(C-1)$ $2 \times 3 = 6$	33.50	$MSSE = \frac{SSE}{(R-1)(C-1)}$ $= \frac{33.5}{6} = 5.59$	
Total	$N-1$ $12-1=11$	124.25		

05/09/19
Thursday

Testing (2) Of Hypothesis

Hypothesis is a general statement about the population parameters which is either true or false.

MBA → Institute → Asan college Best ?

Hypothesis are divided into two types :
> Null Hypothesis (denoted by H_0)
> Alternative Hypothesis (denoted by H_1).

QT calculator (scientific) → Fx casio 100MS - 750
very useful for QT (taking whatever is suggested in mind). — Null Hypothesis. (without interference) (absorb without difference of opinion)

(1) Null Hypothesis - Any Hypothesis which is of no difference Rajini = Kamal
Alternative Hypothesis - Any Hypothesis which is

① maximum 10%. If goes above 10% we have to recheck ② Rejecting good things. 41

Page No:
Date: / /

Opposite to the Null Hypothesis is called Alternative Hypothesis. Rajini ≠ Kamal

> Level of significance: α 1% or 5% or 10% ①

IV $\frac{27}{30}$ (definite knowledge but I are rejecting missing it)

The maximum probability of rejecting the null Hypothesis when the null Hypothesis is true.

alpha
α - percentage of acceptance / Rejection

> Calculated value

Govt. clg $\frac{2}{38}$. asked questions inspection

Judge a person with only 2 persons is not possible. calculated by researcher.

Calculated value is also called test statistic value which is the value calculated by Researcher. Self identified value brought and presented

> Table value

Testing the calculated value with the Tester's knowledge.

eg:- coffee put and given to test

> Result (Inference)

If calculated value is less than Table value accept null Hypothesis (H_0)

If calculated value is greater

than Table value accept Alternative hypothesis (H_1)

17/09/19
Tuesday

sample size $n < 30$.
- t - test for single mean
- t - test for difference between two mean

sample size $n \geq 30$
- z - test for single mean
- z - test for difference between two mean
- z - test for single proportion.
  (or)
  P - test
- P - test for difference between two proportions.

Paired t - test

F test (variance) KW

χ^2 test (Association / Independent) - KW

I Group :- How many groups are there I or II groups.

II Keyword :- Mean, ① proportion are the keywords. (variance is also keyword)

III Sample Size :- $n < 30$	$n \geq 30$
Small sample	Large sample.

* If the sample size < 30 use t - test
* If the sample size $\geq$ 30 use z - test.

(1) degrees of freedom. (4) Sample Standard deviation.
(2) Population mean. (5) Sample size
(3) Sample mean (6) Sample variance

~~t-test for single mean - I group~~

In P-test the sample size is always ≥ 30. P-test cannot be applicable for ~~sample~~ small sample.

t-test for single mean	t-test for diff btw two means
$H_0 : \mu = N_0$. $H_1 : \mu \neq N_0$.	$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$ (or) $\mu_1 > \mu_2$ (or) $\mu_1 < \mu_2$
α : $t_{cal} : \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$	α $t_{cal} : \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} - \dfrac{s_2^2}{n_2}}}$
$t_{tab} : t_{n-1}, \alpha$ df	t_{tab} : $\boxed{t_{n_1+n_2-2}}$, α df. (7) $n_1 - 1$ $n_2 - 1$ $n_1 + n_2 - 2$
Inference. Cal < Tab. accept H_0 Cal > Tab. accept H_1	Inference. Cal < Tab. accept H_0 Cal > Tab. accept H_1

① Parent population
② Population SD.

Page No.:
Date: / /

① Z-test for single mean.	Z-test diff. btw two mean
$H_0 : \mu = NO.$ $H_1 : \mu \neq NO.$	$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$
α :	α : 1% or 5% or 10%
$Z\,cal : \dfrac{\bar{x} - \mu}{② \sigma / \sqrt{n}}$ ↓ compare	$Z\,cal : \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$
Z tab : Z-1% = 2.58 Z-5% = 1.96 Z-10% = 2.33	Z tab : Z-1% = 2.58 Z-5% = 1.96 Z-10% = 2.33
Inference.	Inference.

Q) Find $\bar{x}$ & σ :

I grp.

x	$(x-\bar{x})^2$	
4	$4-6 = -2^2$	4
7	$7-6 = 1^2$	1
9	$9-6 = 3^2$	9
8	$8-6 = 2^2$	4
2	$2-6 = -4^2$	16
30		34

$$\bar{x} = \frac{\Sigma x}{n} = \frac{30}{5} = 6.$$

Page No.: Date: / /

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{\Sigma 34}{5-1}}$$

$$= \sqrt{\frac{34}{4}}$$

$$= 2.92$$

Q) Find $\bar{x}$ & s ?

x	$(x-\bar{x})$	$(x-\bar{x})^2$
12	12-15.4=-3.40	11.56
15	15-15.4=-0.40	0.16
22	22-15.4=6.60	43.56
18	18-15.4=2.60	6.76
10	10-15.4=-5.40	29.16
77		91.20

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

$$= \frac{77}{5}$$

$$= 15.40$$

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{91.2}{5-1}}$$

$$= \sqrt{\frac{91.2}{4}}$$

$$= 4.77$$

Page No.: Date: / /

Q) Find $\bar{x}$ & Δ

x_1	x_2	$(x_1-\bar{x}_1)^2$	$(x_2-\bar{x}_2)^2$
5	4	4	25
7	6	0	9
9	10	4	1
10	20	9	121
4	5	9	16
35	45	26	172

$$\bar{x}_1 = \frac{\Sigma x_1}{n_1} = \frac{35}{5} = 7$$

$$\bar{x}_2 = \frac{\Sigma x_2}{n_2} = \frac{45}{5} = 9$$

$$\Delta_1 = \sqrt{\frac{26}{5-1}} \qquad \sqrt{\frac{\Sigma(x_1-\bar{x}_1)^2}{n_1-1}}$$

$$= \sqrt{\frac{26}{4}}$$

$$= 2.55$$

$$\Delta_2 = \sqrt{\frac{172}{5-1}} \qquad \sqrt{\frac{\Sigma(x_2-\bar{x}_2)^2}{n_2-1}}$$

$$= \sqrt{\frac{172}{4}}$$

$$= 6.56$$

① Mean or expected value or Mean life time.

18/09/19
Wednesday

Q) The life time of electric bulbs for a random sample of 10 from a large consignment gave the following data :-

Item : 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Age in '000 hrs : 4.2, 4.6, 3.9, 4.1, 5.2, 3.8, 3.9,
4.3, 4.4, 5.6.

Can we accept the hypothesis that the average life time of bulbs is 4000 hrs.

X	$(X-\bar{X})^2$
4.2	0.04
4.6	0.04
3.9	0.25
4.1	0.09
5.2	0.64
3.8	0.36
3.9	0.25
4.3	0.01
4.4	0
5.6	1.44
44	3.12

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

$$= \frac{44}{10}$$

$$= 4.4$$

If no value for alpha α is given we have to take 5%.

(1) Found by Gossett [student t distribution]

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{3.12}{10-1}}$$

$$= \sqrt{\frac{3.12}{9}}$$

$$= \underline{0.59}$$

(1) t-test for single mean.

$H_0 = \mu = 4000$ hrs

$H_1 = \mu \neq 4000$ hrs.

$\alpha = 5\%$

$$t_{cal} = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

$$= \frac{4.4 - 4}{0.59/\sqrt{10}}$$

$$= \frac{0.40}{0.59/3.16}$$

$$= \frac{0.40}{0.19}$$

$$= \underline{\underline{2.14}}$$

t table value: t_{n-1}, α df

: $t_{10-1}, 0.05$

: $t_9, 0.05 = \underline{1.833}$

Since tcal > ttab accept alternative Hypothesis (H_1) i.e., the average life time of bulbs is not equal to 4000 hrs i.e., $\mu \neq 4000$ hrs.

Q) Obtain the two regression equations from the following data:

X : 27 27 27 28 28 29

Y : 18 18 19 20 21 21

X	Y	XY	X^2	Y^2
27	18	486 ~~729~~	729	324
27	18	486 ~~729~~	729	324
27	19	513 ~~729~~	729	361
28	20	560 ~~784~~	784	400
28	21	588 ~~784~~	784	441
29	21	609 ~~841~~	841	441
166	117	3242	4596	2291

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

$$= \frac{166}{6}$$

$$= \underline{27.67}$$

Page No.: Date: / /

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

$$= \frac{117}{6}$$

$$= 19.50$$

$$\text{cov}\, x,y = \frac{\Sigma xy}{n} - (\bar{x} \times \bar{y})$$

$$= \frac{3242}{6} - (27.67 \times 19.50)$$

$$= 540.33 - \cancel{548.79}\ 539.57$$

$$= \cancel{8.46}\ 0.76$$

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

$$= \sqrt{\frac{4596}{6} - (27.67)^2}$$

$$= \sqrt{766 - 765.63}$$

$$= \sqrt{0.37}$$

$$= 0.61$$

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

$$= \sqrt{\frac{2291}{6} - (19.50)^2}$$

$$= \sqrt{381.83 - 380.25}$$

Page No.:
Date: / /

$= \sqrt{\cancel{11.40}\ 1.58}$

$= \cancel{8}\ \underline{\underline{1.26}}$

$r = \dfrac{\text{cov } x,y}{\sigma_x \times \sigma_y}$

$= \dfrac{0.76}{0.61 \times 1.26}$

$= \dfrac{0.76}{0.77}$

$= \underline{\underline{0.99}}$

$\underline{\underline{r = 0.99}}$

Regression Lines

X on Y

$X - \bar{X} = r \times \dfrac{\sigma_x}{\sigma_y}(Y - \bar{Y})$

$X - 27.67 = 0.99 \times \dfrac{0.61}{1.26}(Y - 19.5)$

$X - 27.67 = 0.48Y - 9.36$

$X = 0.48Y - 9.36 + 27.67$

$\underline{\underline{X = 0.48Y + 18.31}}$

Y on X

$Y - \bar{Y} = r \times \dfrac{\sigma_y}{\sigma_x}(X - \bar{X})$

$Y - 19.5 = 0.99 \times \dfrac{1.26}{0.61}(X - 27.67)$

$$Y - 19.5 = 2.04x - 56.48$$
$$Y = 2.04x - 56.48 + 19.5$$
$$Y = 2.04x - 36.95$$

Good

24/9/19
Tuesday

Q) Two random samples were drawn from two normal populations and their values are given below :

Sample I	60	65	70	74	76	82	85	87
Sample II	61	66	67	85	78	63	85	86
							88	91

F Test

Test whether the two populations have the same variance at 5 % level of significance. ($F = 3.36$ at 5% level for $V_1 = 10$ and $V_2 = 8$).

F-test is used to test the ratio between 2 variances

① Sample variance of 1^{st} grp

② population variance of 1^{st} grp. people may popul

always for more

Page No.:
Date: / /

only for 2 variance

F-Test (variance)

$H_0 : \sigma_1^2 = \sigma_2^2$ $\quad \boxed{S_1^2}$

~~also~~ H_0 and H_1 should be addressed only for the population not for the sample.

$H_1 : \sigma_1^2 \neq \sigma_2^2$

$\alpha : 5\%$

Fcal : $\dfrac{S_1^2}{S_2^2}$ [always > 1 or = 1]

$$S_1^2 = \frac{\Sigma(X_1 - \bar{X}_1)^2}{n_1 - 1}$$

$$S_2^2 = \frac{\Sigma(X_2 - \bar{X}_2)^2}{n_2 - 1}$$

Ftab : $F_{N-1,\, D-1,\, \alpha}$ df.

x_1	x_2	$(x_1-\bar{x}_1)^2$	$(x_2-\bar{x}_2)^2$
60	61	221.41	256
65	66	97.61	121
70	67	23.81	100
74	85	0.77	64
76	78	1.25	1
82	63	50.69	196
85	85	102.41	64
87	86	146.89	81
	88		121
	91		196
599	770	644.84	1200

$$\bar{x}_1 = \frac{\Sigma x_1}{n_1}$$

$$= \frac{599}{8}$$

$$= 74.88$$

$$\bar{x}_2 = \frac{770}{10} \quad \frac{\Sigma x_2}{n_2}$$

$$= 77$$

$$s_1^2 = \frac{\Sigma(x_1-\bar{x}_1)^2}{n_1-1}$$

$$= \frac{644.84}{8-1}$$

$$= 92.12$$

$$s_2^2 = \frac{\Sigma(x_2-\bar{x}_2)^2}{n_2-1}$$

$$= \frac{1200}{10-1}$$

$$= \underline{133.33}$$

Fcal : $\dfrac{s_1^2}{s_2^2}$

less than 1 $\left[= \dfrac{92.12}{133.33} = \underline{0.69} \right]$ $= \dfrac{133.33}{92.12} = \underline{1.45}$

Ftab : $F_{N-1,\, D-1,\, \alpha\, df}$.

