ASSIGNMENT

MODEL PAPER - ①

MODULE - ①

1 a) A shipmen of 8 similar microcomputers to a retail out-let contains 3 that are defective. of a School makes a random purchase of 2 of these computers. find the probability distribution for the number of defectives. Find the mean and variance of this distribution?

⇒ No. of microcomputers = 8

No. of defective microcomputers = 3

No of ~~non~~ defective micdocomputers = 8 - 3 = 5

∴ x can takes the values 0, 1, 2.

$p(x=0)$ = probability of selecting 2 non-defective computers.

$$p(x=0) = \frac{5C_2}{8C_2} = \frac{5}{14}.$$

$p(x=1)$ = probability of selecting 1 defective and one non-defective computer.

$$p(x=1) = \frac{5C_1\, 3C_1}{8C_2} = \frac{15}{28}$$

$p(x=2)$ = probability of selecting 2 defective micro computers.

$$p(x=2) = \frac{3C_2}{8C_2} = \frac{3}{28}.$$

The probability distribution is

x	0	1	2
$p(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$

mean = $M = \Sigma x\, p(x)$

$$= 0\left(\frac{5}{14}\right) + 1\left(\frac{15}{28}\right) + 2\left(\frac{3}{28}\right)$$

$$\boxed{M = 0.75}$$

variance = $V = \Sigma x^2 p(x) - M^2$

$$= 0\left(\frac{5}{14}\right) + 1\left(\frac{15}{28}\right) + 4\left(\frac{3}{28}\right) - (0.75)^2$$

$$\boxed{V = 0.4018}$$

b) In a factory producing blades, the probability of any blade being defective is 0.002. If blades are supplied in packets of 10, using poisson distribution determine the number of packets containing,

(i) No defective

(ii) One defective and

(iii) Two defectives blades respectively in a consignment of 10000 packets

⇒ Number of blades supplied in packets, $n = 10$.

probability of defective blade $= 0.002 = p$

probability Distribution, $P(x) = \dfrac{m^x e^{-m}}{x!}$

Mean $= np = 10 \times 0.002 = 0.02$.

$$\boxed{P(x) = \frac{0.02^x \, e^{-0.02}}{x!} \rightarrow (1)}$$

consignment of 10,000 packets is,

$f(x) = 10.000 \times P(x)$

$f(x) = 10.000 \left[\dfrac{0.02^x \, e^{-0.02}}{x!}\right]$

(i) probability of no defective

$f(0) = 10000 \left[\dfrac{0.02^0 \cdot e^{-0.02}}{0!}\right]$

$f(0) = 9801.98$

$\boxed{f(0) = 9802}$

(ii) probability of one defective

$f(1) = 10000 \left[\dfrac{0.02^1 \cdot e^{-0.02}}{1!}\right]$

$= 196.0397$

$\boxed{f(1) = 196}$

(iii) probability of two defectives

$f(2) = 10000 \left[\dfrac{0.02^2 \cdot e^{-0.02}}{2!}\right]$

$= 1.9604$

$\boxed{f(2) = 2}$

c) If the mileage (in thousands of miles) of a certain radial tyres is a random variable with exponential distribution with mean 40,000 miles. Determine the probability that the tyre will last

(i) Atleast 20,000 miles.

(ii) Atmost 30,000 miles.

⇒ Exponential Distribution is given by

$$f(x) = \begin{cases} \alpha e^{-\alpha x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Mean $\mu = 40{,}000$

$$\alpha = \frac{1}{\mu} = \frac{1}{40{,}000}$$

$$\boxed{\therefore f(x) = \frac{1}{40{,}000} e^{-\frac{1}{40000}x}}$$

(i) The probability that the tyre will last atleast 20,000 miles.

$$P(x \geq 20000) = 1 - P(x < 20000)$$

$$= 1 - \int^{20000} f(x) \cdot dx$$

$$= 1 - \int^{20000} \frac{1}{40{,}000} e^{-\frac{1}{40{,}000}x}$$

$$= 1 - \frac{1}{40000} \left[\frac{-e^{-\frac{1}{40000}x}}{\frac{1}{40000}} \right]_0^{20000}$$