[as Fcal has changed this will also be changed].

$= F_{D-1,\, N-1,\, \alpha\, df}$

$= F_{10-1,\, 8-1,\, 5\%}$

$= F_{9,\, 7,\, 0.05}$

$= \underline{3.68}$

Inference :

Since Fcal < Ftab,

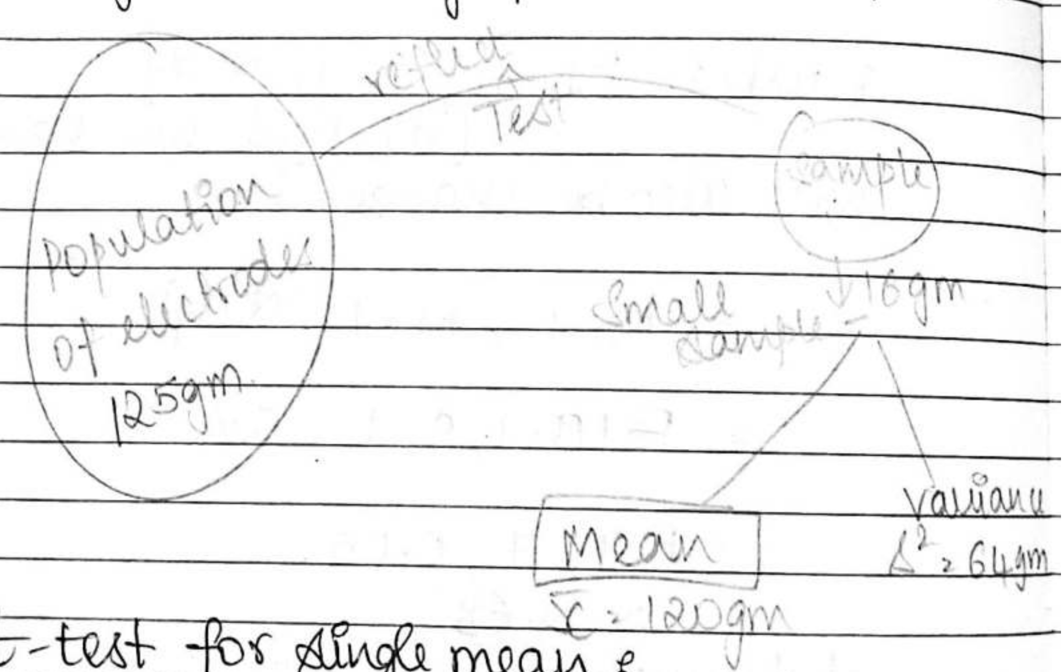
$1.45 < 3.36$

accept H_0. i.e., The two population have same variances.

Page No.: Date: / /

25/9/19
Wednesday

Q) The weight of electrodes purchased by a foundry follows normal distribution. The sales manager of the vendor firm claims that the mean weight of the electrodes is at least 125 gms. The quality manager of the foundry wants to verify this claim. So, he has taken a sample of 16 electrodes. The mean and the variance of the electrodes in the sample are found to be 120 gm (mean) and 64 gm (variance), respectively. Verify the claim of the sales manager at a significance level of 0.01.



t-test for single mean:

$H_0: \mu = 125gms$

$H_1: \mu \neq 125gms.$

$\alpha: 1\% = 0.01$

tcal: $\bar{x} = 120gms$

$s^2 = 64gms$

$s = \underline{\underline{8}}$

Page No.:
Date: / /

$\mu = 125$
$n = 16$.

$$t_{cal} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$= \left| \frac{120 - 125}{8/\sqrt{16}} \right|$$

$$= \left| \frac{-5}{2} \right|$$

$$= \underline{\underline{2.5}}$$

$t_{tab} = t_{\underset{16-1}{n-1}}, \alpha\ df.$

$t_{15, 0.01} = 2.262$ (Approx.).

Inference:
Since $t_{cal} > t_{tab}$
$2.5 > 2.262$
accept alternative hypothesis (H_1).
i.e., $\mu \neq 125$ gms.

The main objective of testing of hypothesis is to check whether the sample will represent the population or not.

Q) ~~There~~ Three varieties of potatoes are being compared for yield. The experiment was carried out by assigning each variety at random to four of 12 equal size plots one being chosen in each of 4 locations. The following yields results in proper unit.

Variety of Potato

Location	A	B	C
1	18	13	12
2	20	23	21
3	14	12	9
4	11	17	10

Test whether there is any difference in the yield.

Sol: ANOVA [one way].

Variety of potatoes

A	B	C
18	13	12
20	23	21
14	12	9
11	17	10

A	18	20	14	11	$R_1 = 63$
B	13	23	12	17	$R_2 = 65$
C	12	21	9	10	$R_3 = 52$
				G	180

$$CF = \frac{G^2}{N} = \frac{(180)^2}{12} = \frac{32400}{12}$$

$$= 2700$$

$$TSS = \Sigma\Sigma x_{ij}^2 - CF$$

$$= 18^2 + 20^2 + 14^2 + 11^2 + 13^2 + 23^2 + 12^2 + 17^2 + 12^2 + 21^2 + 9^2 + 10^2$$

$$= 2938 - 2700$$

$$TSS = 238$$

$$SSR = \frac{\Sigma R_i^2}{n_i} - CF$$

$$i = \left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3}\right] - CF$$

$$= \left[\frac{63^2}{4} + \frac{65^2}{4} + \frac{52^2}{4}\right] - CF$$

$$SSR = 2724.5 - 2700$$

$$= 24.50$$

$$SSE = TSS - SSR$$

$$= 238 - 24.50$$

$$= 213.50$$

Anova Table.

Source of variation	Degree's of freedom	Sum of Square
Row	$R-1 \Rightarrow 3-1 = 2$	
Error	$N-R \Rightarrow 12-3 = 9$	
Total		

One Way ANOVA Table

Source of Variation	Degrees of freedom	Sum of Square	Mean SOS $= \frac{SS}{df}$	F Ratio
Row	$R-1 = 3-1 = 2$	$SSR = 24.50$	$MSSR = \frac{SSR}{R-1} = \frac{24.5}{3-1} = \underline{12.25}$	$\frac{MSSR}{MSSE} = \frac{12.25}{23.72} = 0.52 < 1$ so, $\Rightarrow \frac{MSSE}{MSSR} = \frac{23.72}{12.25} = \underline{1.94}$
Error	$N-R = 12-3 = 9$	$SSE = 213.50$	$MSSE = \frac{SSE}{N-R} = \frac{213.50}{12-3} = \underline{23.72}$	
Total	$N-1 = 12-1 = 11$	$TSS = 238$	*	

H_0: There is no significant difference ~~between~~ in the yield of potatoes.

H_1: There is a significant difference in the yield of potatoes.

α : 5%

Fcal : $F_{row} = \frac{MSSE^{\,9}}{MSSR_{\,2}} = 1.94$

Ftab : $F_{9,2,0.05} = 2.256$ (approx.)

Inference : Since Fcal $\not>$ Ftab accept H_0. i.e, There is no significant difference in the yield of potatoes.

P-test - sample size always ≥ 30.
(Proportion)
① True [biased = cheat].

61

26/9/19
Thursday

P-test or Z-test

P-test proportion test.

$p = \frac{x}{n}$ out of 1000. eg: 1/4 1 out of 4

Numerator will always be less than denominator.

A no. which is always less than 1

Q) A coin is tossed 1000 times and the number of heads turns up is 600 times. Test whether the coin is unbiased or not.

$\frac{1}{2} \times 1000 = 500$ but we are getting 600 times.

$H_0 : P = 1/2 = 0.5$

$H_1 : P \neq 1/2$

$\alpha : 5\%$

$Z_{cal} : \frac{p - P}{\sqrt{PQ/n}}$

$p = \frac{600}{1000} = 0.6$

$P = 0.5$

$n = 1000$

$Q = 1 - P = 1 - 0.5 = 0.5$

Page No.:
Date: / /

$Z_{5\%} = 1.96$
$Z_{1\%} = 2.58$
$Z_{10\%} = 2.33$.

$$= \frac{0.60 - 0.50}{\sqrt{0.5 \times 0.5 / 1000}}$$

$= \underline{0.60}$

$Z_{tab} : Z_{5\%} = \underline{1.96}$.

Inference : Since $Z_{cal} < Z_{tab}$ accept H_0
$\qquad 0.60 \qquad 1.96$

i.e., $H_0 = p = 1/2$ i.e, the coin is unbiased.

Q) In town A 400 people are smokers out of sample of 1200. In another town B 600 people are smokers out of of 1400 samples. Check whether the proportion of smokers are equal or not.

$TA = \frac{400\ x_1}{1200\ n_1} \quad TB = \frac{600\ x_2}{1400\ n_2}$

P test for difference btw 2 proportions.

$H_0 : P_1 = P_2$

$H_1 : P_1 \neq P_2$

$\alpha : 5\%$.

$$Z_{cal} : \frac{p_1 - p_2}{\sqrt{PQ\left(1/n_1 + 1/n_2\right)}}$$

$p_1 = \frac{400 x_1}{1200 n_1} = 0.33$

$p_2 = \frac{600 x_2}{1400 n_2} = 0.43$

$P = \frac{x_1 + x_2}{n_1 + n_2} = \frac{400+600}{1200+1400} = \underline{0.38}.$

$Q = 1 - P$
$= 1 - 0.38$
$= \underline{0.62}$

Zcal : $\left| \frac{0.33 - 0.43}{\sqrt{0.38 \times 0.62 \left(\frac{1}{1200} + \frac{1}{1400}\right)}} \right|$

$= \cancel{\frac{-0.10}{0.24}}$

$= \underline{0.33}$

Ztab : $Z_{5\%} = \underline{1.96}$

Inference : Since Zcal < Ztab accept ~~H~~o.
0.33 < 1.96.
i.e., $\underline{P_1 = P_2}$

① action, Strategy

Page No.:
Date: / /

28/9/19
Saturday

DECISION MAKING ANALYSIS :→

IMBA II MBA ⟨ Finance, Market, HR ⟩ Staff. choose the best

Ds.

Decision making analysis is the process of selecting the best alternatives from the given set of alternatives.

Classified:-

DMA (RCU)

Risk → Sure → Certainity 100% sure risk. it is going to happen.

Uncertainity. unknown Not sure

Home $\xrightarrow{45min}$ Clg 8:25

taking risk to come by 9:00 under 80km/hr. Know what is going to happen due to the risk.

DMA

Probability Known
1. Risk
2. Certainity

taken after knowing what is going to happen.

Probability unknown.
1. Uncertainity.

DMA under Uncertainity.
Models/Method/Principle/criterion:

1. Maximin Criterion ①
2. Maximax criterion ③
3. Lapalace criterion ④ (equal probability)
4. Minimax Regret criterion ⑤
5. Hurwitz Criterion. ⑥

① 40 — old → use as wellwishers.
400 emp.

Maximin Criterion - Pessimistic criterion ②
Maximax - optimistic
Lapalace - criterion of Rationalism
Minimax - Opportunity loss (Regret).
Hurwits - Criterion of realism.

② Ravi → Diwali → Amma - saree ₹10000
Iyyanar → " → " saree.
Total salary to mom ₹5000
wife
50 100 150 ✓

③ Bcom 18K salary → MBA 1L highest, 45K higher.

④ Consider everyone and then take decision. eg: BIKE (Honda shine) consider Family and then get the bike.

(1) Trying to do a work but something is blocking way. An event or a situation which affects the outcomes of the alternative. 66

(2) instigate to do good.

(5) VIPS

	Salary	
	V - 70000	180000
	I - 200000	50000
	P - 60000	190000
	S - 250000	0

Try to Minimize the maximum loss (Regret).

(6) Consider only good and bad. Do not consider all the concept.

class:

6 Good excellent ← Average → 10 ↓ Not good

consider these 2.