$$= 1 + \left[e^{-\frac{1}{40000}x} \right]_0^{20000}$$

$$= 1 + e^{-\frac{20000}{40000}} - e^{-0}$$

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

$$\boxed{P(x \geq 20000) = 0.607}$$

(ii) The probability that the tyre will last atmost 30.000 miles

$$P(x \leq 30000) \int_0^{30000} f(x) \cdot dx$$

$$= \int_0^{30000} \frac{1}{40000} e^{-\frac{1}{40000}x} \cdot dx$$

$$= \frac{1}{40000} \left[\frac{-e^{-\frac{1}{40000}x}}{\frac{1}{40000}} \right]_0^{30000}$$

$$= -\left[e^{\frac{-30000}{40000}} - e^{-0} \right]$$

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

$$= -e^{\frac{3}{4}} + 1$$

$$\boxed{P = (x \leq 30000) = 0.5276}$$

OR

2 a) The density function of a random variable x is given by

$$f(x) = \begin{cases} K\sqrt{x}, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

(i) find K.

(ii) Find the cdf. F(K) and use it to evaluate $P[0.3 < x < 0.6]$

⇒ (i) To find K, we use the formula.

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

$$\Rightarrow \int_0^1 K\sqrt{x} \cdot dx = 1 \Rightarrow K \int_0^1 \sqrt{x} \cdot dx = 1 \Rightarrow K \left[\frac{x^{3/2}}{3/2} \right]_{x=0} = 1$$

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

$$\boxed{K = \frac{3}{2}} \qquad \therefore f(x) = \begin{cases} \frac{3}{2}\sqrt{x} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

(ii) cdf $F(x) = \int_{-\infty}^{x} k\sqrt{t}\cdot dt = \frac{3}{2}\int_{-\infty}^{x} \sqrt{t}\cdot dt = \frac{3}{2}\int_{-\infty}^{0} \sqrt{t}\cdot dt + \frac{3}{2}\int_{0}^{x} \sqrt{t}\cdot dt$

$$= 0 + \frac{3}{2}\left[\frac{t^{3/2}}{3/2}\right]_0^x = \left(t^{3/2}\right)_0^x = x^{3/2}$$

$$P(0.3 < x < 0.6) = F_x(0.6) - F_x(0.3)$$

$$= (0.6)^{3/2} - (0.3)^{3/2}$$

$$\boxed{P = (0.3 < x < 0.6) = 0.3004}$$

---

b) Find the mean and variance of Binomial Distribution?

⇒ Mean = $\mu = \Sigma x \; p(x)$

$P(x)$ for Binomial Distribution is.

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

$$\mu = \sum_{x=0}^{n} x\, {}^nC_x\, p^x\, q^{n-x}$$

$$= \sum_{x=0}^{n} x \frac{n!}{x!\,(n-x)!} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \frac{n!}{(x-1)!\,(n-x)!} p^x q^{n-x}$$

$$= \sum_{x=1}^{n} \frac{n!}{(x-1)!\,(n-x)!} p^x q^{n-x}$$

$$= \sum_{x=1}^{n} \frac{n(n-1)!}{(x-1)!\,[(n-1)-(x-1)]!} p^{x-1}\cdot p\, q^{[(n-1)-(x-1)]}$$

$$= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,[(n-1)-(x-1)]!} p^{x-1} q^{[(n-1)-(x-1)]}$$

$$= np \sum_{x=1}^{n} {}^{(n-1)}C_{x-1}\, p^x\, q^{[(n-1)-x-1)]}$$

$$\mu = np[(q+p)^{n-1}] \qquad \mu = np\,(1)^{n-1} \qquad \boxed{\mu = np}$$

VARIANCE

Variance = $V = \Sigma x^2 p(x) - \mu^2$

Probability for Binomial Distribution is,

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

$$V = \sum_{x=0}^{n} (x^2+x-x)\, p(x) - \mu^2$$

$$= \sum_{x=0}^{n} (x^2-x)\, p(x) + \sum_{x=0}^{n} x\, p(x) - \mu^2$$

$$= \sum_{x=0}^{n} (x^2-x)\, p(x) + \mu - \mu^2$$

$$= \sum_{x=0}^{n} x(x-1)\, p(x) + \mu - \mu^2$$

$$= \sum_{x=0}^{n} = x(x-1)\left[{}^nC_x\, p^x q^{n-x}\right] + \mu - \mu^2$$

$$= \sum_{x=0}^{n} x(x-1) \frac{n!}{x!\,(n-x)!}\, p^x q^{n-x} + \mu - \mu^2$$

$$= \sum_{x=0}^{n} \frac{n!}{x(x-1)\,(x-2)!\,[(n-x)!]}\, x(x-1)\, p^x q^{n-x} + \mu - \mu^2$$