(3) If no value for α is given we have to take 0.50 (50%)

Hurwitz Criterion

$$H = \alpha(\max) + (1-\alpha)\min.$$

Where,

α = Optimistic factor value question (2)

Q)

State of Nature (1)

		N_1	N_2	N_3	N_4
Strategy	S_1	27	35	-15	10
	S_2	18	25	20	32
	S_3	15	20	24	12.

The various options given to a decision maker in order to select the best

Payoff take $\alpha = 0.75$ (3) Payoff

(9)

Hari 1 coin

H T

Tyyenar 1 coin H

The game

Explaining

HH → I receive 1 rupee → +1

TH → 0

HT → 0

TT → I pay to Hari → -1.

Payoff is a conditional payment which is a combination of an alternati and state of nature.

Select the best strategy by using the methods of DMA.

Sol:

	N_1	N_2	N_3	N_4	Min	Max
S_1	27	35	-15	10	-15	35
S_2	18	25	20	32	18	32
S_3	15	20	24	12	12	24

(1) Under Maximin criterion the strategy S_2 is the best strategy.

(2) Under Maximax criterion the strategy S_1 is the best strategy.

(3) Lapalace.

$$L_1 = \frac{27+35+(-15)+10}{4} = 14.25$$

$$L_2 = \frac{18+25+20+32}{4} = 23.75$$

$$L_3 = \frac{15+20+24+12}{4} = 17.75$$

Under Lapalace criterion the strategy S_2 is the best strategy.

check for row

(4) Hurwitz

$$H = \alpha(\text{Max}) + (1-\alpha)\text{Min}.$$

(1) $24-(-15) = 39$.

$H_1 = 0.75(35) + 0.25 \times -15$
$= \underline{22.50}$

$H_2 = 0.75(32) + 0.25 \times 18$
$= 28.50$

$H_3 = 0.75(24) + 0.25 \times 12$
$= \underline{21}$

Under Hurwitz criterion the strategy S_2 is the best strategy.

(5) Minimax Regret (loss) Criterion We have to check the column. take Biggest number and subtract the rest with that.

N_1	N_2	N_3	N_4
27	35	-15	10
18	25	20	32
15	20	24	12

Minimax Regret

					Max
S_1	0	0	39 (1)	22	39
S_2	9	10	4	0	10
S_3	12	15	0	20.	20

Under Minimax Regret criterion the strategy S_2 is the best strategy.

① the thing that was supposed to get

69

50 in the steps of 10
50, 60, 70, 8..

Page No.: Date:

(Q20) order size == Demand ⇒ ₹20 profit
① OR > D ⇒ ₹10 profit
opportunity cost. Loss = ₹15
OR < D

Sol:

Demand.

Order size	50	100	150	200	250
① 75 Kg	1250	1125	375	-375	-1125
② 150 Kg	2000	2500	3000	2250	1500
③ 225 Kg	2750	3250	3750	4250	4125
④ 300 Kg	3500	4000	4500	5000	5500

① $(50Kg \times 20) + (25 \times 10) = 1250$
$(75 \times 20) + (25 \times -15) = 1125$
$(75 \times 20) + (75 \times -15) = 375$
$(75 \times 20) + (125 \times -15) = -375$
$(75 \times 20) + (175 \times -15) = -1125$

② $(50 \times 20) + (100 \times 10) = 2000$
$(100 \times 20) + (50 \times 10) = 2500$
$(150 \times 20) = 3000$
$(150 \times 20) + (50 \times -15) = 2250$
$(150 \times 20) + (100 \times -15) = 1500$

③ $(50 \times 20) + (175 \times 10) = 2750$
$(100 \times 20) + (125 \times 10) = 3250$
$(150 \times 20) + (75 \times 10) = 3750$
$(200 \times 20) + (25 \times 10) = 4250$
$(225 \times 20) + (25 \times -15) = 4125$

④ $(50 \times 20) + (250 \times 10) = 3500$
$(100 \times 20) + (200 \times 10) = 4000$
$(150 \times 20) + (150 \times 10) = 4500$
$(200 \times 20) + (100 \times 10) = 5000$
$(250 \times 20) + (50 \times 10) = 5500$

Lapalace Criterion

$75Kg = \frac{1250 + 1125 + 375 + (-375) + (-1125)}{5}$

$= 1250$

$150 Kg = \frac{2000 + 2500 + 3000 + 2250 + 1500}{5}$

$= 2250$

$225 Kg = \frac{2750 + 3750 + 3750 + 4250 + 4125}{5}$

$= 3625$

$300 kg = \frac{3500 + 4000 + 4500 + 5000 + 5500}{5}$

$= 4500$

Thus, under Lapalace Criterion the ~~sta~~ order size 300 Kg is the best strategy.

$X \rightarrow$ Random variable marks

0,1,2, ... 100 $x \rightarrow$ Different values of R:V or success

success rate $p \rightarrow$ Probability of success thinking to attain %

$q = 1 - p$. $q \rightarrow$ probability of failure failed %.

$n \rightarrow$ No. of independent trials.

^nC_x means select x out of n.

$\lambda \rightarrow$ Lamda $\rightarrow$ Average or mean or expected value or variance of poisson distribution.

Random variable is defined as assigning the real number to different outcomes of a trail.

$X - 0, 1$ (only 2 alternatives) – Bernoulli R.V

$X - 0, 1, 2, \ldots\ldots n$ (particular no.) – Binomial R.V. (ended with finite no.)

$X - 0, 1, 2, \ldots\ldots, \infty$ (countless)(endless) (infinite) Poisson R.V.

01/10/19
Tuesday

F-Test

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

$\alpha : 5\%$

$$F_{cal} = \frac{S_1^2}{S_2^2}$$

$$S_1^2 = \frac{\Sigma(x_1 - \bar{x}_1)^2}{n_1 - 1}$$

X_1	X_2	$(X_1 - \bar{X_1})^2$	$(X_2 - \bar{X_2})^2$
6.2	5.6	0	0.01
5.7	5.9	0.25	0.04
6.5	5.6	0.09	0.01
6.0	5.7	0.04	0
6.3	5.8	0.01	0.01
5.8	5.7	0.16	0
5.7	6.0	0.25	0.09
6.0	5.5	0.04	0.04
8.0	5.7	3.24	0
5.8	5.5	0.16	0.04
62	57	4.24	0.24

$$\bar{X_1} = \frac{\Sigma X_1}{n}$$

$$= \frac{62}{10}$$

$$= 6.2$$

$$\bar{X_2} = \frac{\Sigma X_2}{n}$$

$$= \frac{57}{10}$$

$$= 5.7$$

$$S_1^2 = \frac{4.24}{10-1}$$

$$= 0.47$$

$$S_2^2 = \frac{0.24}{10-1}$$

$$= 0.03$$

$F_{cal} = \frac{S_1^2}{S_2^2}$

$= \frac{0.47}{0.03}$

$= 15.67$

$F_{tab} = F_{N-1, D-1, \alpha df}$

$= F_{10-1, 10-1, 0.05}$

$= F_{9, 9, 0.05}$

$= 3.18$

Inference:
15.67 3.18
Since Fcal > Ftab, accept H_1 (alternative hypothesis) i.e. The two yields have same variance.

Pg. 10)
PART C

15) Regression Lines:

X	Y	XY	X^2	Y^2
27	18	486	729	324
27	18	486	729	324
27	19	513	729	361
28	20	560	784	400
28	21	588	784	441
29	21	609	841	441
29	22	638	841	484
29	23	667	841	529
30	24	720	900	576
31	25	775	961	625
285	211	6042	8139	4505

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

$$= \frac{285}{10}$$

$$= 28.5$$

$$\bar{Y} = \frac{\Sigma y}{n}$$

$$= \frac{211}{10}$$

$$= 21.1$$

$$\mathrm{cov}(x,y) = \frac{\Sigma xy}{n} - (\bar{x} \times \bar{y})$$

$$= \frac{6042}{10} - (28.5 \times 21.1)$$

$$= 604.2 - 601.35$$

$$= 2.85$$

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

$$= \sqrt{\frac{8139}{10} - (28.5)^2}$$

$$= \sqrt{813.9 - 812.25}$$

$$= \sqrt{1.65}$$

$$= 1.28$$

Page No.:
Date: / /

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

$$= \sqrt{\frac{4505}{10} - (21.1)^2}$$

$$= \sqrt{450.5 - 445.21}$$

$$= \sqrt{5.29}$$

$$= \underline{\underline{2.30}}$$

$$r = \frac{\text{cov } x, y}{\sigma x \times \sigma y}$$

$$= \frac{2.85}{1.28 \times 2.30}$$

$$= \frac{2.85}{2.94}$$

$$= \underline{\underline{0.97}}$$

<u>RL</u>

<u>X on Y</u>

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

$$X - 28.5 = 0.97 \times \frac{1.28}{2.30}(Y - 21.1)$$

$$X - 28.5 = 0.54Y - 11.39$$

$$X = 0.54Y - 11.39 + 28.5$$

$$x = 0.54y + 17.11$$

y on x

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

$$y - 21.1 = 0.97 \times \frac{2.30}{1.28} (x - 28.5)$$

$$y - 21.1 = 1.74x - 49.59$$

$$y = 1.74x - 49.59 + 21.1$$

$$y = 1.74x - 28.49$$

Good

Page No.:
Date: / /

04/10/19
Friday

CHI Square Test

CHI square test is used to test is there any significant difference between observed & expected frequencies.

$O_i \rightarrow$ observed frequency
$E_i \rightarrow$ Expected frequency

eg: A_1, A_2
Age ⟷ Marital Status

Keyword:

* Association
* Dependent and Independent
* 2x2 contingency table.

eg: Gender M F ; Yes No ; Do you like me

Pg. 21)

21) H_0: The New treatment not superior to conventional treatment.

H_1: The New treatment superior to conventional treatment.

α : 5%

χ^2 =

Calculated value = $\sum \frac{(O_i - E_i)^2}{E_i}$ (founded by Karl pearson)

O_i	$E_i = \frac{RT \times CT}{GT}$	$(O_i - E_i)^2 \div E_i$
140	$= \frac{170 \times 200}{250} = 136$	0.12
30	34	0.47
60	64	0.25
20	16	1.00
		1.84

Page No: ___
Date: / /

$x^2 cal = 1.84$

x tab value $= x^2 (R-1)(C-1) \alpha df$

$\underbrace{(R-1)}_{\text{No. of rows}} \underbrace{(C-1)}_{\text{No. of Column}}$

$= (2-1)(2-1), 0.05$

$= 1 \times 1, 0.05$

$= 1, 0.05$

$= \underline{3.84}$//

Result :

Since $x^2 cal < x^2 tab$; accept Ho. i.e; The new treatment is not superior to the conventional treatment

H_0 : The N.T is not superior to the C.T

H_1 : The N.T is superior to C.T

$\alpha = 5\%$.

$$x^2 cal = \sum \frac{(O_i - E_i)^2}{E_i}$$

O_i	$E_i = \frac{RT \times CT}{GT}$	$(O_i - E_i)^2 \div E_i$
140	136	0.12
30	34	0.47
60	64	0.25
20	16	1.10
		1.84

$\chi_{tab} = \chi^2_{(R-1)(C-1)\,\alpha\,d.f}$

$(2-1)(2-1), 0.05$

$1 \times 1, 0.05$

$1, 0.05$

$= \underline{3.84}$

Result:

Since $\chi^2 cal < \chi^2 tab$, accept H_0 (i.e.)
The New treatment is not superior to the conventional treatment.

Pg. 14)

21)

O_i	E_i	$(O_i - E_i)^2 \div E_i$
315	312.75	0.02
101	104.25	0.10
108	104.25	0.13
32	34.75	0.22
556	556	0.47

$9 : 3 : 3 : 1$

$16x = 556$

$x = \frac{556}{16}$

$= \underline{34.75}$

$\chi\, tab = \chi^2_{(n-1),\,\alpha\, df}$

$4-1, 0.05$

$3, 0.05 = \underline{7.82}$

H_0: There is no evidence to tought the theory at 5% level of significance.

H_1 :- There is a evidence to tought the theory.

Result :

Since $\chi_{cal} < \chi_{tab}$, accept H_0

i.e., there is no evidence.

09/10/19

Wednesday

Baye's Theoreum :-

A, B, C, D, ...

Events

Favourable outcomes.

Eg.... 100 marks - The marks that we expecting is favourable outcomes - what I need.

P(A) Probability of A

P(B) " " B

P(C) " " C

$$P(A) = \frac{n(A)}{n(S)_{\text{total chance}}} = \frac{n_A}{n(S)} = \frac{n_A}{N}$$

1 stick

event = To hit the stick

Probability of an event is defined as the ratio between no. of favourable outcomes to the total no. of outcomes.

① Sure / certain / universal event

∩ - Intersection
∪ - Union.

Page No.: ____
Date: / /

Rules of Probability :-

(ϕ)	① S
No chance of happening: Null or Empty or Void	Surely happening: Sure or Certain or Universal
$P(\phi) = 0.$	$P(S) = 1$

1. Probability of an impossible event ϕ (Phi) $= 0$ i.e.,

$$P(\phi) = 0.$$

2. Probability of an event will always lie between 0 to 1.
   $0 \leq P(A) \leq 1$

3. Probability of possible event $S = 1$ i.e.,

$$P(S) = 1.$$

4. Let A & B be any two events such that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Atleast

↳ is called Law of addition of probability.

$A \cup B$ - A union B. $A \cap B$ - A intersection B.

②

$$P(A \cap B) = P(A) \times P(B)$$

Both. ↳ Law of multiplication of probability.

② wedding multiplication theorem.

invite

∩ - family / invite

Me Wife → A ∩ B

and maths × multiply

$\cap \rightarrow$ and $-$ $\times$ (multiply)

$\cup \rightarrow$ OR $-$ $+$ (Addition)

Page No.:
Date:

$\cup$

$\cap$

Family of 10

call death

A or B or both

Addition theorum

$A \cup B - P(A) + P(B) - P(A \cap B)$

$x \in A,\ x \in B,\ x \in AB$

① $A/B \rightarrow$ A given B $\frac{1^{st} \text{ sem exam}}{\text{Joined MBA}}$

B is achieved

A don't know if it will be achieved.

<u>Conditional Probability.</u>

Let A and B be any two events such that

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

(or)

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

P(A/B) denotes probability for the happening of event A when event B is already happened.

P(B/A) denotes probability for the happening of event B when event A is already happened.

① B is already over, what is the probability of happening of A.

10/10/19
Thursday

Mutually Exclusive Events
Collectively Exhaustive Events.
Dependent and Independent Events.

① The thing that is currently happening will stop other things.
eg:- Staff. The happening of Iyyanar sir will stop the happening of others.

* The happening of one event which stops the happening of all other events, then the events are said to be Mutually exclusive events.

② Any number of event can be taken. all the event will have probability. the total of the probabilities of the events should be 1.

* If the total of probability of all the events is equal to 1, then the events are said to be collectively Exhaustive events.

Coin H, T. $P(H) = \frac{1}{2}$ $P(T) = \frac{1}{2}$

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

Auto - 1000
Milk - 1000
Veg - 2000
Fruit - 6000
10000

E_1 E_2 E_3 E_4

$\frac{1000}{10000} + \frac{1000}{10000} + \frac{2000}{10000} + \frac{6000}{10000}$

$\frac{10000}{10000} = 1.$

① Subset / contains in
$\forall$ - for all

Page No.:
Date: / /

* If the happening of one event which depends on the happening of another event, then the events are said to be dependent events.

conditional / dependent.

③ eg: UG $\longrightarrow$ MBA

Teacher $\longrightarrow$ Students.

* If the happening of one event which does not depend upon another event is said to be independent event.

eg: Teacher not dependent on another Teacher to take class.

Baye's Theoream :-

Let $E_1, E_2, \ldots\ldots E_n$ be 'n' mutually exclusively and collectively Exhaustive events such that $P(E_i) \neq 0 \; \forall \; i = 1, 2, \ldots\ldots n$. & $\Sigma P(E_i) = 1$, then there exists an arbitrary event A such that $P(A) \neq 0$ and $A \subset \bigcup_{i=1}^{n} E_i$ then

$$P\left(E_i/A\right) = \frac{P(A/E_i) \times P(E_i)}{\sum_{i=1}^{n} P(A/E_i) \times P(E_i)}.$$

Page No.:
Date: / /

15/10/19
Tuesday
Pg.2)

(Q 10) BAYE'S THEOREUM.

Bolt

	A	B	C	
	25%	35%	40%	⇒ 100%
Defective bolts -	5%.	4 %.	2 %	⇒ 11%

within 100%

eg:- water bottle

Baye's theoreum is used to find the new information from the ~~alternative~~ already available information.

already available info.	New information
our personal details	whether we will pass in the 1st sem.

Pg. 2)

10) E_1 = Bolt manufactured by MA
E_2 = " " " MB
E_3 = " " " MC
A = Defective bolts.

E_i	$P(E_i)$	A	$P(A/E_i)$	$P(A/E_i) \times P(E_i)$
E_1	$P(E_1) = 0.25$	D	$P(A/E_1) = 0.05$	$P(A/E_1) \times P(E_1) =$
		E		0.25×0.05
		F		$= 0.0125$
		E		
E_2	$P(E_2) = 0.35$	C	$P(A/E_2) = 0.04$	$P(A/E_2) \times P(E_2)$
		T		$= 0.35 \times 0.04$
		I		$= 0.0140$
		V		
E_3	$P(E_3) = 0.40$	E	$P(A/E_3) = 0.02$	$P(A/E_3) \times P(E_3)$
		B		$= 0.40 \times 0.02$
		O		$= 0.0080$
		L		
		T		
		S		
				0.0345

By using Baye's Theorem :

$$P(E_i/A) = \frac{P(A/E_i) \times P(E_i)}{\sum_{i=1}^{3} P(A/E_i) \times P(E_i)}$$

$$P(E_1/A) = \frac{P(A/E_1) \times P(E_1)}{P(A/E_1) \times P(E_1) + P(A/E_2) \times P(E_2) + P(A/E_3) \times P(E_3)}$$

$$= \frac{0.0125}{0.0345}$$

$$= \underline{\underline{0.3623}}$$

$$P(E_2/A) = \frac{P(A/E_2) \times P(E_2)}{P(A/E_1) \times P(E_1) + P(A/E_2) \times P(E_2) + P(A/E_3) \times P(E_3)}$$

$$= \frac{0.0140}{0.0345}$$

$$= 0.4058$$

$$P(E_3/A) = \frac{P(A/E_3) \times P(E_3)}{P(A/E_1) \times P(E_1) + P(A/E_2) \times P(E_2) + P(A/E_3) \times P(E_3)}$$

$$= \frac{0.0080}{0.0345}$$

$$= \underline{\underline{0.2319}}$$

<u>Baye's Theorem</u>

Let $E_1, E_2, \ldots E_n$ be 'n' mutually exclusive and collectively exhaustive events where $P(E_i) \neq 0 \; \forall \; i = 1, 2, \ldots n$ & $\sum P(E_i) = 1$, then there exists an arbitrary event A where $P(A) \neq 0$ & $A \subset \bigcup_{i=1}^{n} E_i$, then

$$P(E_i/A) = \frac{P(A/E_i) \times P(E_i)}{\sum_{i=1}^{n} P(A/E_i) \times P(E_i)}.$$

where, P = Probability

E_i = Events [mutually exclusive & collectively exhaustive]

A = Defective event / arbitrary event

Page No.:
Date: / /

Pg. 21)
20)

E_i	$P(E_i)$	A	$P(A/E_i)$	$P(A/E_i) \times P(E_i)$
E_1	$P(E_1) = 0.30$	D E	$P(A/E_1) = 0.05$	$P(A/E_1) \times P(E_1) = 0.30 \times 0.05 = 0.0150$
E_2	$P(E_2) = 0.70$	F E C T I V E	$P(A/E_2) = 0.01$	$P(A/E_2) \times P(E_2) = 0.70 \times 0.01 = 0.0070$
				0.0220

$$P(E_i/A) = \frac{P(A/E_i) \times P(E_i)}{\sum_{i=1}^{n} P(A/E_i) \times P(E_i)}$$

$P(E_1/A)$

$$= \frac{0.0150}{0.0220}$$

$= 0.6818$

$P(E_2/A)$

$$= \frac{0.0070}{0.0220}$$

$= 0.3182$

89

Page No.:
Date: / /

21/10/19
Monday
Pg. 28
Anniversary QP

11)(a) E_1 = Let E_1 be the intial repair done by Ravi.
E_2 = Initial repair done by Tarun
E_3 = Initial repair done by Gautam.
E_4 = Initial repair done by Prasad.
A = Incomplete repair.

E_i	$P(E_i)$	A	$P(A/E_i)$	$P(A/E_i) \times P(E_i)$
E_1	$P(E_1) = 0.20$	Incomplete Repair	$P(A/E_1) = 0.05$	$0.05 \times 0.20 = 0.0100$
E_2	$P(E_2) = 0.60$		$P(A/E_2) = 0.10$	$0.10 \times 0.60 = 0.0600$
E_3	$P(E_3) = 0.15$		$P(A/E_3) = 0.10$	$0.10 \times 0.15 = 0.0150$
E_4	$P(E_4) = 0.05$		$P(A/E_4) = 0.05$	$0.05 \times 0.05 = 0.0025$
				~~0.0850~~
				0.0875

$$P(E_i/A) = \frac{P(A/E_i) \times P(E_i)}{\sum_{i=1}^{4} P(A/E_i) \times P(E_i)}$$

$P(E_1/A) = 0.1143$
$P(E_2/A) = 0.6857$
$P(E_3/A) = 0.1714$
$P(E_4/A) = 0.0286$

Pg. J1102. Q14(b) chi square test

Q15(a) Regression.

Q16(a) Regression, 1st y on x. - 50 crs x.

Pg. BS2102 Q15(a) 2nd x on y - 10 lakhs y,

Random Variable

X - 0, 1, 2, 3, ... - Discrete random variable.

X - 0 to 5, 5 to 10, -15 to 20, 30 to 45 - Continous R.V.

Distributions - Pattern.

2019 Jeans

2019 → 1990's Jeans

2090's Ink Pen ball

1970's ball point ✗

n → no. of independent trials. tossing

p → prob. of success [max value 1] [least 0]

q → Prob. of failure.

x → no. of success.

Bernoulli	Binomial	Poisson
n = 1	n = finite	n = Infinite ∞
p = 0.5	p → 1 (tends to 1)	p → 0 (tends to 0)
q = 0.5	q → 0 (tends to 0)	q → 1 (tends to 1)
		λ - indirect ⇓ Mean $\bar{x}$

X

Q) ~~Ten~~ Five coins are tossed simultaneously. What is the probability of getting:
1) Exactly 2 heads
2) Atleast 4 heads
3) Atmost 1 head.

X > 30 - Binomial
X > 20 - Poisson

X - No. of heads.
X - 0, 1, 2, 3, 4, 5 - finite
Discrete R.V.
Binomial Distribution.

Exactly =
Starting Atleast $\geq$
Maximum Atmost $\leq$
lies btw $\leq x \leq$
less than $<$
greater than $>$

Discrete

Sol:
1) Exactly 2 heads $P[X = 2]$
2) Atleast 4 heads $P[X \geq 4]$
3) Atmost 1 heads $P[X \leq 1]$

$X = \underbrace{0, 1, 2, 3, 4, 5}_{\text{success}}$ n

1) ~~We know that~~ $P[X = 2]$

we know that by using binomial distributions

$$P[X = x] = nC_x \, p^x \, q^{n-x}$$

$q = 1-p$

$$P[X = 2] = 5C_2 \left(1/2\right)^2 \left(1/2\right)^{5-2}$$

$$= 5C_2 \left(1/2\right)^2 \left(1/2\right)^3$$

$x = 2$
$n = 5$
$p = 1/2$
$q = 1 - 1/2 = 1/2$.

(1) included.

$nC_0 = 1$

$nC_n = 1$

$nC_1 = n$

Page No.: Date: / /

$5C_2 = 10.$

$= 10 \times \frac{1^2}{2^2} \times \frac{1^3}{2^3}$

$= 10 \times \frac{1}{4} \times \frac{1}{8}$

$= \cancel{\frac{10}{32}}$ $5/16$ //.

$x = 4 + x = 5$

(2) $P[X \geq 4] = (P[X=4] + P[X=5])$

$= 5C_4 (1/2)^4 (1/2)^1 + 5C_5 (1/2)^5 (1/2)^0$

$= 5/32 + 1/32 = 3/16.$

(3) $P[X \leq 1] = P[X=0] + P[X=1]$

$= 5C_0 (1/2)^0 (1/2)^5 + 5C_1 (1/2)^1 (1/2)^4$

$= 3/16$

Anna university Pg. 15) 11(b)(i)

2 weeks $p = 0.45$ $q = 0.55$

1 week $p = 0.10$ $q = 0.90$

$\geq$ 3 weeks $p = 0.20.$ $q = 0.80.$

$X = 0, 1, 2, 3, 4, \ldots 20$

$n = 20.$

Page No. :
Date : / /

(i) $P[X=8]$ $n=20$ $p=0.45$ $q=0.55$
$x=8$.