$$= \sum_{x=0}^{n} \frac{n!}{(x-2)!\,[(n-x)]!}\, p^x q^{n-x} + \mu - \mu^2$$

$$= \sum_{x=2}^{n} \frac{n!}{(x-2)!\,[(n-2)-(x-2)]!}\, p^{x-2} \cdot p^2 \cdot q^{[(n-2)-(x-2)]} + \mu - \mu^2$$

$$= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!\,[(n-2)-(x-2)]!}\, p^{x-2} \cdot q^{[(n-2)-(x-2)]} + \mu - \mu^2$$

$$= n(n-1)\, p^2 (q+p)^{n-2} + np - (np)^2$$

$$= n(n-1)\, p^2 + np - n^2p^2$$

$$= n^2p^2 - np^2 + np - n^2p^2$$

$$= np - np^2$$

$$= np(1-p)$$

$$\therefore \boxed{V = npq}$$

c) In a test on 2000 electric bulbs. it was found that the life of a particular make, was normally distributed with an average life 2040 hours and standard deviation of 60 hours. Estimate the number of bulbs likely to burn for.

(i) More than 2150 hours.

(ii) Less than 1950 hours.

(iii) In between 1920 and 2160 hours.

⇒ Given:- Average life of electric bulbs = $\mu = 2040$

Standard deviation σ = 60hrs.

Let 'x' be a variate of duration of the bulb.

(i) More than 2150 hours.

$P(x > 2150)$

when, $x = 2150$, $z = \frac{x-\mu}{\sigma} = \frac{2150 - 2040}{60} = 1.83$

$P(z > 1.83)$

$z(0 \text{ to } \infty) - z(0 \text{ to } 1.83)$

$0.5 - \phi(1.83)$

$0.5 - 0.4664$

$= 0.0336$

$$\boxed{P(x > 2150) = 0.0336}$$

No. of 2000 bulbs that are likely to last more than 2150 hours = $2000 \times 0.0336 = 67.2 = \underline{67}$ bulbs.

(ii) $P(x < 1950)$

when $x = 1950$, $z = \frac{x-\mu}{\sigma} = \frac{1950 - 2040}{60} = -1.5$

$P(z < -1.5)$

$z(0 \text{ to } \infty) - z(0 \text{ to } 1.5)$

$= 0.5 - \phi(1.5)$

$= 0.5 - 0.4332$

$\boxed{P(z < -1.5 = 0.0668}$

No. of 2000 bulbs that are likely to last less than 1950 hours $= 2000 \times 0.0668 = 133.6 \simeq$ 134 bulbs

(iii) $P(1920 < x < 2160)$

when $x = 1920$, $z = \frac{1920 - 2040}{60} = -2$

when $x = 2160$ $z = \frac{2160 - 2040}{60} = 2$

$P(-2 < z < 2) = z\,(-2 \text{ to } 0) + z\,(0 \text{ to } 2)$

$= 2 \times z\,(0 \text{ to } 2)$

$= 2\phi\,(2)$

$= 2(0.4772)$

$\boxed{P(-2 < z < 2) = 0.9544}$

No. of 2000 bulbs that are likely to born in between 1920 and 2160 hours $= 2000 \times 0.9544 = 1908 \simeq$ 1909 bulbs

MODULE - 2

3 a) The joint probability distribution of two random variables. x and y is.

$y \backslash x$	-4	2	7
1	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$
5	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

Find the marginal distribution obtain the condition co-efficient between x and y.

⟹ Marginal distribution of x and y.

Distribution of x

$x = x_i$	$x_1 = 4$	$x_2 = 2$	$x_3 = 4$
$f(x_i)$	$f(x_1) = \frac{3}{8}$	$f(x_2) = \frac{3}{8}$	$f(x_3) = \frac{1}{4}$

Distribution of y

$y = y_i \quad = y_1 = 1 \quad y_2 = 5$

$g(y_i) \quad g(y_1) = \frac{1}{2} \quad g(y_2) = \frac{1}{2}$.

corelation. $\rho(x, y) = \dfrac{cov(x, y)}{\sigma_x \cdot \sigma_y}$

$E(x) = \Sigma x_i f(x_i)$

$E(x) = -\left(4 \times \frac{3}{8}\right) + \left(2 \times \frac{3}{8}\right) + \left(7 \times \frac{1}{4}\right)$