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

$= 20C_8 (0.45)^8 (0.55)^{20-8}$

$= \underline{0.1623}$

(ii) $P[X=1]$ $n=20$ $p=0.10$ $q=0.90$
$x=1$

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

$= 20C_1 (0.10)^1 (0.90)^{20-1}$

$= \underline{0.2702}$

(iii) $P[X\leq 2]$ $n=20$ $p=0.20$ $q=0.80$
$x\leq 2$.

$= P[X=0] + P[X=1] + P[X=2]$

$= 20C_0 (0.20)^0 (0.80)^{20} + 20C_1 (0.20)^1 (0.80)^{19} + 20C_2 (0.20)^2 (0.80)^{18}$

$= \underline{0.2061}$

(iv) $P[X\geq 2]$ $n=20$ $p=0.10$ $q=0.90$

$P[X\geq a] = 1-P[X<a]$

$P[X>a] = 1-P[X\leq a]$

$P[X\leq a] = 1-P[X>a]$

$P[X<a] = 1-P[X\geq a]$

$= P[X\geq 2] = 1-P[X<2]$

$= 1-P[X=0]+P[X=1]$

① Average or mean or expected value or expected
② Factorial

94

Indirect, λ, $n \times p$

$= 1 - {}^{20}C_0 (0.10)^0 (0.90)^{20} + {}^{20}C_1 (0.10)^1 (0.90)^{19}$

$= 1 - 0.3917$

$= 0.6083$

22/10/19
Tuesday

Poisson Distribution.

$X \rightarrow$ Discrete

$X = 0, 1, 2, 3, \ldots \infty$

$\lambda \rightarrow$ Parameter of P.D.

Direct $\rightarrow$ [Mean of P.D. λ / Variance of P.D λ / Standard deviation of P.D $\sqrt{\lambda}$, $(\sqrt{\lambda})^2 = \lambda$

$$P[X = x] = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

$\lambda = n \times p$

Pg. 19)
11)(a)(i)

$n = 30$ $\quad q = 0.96$.

$p = 0.04$

$\lambda = n \times p$

$= 30 \times 0.04$

$= 1.20$

$P[X = 25]$

(1) $P[X = x] = \dfrac{e^{-\lambda} \cdot \lambda^x}{x!}$

$$= \frac{e^{-1.20} \times (1.20)^{25}}{25!} = 0$$

0.0355
+0.2020

Page No.:
Date: / /

(2) $P[X=3] = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$

$= \frac{e^{-1.20} \times 1.20^3}{3!}$

$= \underline{0.0867}$

24/10/19 Thursday

$X = 0, 1, 2, 3, \ldots 20.$

11) (b) (i) $P[X = x] = nC_x\, p^x q^{n-x}$ $n=20$ $q=0.70$

(Pg. 5) $p = 0.30$

$P[X \leq 4] = P[X=0] + P[X=1]\, P[X=2] + P[X=3],\ P[X=4].$

$= 20C_0 \times (0.30)^0 (0.70)^{20} + 20C_1 (0.30)^1 (0.70)^{19} + 20C_2 (0.30)^2 (0.70)^{18} + 20C_3 (0.30)^3 (0.70)^{17} + 20C_4 (0.30)^4 (0.70)^{16}$

$= \underline{0.2375}$

(ii) Atleast to 3 $[X \geq 3]$.

$X = 0, 1, 2, 3, \ldots 20.$

$P[X \geq 3]$.

$P[X \geq a] = 1 - P[X < a]$

$P[X \geq 3] = 1 - [P[X=0] + P[X=1] + P[X=2]]$

$= 1 - [20C_0 (0.30)^0 (0.70)^{20} + 20C_1 (0.30)^1 (0.70)^{19} + 20C_2 (0.30)^2 (0.70)^{18}]$.

$= 1 - [0.0008 + 0.0068 + 0.0278]$

$= 1 - 0.0354$

$= \underline{0.9646}$

Binomial Discrete Poisson | Normal continuous variable

If any problem related to time, this problem comes under poisson distribution

Pg. 8) 11(b) $\lambda = 6$ people per hr. [P.D] $\frac{e^{-\lambda} \cdot \lambda^x}{x!}$

(i) $P[X=6] = \frac{e^{-6} \cdot 6^6}{6!} = \underline{\underline{0.1606}}$

(ii) $P[X<5] = P[X=0] + P[X=1] + P[X=2] + P[X=3] + P[X=4]$.

$= \frac{e^{-6} \cdot 6^0}{0!} + \frac{e^{-6} \cdot 6^1}{1!} + \frac{e^{-6} \cdot 6^2}{2!} + \frac{e^{-6} \cdot 6^3}{3!} + \frac{e^{-6} \cdot 6^4}{4!}$

$= 0.0025 + 0.0149 + 0.0446 + 0.0892 + 0.1339$.

$= \underline{\underline{0.2851}}$

(iii) $\underline{\underline{\lambda = 1}}$ [10 min = 1 person] $= \cancel{6} \times \frac{1\cancel{0}}{\cancel{6}\cancel{0}}$

$P[X=0] = \frac{e^{-1} \cdot 1^0}{0!} = \underline{\underline{0.3679}}$

(iv) $\lambda = 0.5$ [5 min = 1 person] $= 6 \times \frac{5}{60} = 0.5$

$P[X=0] = \frac{e^{-0.5} \cdot 0.5^0}{0!} = \underline{\underline{0.6065}}$

(1) Population mean
(2) Population S.D
(3) Random variable

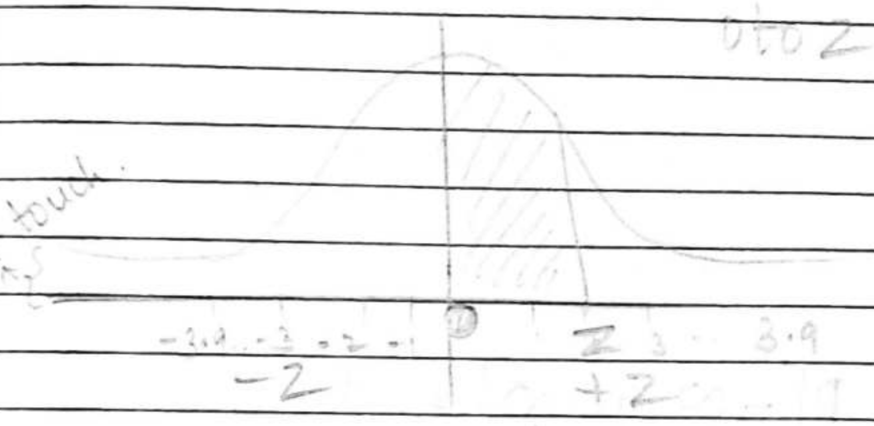
1:00

25/10/19
Friday

(1) $Z = \dfrac{X - \mu}{\sigma}$ is called standard Normal variate.

Standard Normal Table :→

0 to Z



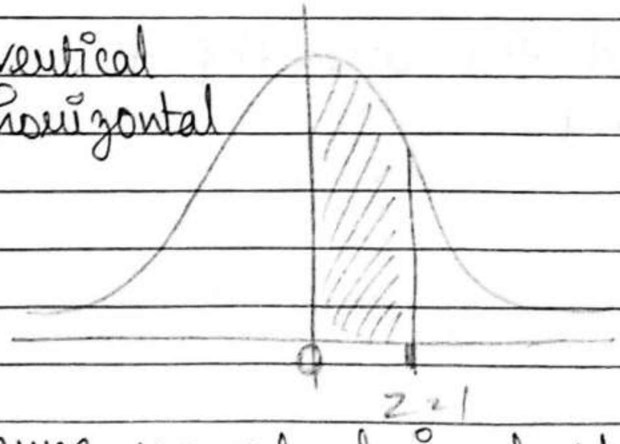
2nd decimal

Z	0.00	0.01	0.02	----
0.0				
0.1				
0.2				
0.3				
⋮				
1.0	0.3413			
⋮				
3.9				

1st decimal

Q) Find the area of $Z = 1$.

1st decimal - vertical
2nd decimal - horizontal

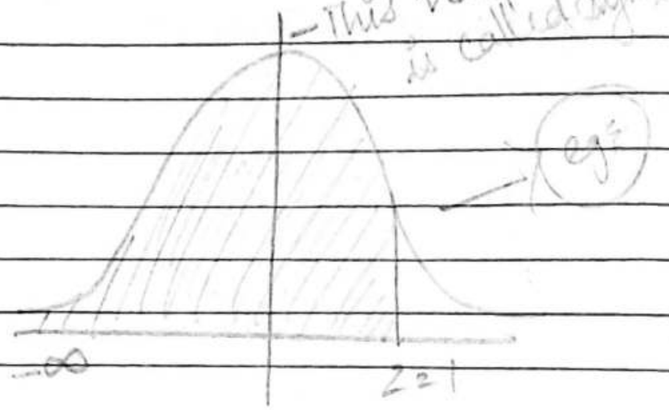


Maximum no. of decimal places for Z is 2.

① Atmost, Not more than, upto that
② Atleast, Not less than.

$P[Z=1] = P[0 \text{ to } 1.00]$
$= \underline{0.3413}$

Q) Find the area to the left of $Z=1$.



$P[-\infty \text{ to } \infty] = 1$

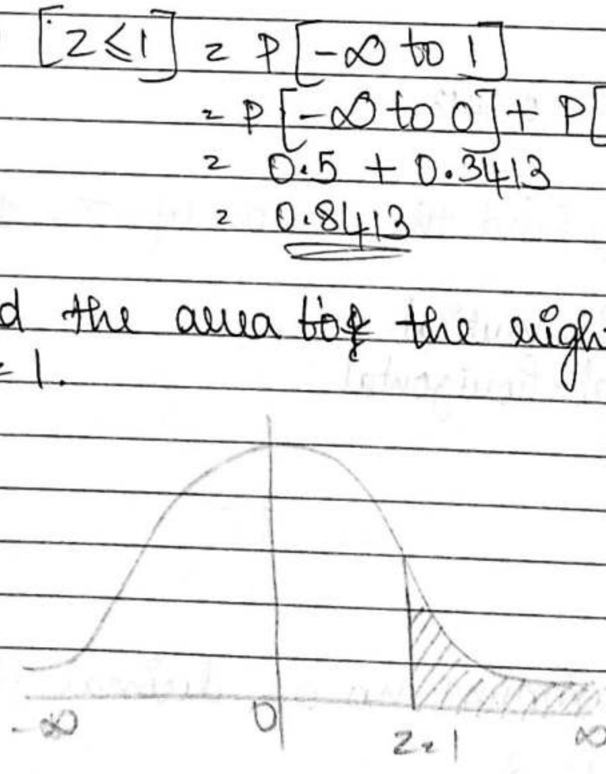
$P[-\infty \text{ to } 0] = 0.5$

$P[0 \text{ to } \infty] = 0.5$

left = $<$ ① Left of $Z=1$

Right = $>$ ② $P[Z \leq 1] = P[-\infty \text{ to } 1]$
$= P[-\infty \text{ to } 0] + P[0 \text{ to } 1]$
$= 0.5 + 0.3413$
$= \underline{0.8413}$

Q) Find the area ~~tof~~ to the right of $Z=1$.



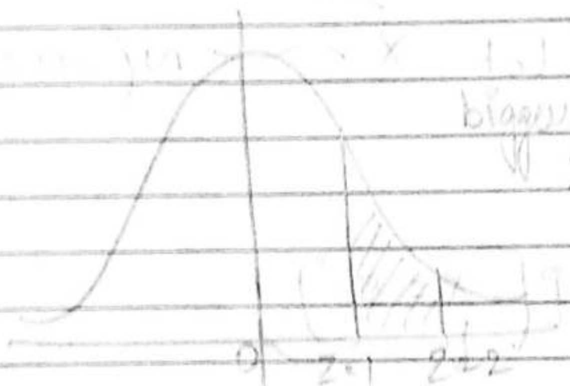
$P[z \geq 1] = P[1 \text{ to } \infty]$

$= 0.5 - \text{Table value of } z$

$= 0.5 - 0.3413$

$= \underline{0.1587}$

Q) Find the area



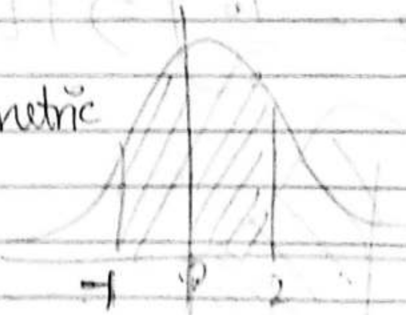
$= P[z = 2] - P[z = 1]$

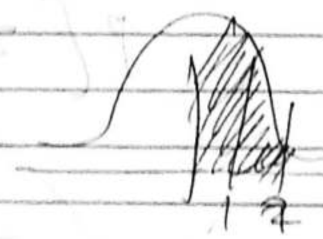
$= 0.4772 - 0.3413$

$= \underline{0.1359}$

Q) Find the area of $[-1 \leq z \leq 2]$ lies between -1 to 2.

$\because$ symmetric





$P[-1 \leq z \leq 2] = P[-1 \text{ to } 0] + P[0 \text{ to } 2]$

$= P[0 \text{ to } 1] + P[0 \text{ to } 2]$

Variance - σ^2 — Normal distribut

S.D - σ.

$= 0.3413 + 0.4772$

$= \underline{0.8185}$

Spoilt

	A	B
Ban	3%	5%
Jack	5%	7%
	E_1	E_2

E_1 E_2

Mathematical Aptitude

29/10/19
Tuesday

Pg. 2)
11(b).

Let $X \sim N(\mu = 70, \sigma = 5]$

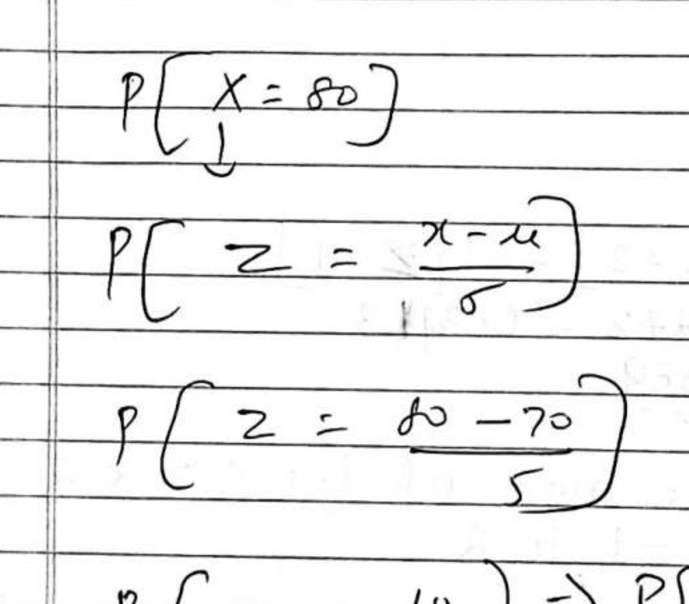
1) $P[X = 80]$

$P\left[Z = \frac{x-\mu}{\sigma}\right]$

$P\left[Z = \frac{80-70}{5}\right]$

$P\left[Z = \frac{10}{5}\right] \Rightarrow P[Z = 2]$

$= \underline{0.4772}$



Class of students

How many - Ans x no of students.

% - Ans x 100.

P - Ans only

-3.5 to 3.5

$70 \le z \le 86$

101 $\quad p = 0.80$

(1)

$P[X = 60]$

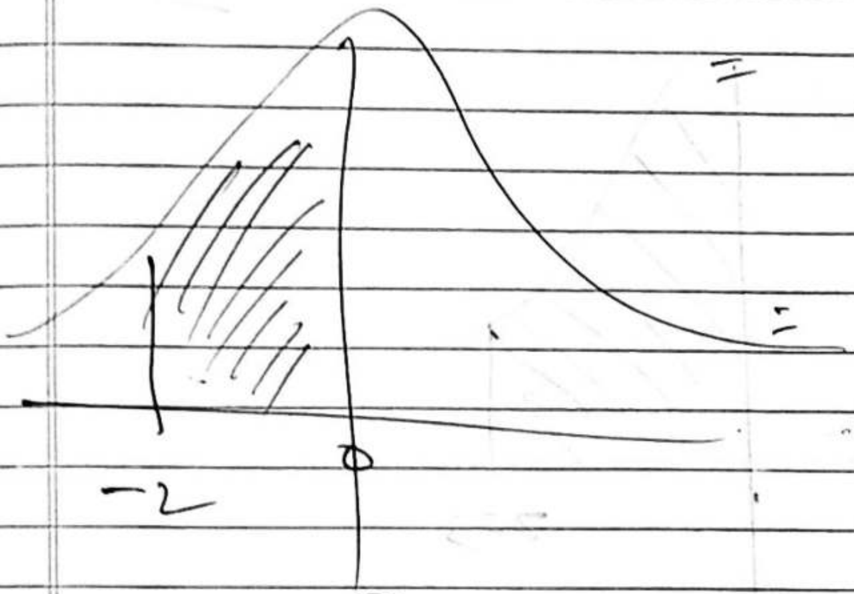
$$P\left[z = \frac{60 - 70}{5}\right]$$

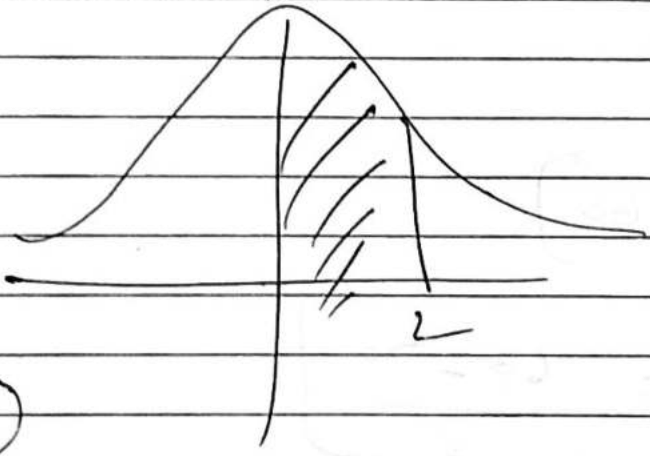
$$P[z = -2] = P[\cancel{-\infty \text{ to } -2}] \quad -2 \text{ to } 0$$

$$= P[0 \text{ to } 2]$$

$\because$ Symmetric

$$= 0.4772$$





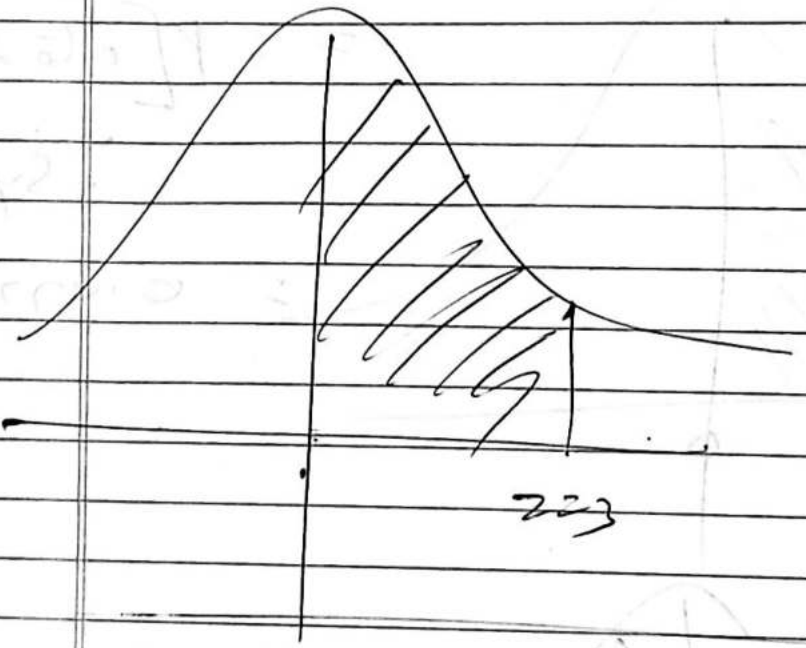
(5)

$\overline{100} \quad \overline{5}$

$$P[70 \le X \le 85]$$

$$P\left[\frac{70-70}{5} \le Z \le \frac{85-70}{5}\right]$$

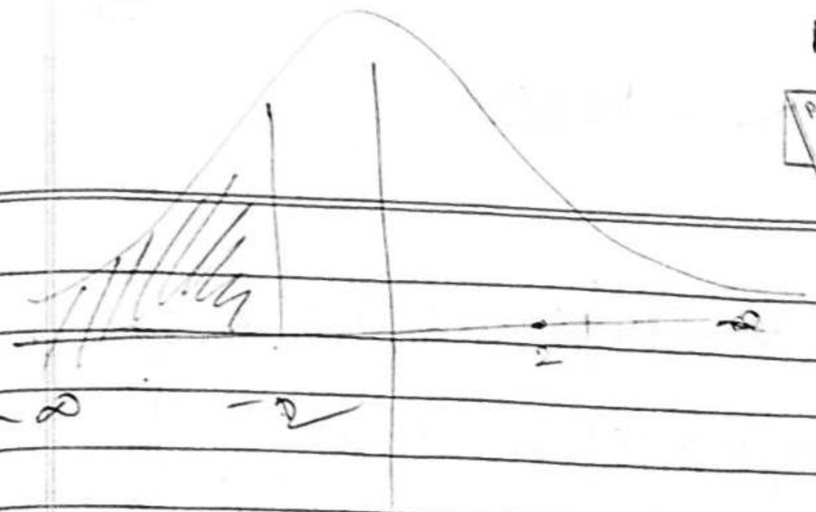
$$P[0 \le Z \le 3] = 0.49875$$



$$P[X \le 60]$$

$$P\left[Z \le \frac{60-70}{5}\right]$$

$$P[Z \le -2]$$



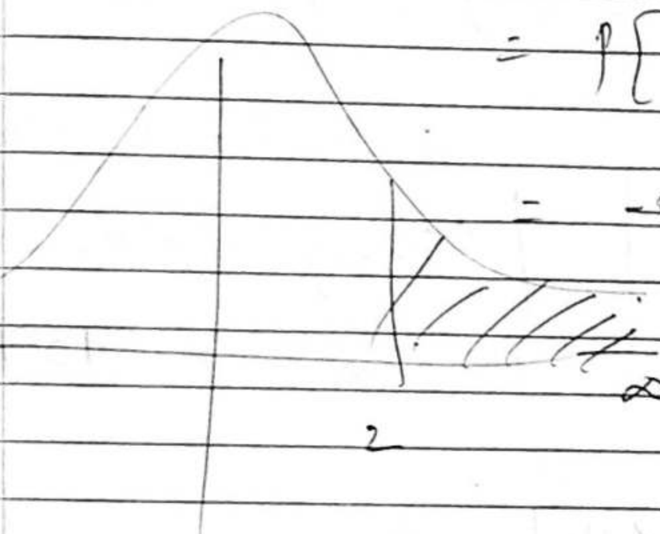
10-2=8

10-1=9

$P[z \le -2] = P[-\infty \text{ to } -2]$

$= P[2 \text{ to } \infty]$ ∵ symmetry

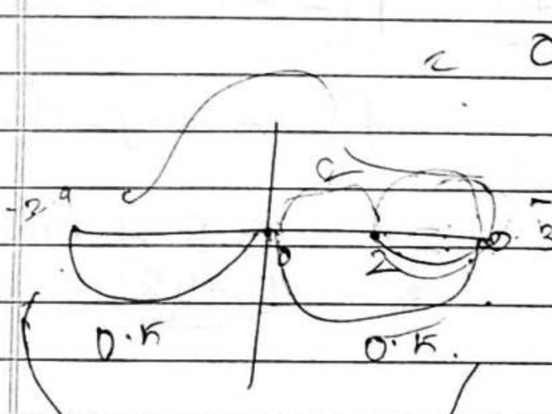
$= \cancel{0.5} - P[0 \text{ to } 2]$



$= P[0 \text{ to } \infty) - P[0 \text{ to } 2]$

$= 0.5 - 0.4772$

$= 0.0228$



$0.5 - (P$ [illegible]

$0.5 - 0.4772$

$= 1.$

$\sim$ - Follows

Pg 4)

11(a) $X \sim N(\mu = 12 \ \& \ \sigma = 3)$

(i) $P[X \geq 15] = P\left[\cancel{Z = \frac{x-\mu}{\sigma}}\right]$

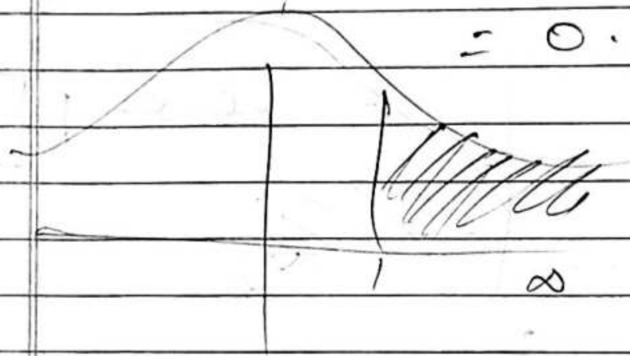
$P\left[Z \geq \frac{15-12}{3}\right]$

$P[Z \geq 1] = P[1 \text{ to } \infty]$

$= 0.5 - P[0 \text{ to } 1]$

$= 0.5 - 0.3413$

$= \underline{0.1587}$



(ii) $P[X \leq 14]$

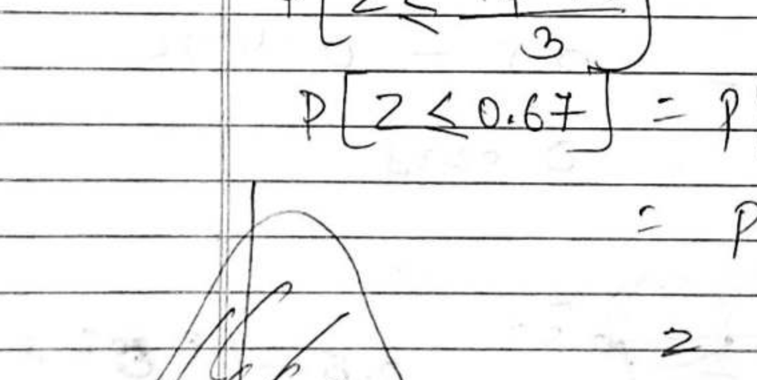
$P\left[Z \leq \frac{14-12}{3}\right]$

$P[Z \leq 0.67] = P[-\infty \text{ to } 0.67]$

$= P[-\infty \text{ to } 0] + P[0 \text{ to } 0$

$= 0.5 + 0.2486.$

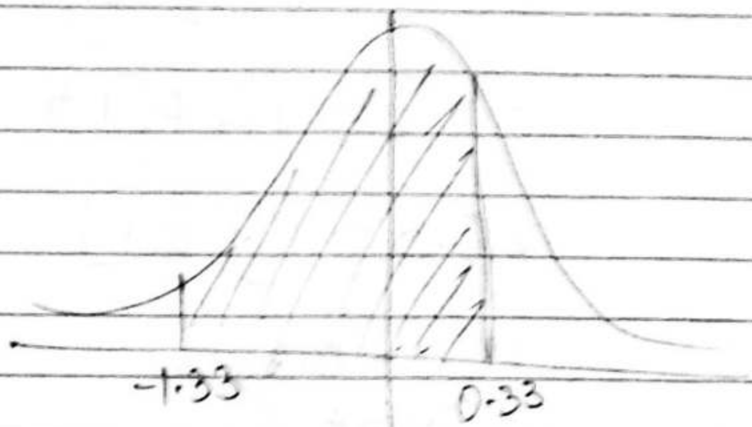
$= \underline{0.7486}$



(iii) $P[8 \leq X \leq 13]$

$P\left[\frac{8-12}{3} \leq Z \leq \frac{13-12}{3}\right]$

$P[-1.33 \leq Z \leq 0.33] =$



$P[-1.33 \text{ to } 0] + P[0 \text{ to } 0.33]$.

$\because$ symmetric $P[0 \text{ to } 1.33] + P[0 \text{ to } 0.33]$.

$= 0.4082 + 0.1293$

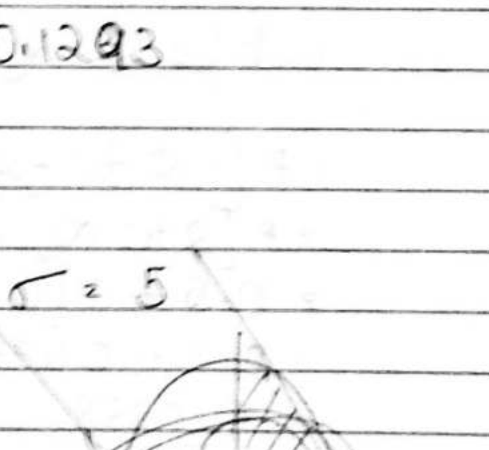
$= \underline{0.5375}$

Pg 4)

11(b) $X \rightarrow \mu = 65 \quad \sigma = 5$

(i) $P[X \geq 70]$

$P\left[Z \geq \frac{70-65}{5}\right]$



$= P[Z \geq 1] = P[$

Page No. Date:

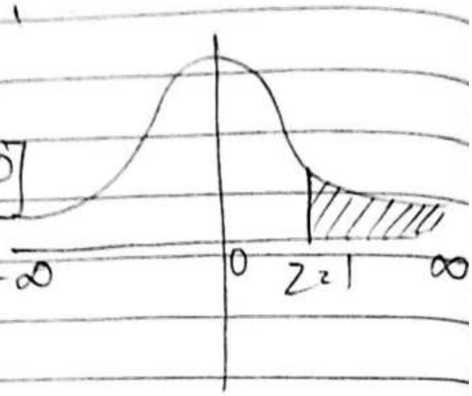
24/10/19 Tuesday Pg. 11)

11 (b) $X \sim N(\mu = 65 \;\&\; \sigma = 5)$

(i) $P[X \geq 70]$.

$= P\left[Z \geq \frac{70-65}{5}\right]$

$= P[Z \geq 1] = P[1 \text{ to } \infty]$



$= 0.5 - P[0 \text{ to } 1]$

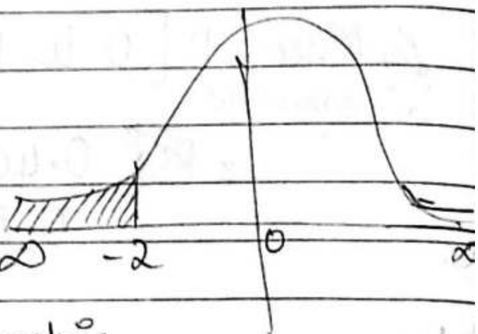
$= 0.5 - 0.3413$

$= \underline{0.1587}$

(ii) $P[X \leq 55 \text{ Kg}]$

$= P\left[Z \leq \frac{55-65}{5}\right]$

$= P[Z \leq -2] =$



$= P[-\infty \text{ to } -2] =$

$= P[2 \text{ to } \infty]$ ~~since~~ ∵ symmetric.

$= 0.5 - P[0 \text{ to } 2]$.

$= 0.5 - 0.4772$

$= \underline{0.0228}$

(iii) $P[60 \leq X \leq 75]$.

$= P\left[\frac{60-65}{5} \leq Z \leq \frac{75-65}{5}\right]$

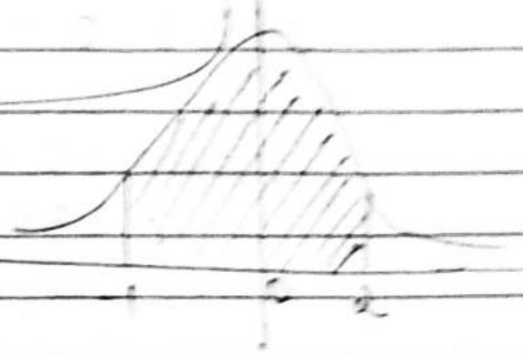
$= P[-1 \leq Z \leq 2]$.

$= P[-1 \text{ to } 0] + P[0 \text{ to } 2]$

$\therefore$ Symmetric $= P[0 \text{ to } 1] + P[0 \text{ to } 2]$

$= 0.3413 + 0.4772$

$= 0.8185$



0.1587×50

$= \underline{7.9350}$

Q.35) Z test for diff btw two means

3) (a) (i) $n_1 = 40$, $\bar{x}_1 = 647$ hrs, $\sigma_1 = 27$ hrs

competitor (ii) $n_2 = 40$, $\bar{x}_2 = 638$ hrs, $\sigma_2 = 31$ hrs.

(i) $\bar{x}_1 = \dfrac{\Sigma x_1}{n_1} = \dfrac{647}{\cancel{40}} = \cancel{16.18}$ 647

(ii) $\bar{x}_2 = \cancel{\dfrac{\Sigma x_2}{n_2} = \dfrac{638}{40}} = \cancel{15.95}$ 638

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

$\alpha: 0.05$

$$Z_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

$$= \frac{\overset{647}{\cancel{16.18}} - \overset{638}{\cancel{15.95}}}{\sqrt{\dfrac{(27)^2}{40} + \dfrac{(31)^2}{40}}}$$

$$= \frac{\cancel{0.23}\ 9}{\sqrt{18.23 + 24.03}}$$

$= \dfrac{\cancel{0.239}}{\cancel{2.4}\ 6.50}$

$= \cancel{0.095}\ 1.38$

Ztab : 0.05 (5%) = 1.96.

Inference : Since Zcal < Z-tab accept H_0
1.38 ~~0.21~~ ~~0.095~~ 1.96

i.e., $\mu_1 = \mu_2$. This substantiates the claim at 0.05 level of significance.

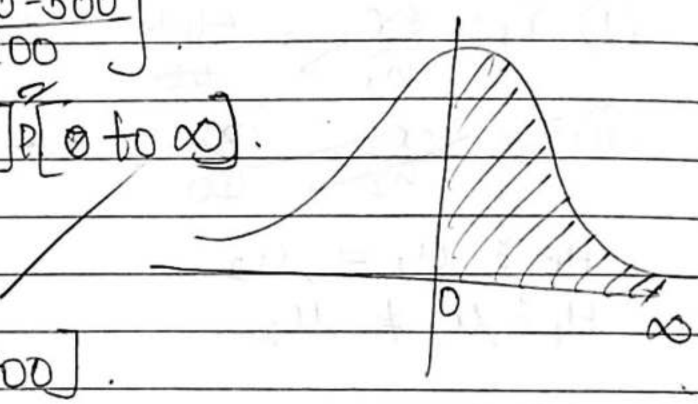
Pg. 28)
Q. 11(b) $X \rightarrow N(\mu = 500$ & $\sigma = 100)$.

(i) $P[X \geq 500]$.

$= P\left[Z \geq \dfrac{500-500}{100}\right]$.

$= P[Z \geq 0]$ or $P[0 \text{ to } \infty]$.

$= 0.5$



(ii) $P[X \geq 700]$.

$= P\left[Z \geq \dfrac{700-500}{100}\right]$

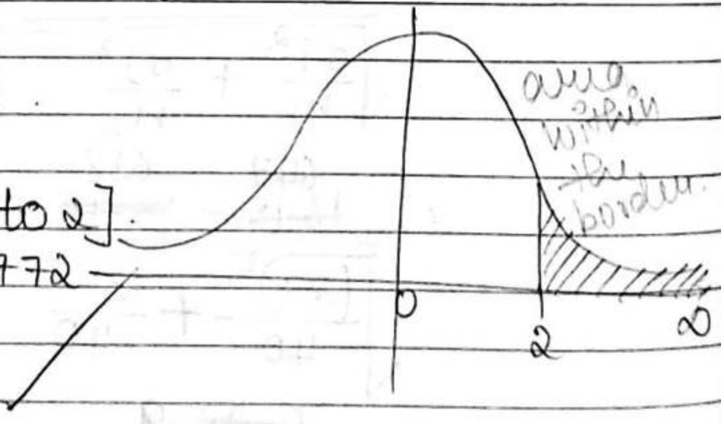
$= P[Z \geq 2]$.

$= P[2 \text{ to } \infty]$

$= 0.5 - P[0 \text{ to } 2]$.

$= 0.5 - 0.4772$

$= 0.0228$



(iii) $P\left[550 \leq X \leq 650\right]$ Bigger - Smaller

$= P\left[\frac{550-500}{100} \leq Z \leq \frac{650-500}{100}\right]$

$= P\left[0.50 \leq Z \leq 1.5\right]$

~~$= P[0 \text{ to } 0.50] + P[0 \text{ to } 1.50]$~~

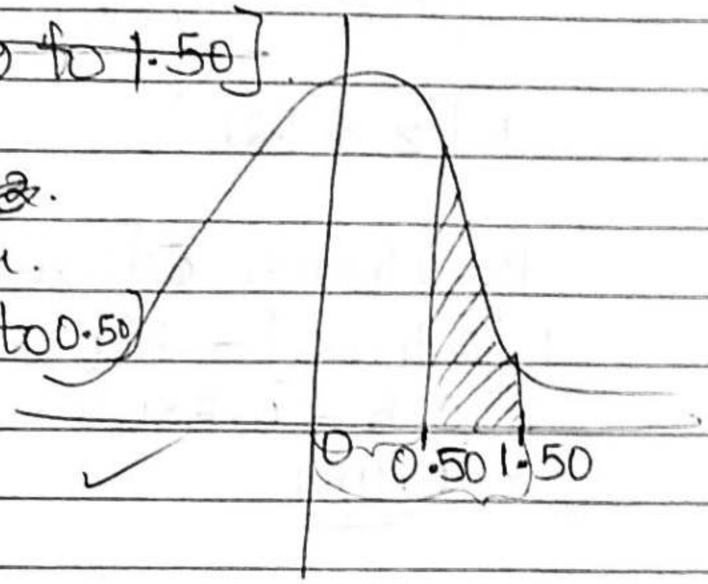
~~$= 0.1915 + 0.4332$~~

= Bigger - smaller.

$= P[0 \text{ to } 1.50] - P[0 \text{ to } 0.50]$

$= 0.4332 - 0.1915$

$= \underline{0.2417}$



(iv) $P\left[X < 580\right]$

$P\left[Z \leq \frac{580-500}{100}\right]$

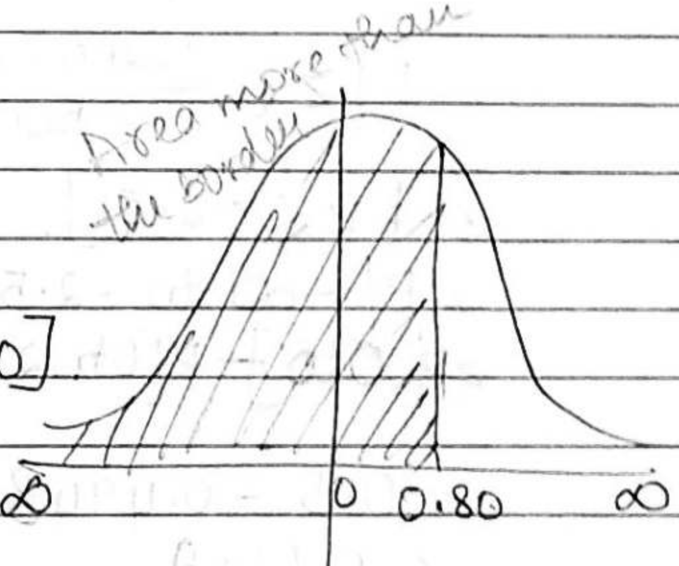
$= P\left[Z \leq 0.80\right]$

$= P[-\infty \text{ to } 0.80]$

$= P[-\infty \text{ to } 0] + P[0 \text{ to } 0.80]$

$= 0.50 + 0.2881$

$= \underline{0.7881}$



Pg. 35)

11(b) $X \sim N(\mu = 25000 \; \& \; \sigma = 5000)$

(i) $P[X \geq 31000]$

$P\left[Z \geq \frac{31000 - 25000}{5000}\right]$

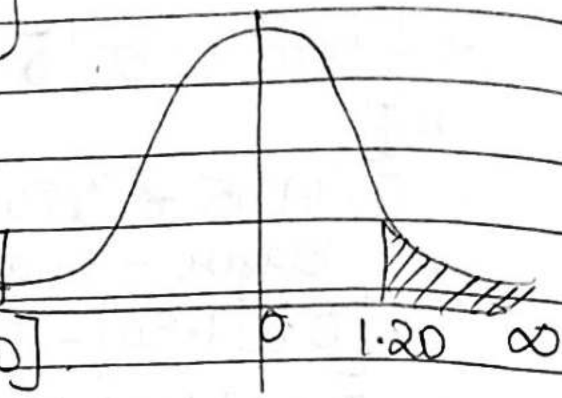
$P[Z \geq 1.20]$

$\cancel{P[0 \text{ to } 1.20] \; P[1.20 \text{ to } \infty]}$

$P[0 \text{ to } \infty] - P[0 \text{ to } 1.20]$

$= 0.5 - 0.3849$

$= \underline{0.1151}$



(ii) $P[X \leq 12200]$

$P\left[Z \leq \frac{12200 - 25000}{5000}\right]$

$= P[Z \leq -2.56]$

$= P[-\infty \text{ to } -2.56]$

$= P[0 \text{ to } \infty] - P[0 \text{ to } 2.56]$ since symmetric

$= 0.5 - 0.4948$

$= \underline{0.0052}$

Random chance of occurance — Probability

Fit — expected frequency

111

Continuous

Atleast $\geq$

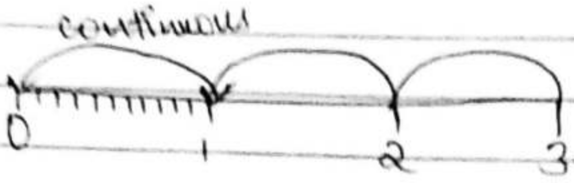
Atmost $\leq$

Lessthan $<$

greaterthan $>$

Like btw $\leq X \leq$

exactly equal to $=$



2/6/10
Thursday

Fitting of Poisson Distribution means calculating the expected frequency.

$$E[X] = N \times P[X = x]$$

N = total frequency.

P – Probability mass function of P.D

$$\bar{X} = \frac{\Sigma FX}{\Sigma F} \text{ or } \frac{\Sigma FX}{N}$$

X	F	FX
0	123	0
1	59	59
2	14	28
3	3	9
4	1	4
	200	100

N = 200

$$\bar{X} = \frac{100}{200} = 0.5 = \lambda$$

$$\boxed{\text{In P.D mean} = \lambda = 0.5.}$$

(i) $P[X=0] = \dfrac{e^{-\lambda} \cdot \lambda^{x}}{x!}$

$= \dfrac{e^{-0.5} \cdot 0.5^{0}}{0!}$

$= \underline{\underline{0.6065}}$

(ii) $P[X=1] = \dfrac{e^{-\lambda} \cdot \lambda^{x}}{x!}$

$= \dfrac{e^{-0.5} \cdot 0.5^{1}}{1!}$

$= 0.3033$

(iii) $P[X=2] = \dfrac{e^{-0.5} \cdot 0.5^{2}}{2!}$

$= \underline{\underline{0.0758}}$

(iv) $P[X=3] = \dfrac{e^{-0.5} \cdot 0.5^{3}}{3!}$

$= \underline{\underline{0.0126}}$

(v) $P[X=4] = \dfrac{e^{-0.5} \cdot 0.5^{4}}{4!}$

$= \underline{\underline{0.0016}}$

$E[X] = N \times P[X=x]$

<u>Fit</u>

$E[X] = N \times P[X=x]$.

$E[0] = 200 \times 0.6065$

$= \cancel{121.300}$ 121

$E[1] = 200 \times 0.3033$

$= \underline{\underline{60.66}}$ 61

(1) Observed frequency
(2) Expected frequency

113

$E[2] = 200 \times 0.0758$
$= 15.16$ — 15

$E[3] = 200 \times 0.0126$
$= 2.520$ — 3

$E[4] = 200 \times 0.0016$
$= 0.32$ — 0

200

Ei

(1) Observed frequency
(2) Expected frequency

X	0	1	2	3	4
O	123	59	14	3	1
E	121	61	15	3	0

~~Assignment~~
~~Nov-2000~~
November 2011

Madras University

Pg. 21) CHI SQUARE TEST

Q1) $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

H_0: The new treatment is not superior to conventional treatment.
H_1: The new treatment is superior to conventional treatment.

α : 5%/0.05

χ^2 cal : $\sum \frac{(O_i - E_i)^2}{E_i}$

O_i	$E_i = \frac{RT \times CT}{GT}$	$(O_i - E_i)^2 \div E_i$
140	136	0.12
30	34	0.47
60	64	0.25
20	16	1
	χ^2cal =	1.84

χ^2tab $= \chi^2 (R-1),(C-1), \alpha$ df

$= (2-1) \times (2-1), 0.05$

$= 1 \times 1, 0.05$

$= 1, 0.05$

$= \underline{3.84}$

Result : χ^2cal < χ^2tab, accept H_0.
i.e., The new treatment is not superior to the conventional treatment.

Pg. 18)

23)

X_1	X_2	$(X_1 - \bar{X}_1)^2$	$(X_2 - \bar{X}_2)^2$
60	61	221.41	256
65	66	97.61	121
70	67	23.81	100
74	85	0.77	64
76	78	1.25	1
82	63	50.69	196
85	85	102.41	64
87	86	146.89	81
	88		121
	91		196
599	770	644.84	1200

$\bar{X}_1 = \frac{\Sigma X_1}{n_1} = \frac{599}{8} = 74.88$

$\bar{x}_2 = \frac{\Sigma x_2}{n} = \frac{770}{10} = 77$

$H_0 : \sigma_1^2 = \sigma_2^2 \quad H_1 : \sigma_1^2 \neq \sigma_2^2$

$F_{cal} = \frac{s_1^2}{s_2^2}$

$s_1 = \frac{\Sigma(x_1 - \bar{x}_1)^2}{n_1 - 1}$

$= \frac{644.84}{8-1}$

$= 92.12$

$s_2 = \frac{\Sigma(x_2 - \bar{x}_2)^2}{n_2 - 1}$

$= \frac{1200}{10-1}$

$= 133.33$

$F_{cal} : \frac{s_1^2}{s_2^2}$

less than 1 $\left[= \frac{92.12}{133.33} = 0.69 \right]$ $= \frac{133.33}{92.12} = 1.45$

Ftab : $F_{N-1,\, D-1,\, \alpha\, df}$

(less than 1 reciprocated)

$F_{D-1,\, N-1,\, \alpha\, df}$

$F_{10-1,\, 8-1,\, 0.05}$

$= F_{9,\, 7,\, 0.05}$

$= 3.68$

Inference : Fcal < Ftab, accept H_0,
i.e., The two samples have same variance
$\sigma_1 = \sigma_2$.

Pg. 21)

20) E_1 = ~~Defective~~ items produced by M I
E_2 = ~~Defective~~ items produced by M II.

A = Defective items.

E_i	$P(E_i)$	A	$P(A/E_i)$	$P(A/E_i) \times P(E_i)$
E_1	$P(E_1) = 0.30$	Defective	$P(A/E_1)$ 0.05	$0.05 \times 0.30 = 0.0150$
E_2	$P(E_2) = 0.70$	items	$P(A/E_2)$ 0.01	$0.01 \times 0.70 = 0.0070$
				0.0220

By Using Baye's Theorem

$$P(E_i/A) = \frac{P(A/E_i) \times P(E_i)}{\sum_{i=1}^{2} P(A/E_i) \times P(E_i)}$$

$$P(E_1/A) = \frac{0.0150}{0.0220}$$

$= 0.6818$

$$P(E_2/A) = \frac{0.0070}{0.0220}$$

$= 0.3182$

Page No.:
Date: / /

9/11/19
~~Tuesday~~
Friday

Multiple Regression →

Multiple regression analysis is one of the most important multivariate analysis which helps us to study the average relationship btw one dependent variable with one or more independent variables.

example

Washing Machine → capacity, Price, Colour, Brand, Advertisements } changes from time to time ↓ Independent variable

Dealer (to attain profit - have to sell) ↓ → Dependent variable

↓ Avg relationship.

Simple regression

$y = a_0 + a_1 x_1$

y ↓ DV (or) unknown variable

a_0, a_1 ↓ regression coefficient

x_1 ↓ IV ↓ Known variable

Multiple regression

$y = \underbrace{a_0 + a_1x_1 + a_2x_2 + \dots a_nx_n}$

y ↓ DV

Regression coefficient

x_n ↓ IV

① Intercept (turning point).

Factor Analysis

Identifying the underlying ~~roots~~.

Page No.: Date: / /

eg:- Asay (Iyyanar sir) perspective

1. Develop
2. Fees
3. Faculty
4. Work pressure (No)

Factors

Mango

1. Size
2. weight
3. Colour.

reducing on number mangoes to minimum to purchase.