$\boxed{E(x) = 1}$

$E(y) = \Sigma y_j g(y_j)$

$E(y) = \left(1 \times \frac{1}{2}\right) + \left(5 \times \frac{1}{2}\right)$

$\boxed{E(y) = 3}$

$E(xy) = \left(-4 \times 1 \times \frac{1}{8}\right) + \left(2 \times 1 \times \frac{1}{4}\right) + \left(7 \times 1 \times \frac{1}{8}\right) + \left(-4 \times 5 \times \frac{1}{4}\right) + \left(2 \times 5 \times \frac{1}{8}\right)(7 \times 5 \times$

$\boxed{E(xy) = \frac{3}{2}}$

$E(x^2) = \left((-4)^2 \times \frac{3}{8}\right) + \left((2)^2 \times \frac{3}{8}\right) + \left((7)^2 \times \frac{1}{4}\right)$

$$\boxed{E(x^2) = \frac{79}{4}}$$

$V(x) = E(x^2) - [E(x)]^2$

$= \frac{79}{4} - 1$

$$\boxed{V(x) = \frac{75}{4}}$$

$$\boxed{\sigma(x) = \sqrt{\frac{75}{4}}}$$

$E(y^2) = \left(1^2 \times \frac{1}{2}\right) + \left(5^2 \times \frac{1}{2}\right)$

$$\boxed{E(y^2) = 13}$$

$V(y) = E(y^2) - [E(y)]^2$

$13 - 9$

$$\boxed{V(y) = 4}$$

$$\boxed{\sigma(y) = 2}$$

$\text{cov}(x, y) = E(xy) - E(x)\, E(y) \Rightarrow \frac{3}{2} - 1 \times 3$

$$\boxed{\text{cov}(x, y) = -\frac{3}{2}}$$

$P(x, y) = \frac{\text{cov}(x, y)}{\sigma x \cdot \sigma y}$

$= \frac{-3/2}{\sqrt{\frac{75}{4}} \times 2}$

$\therefore \boxed{P(x, y) = -0.1732}$

b) Find the unique fixed probability vector of

$$P = \begin{bmatrix} 0 & 3/4 & 1/4 \\ 1/2 & 1/2 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$\Rightarrow$ Let $v = (x, y, z)$ be the unique fixed probability vector.

$\boxed{x+y+z=1}$ $\ni$ $vP = v \rightarrow ①$.

$vP = v$.

$$[x \;\; y \;\; z] \begin{bmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{bmatrix} = [x, y, z]$$

$$[0+y/2+0, \; 3x/4+y/2+z, \; x/4] = [x, y, z]$$

$$[y/2, \; 3x/4+y/2+z, \; x/4] \quad [x, y, z]$$

$$\frac{y}{2} = x, \quad \frac{3x}{4}+\frac{y}{2}+z = y, \quad \frac{x}{4} = 3$$

$\boxed{y = 2x} \rightarrow ②$ $\qquad$ $\boxed{z = \frac{x}{4}} \rightarrow ③$

using ② and ③ in equation ①

$$x+y+z = 1$$

$$x+2x+\frac{x}{4} = 1$$

$$\frac{4x+8x+x}{4} = 1$$

$$\frac{13x}{4} = 1$$

$$13x = 4$$

$$x = \frac{4}{13}, \quad y = \frac{8}{13}, \quad z = \frac{4}{52} = \frac{1}{13}$$

$\boxed{\text{Thus, } v = \frac{4}{13}, \frac{8}{13}, \frac{1}{13}}$

c) Every year, a man trades his car for a new car, if he has a Maruthi, he trades it for an Ambassador. if he has ambassador, he trades it for Santro however if he had Santro he is just as likely to trade it for a new santro as to trades it for a Maruthi or an Ambassador. In 2000 he bought his 1st car which was a santro. Find the probability that he has.

(i) 2002 Santro

(ii) 2002 Maruti

(iii) 2003 Ambassador.

(iv) 2003 Santro.

⇒ State space = {Maruti, Ambassador, Santro}

Associated Transition = $P = P_{ij} =$

$$\begin{array}{c|ccc} & M & A & S \\ \hline M & 0 & 1 & 0 \\ A & 0 & 0 & 1 \\ S & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}$$

He bought his 1st car in 2000

2002 ⇒ This implies 2 years after 1st car.

2000 → 2002.

$$P^2 = P \cdot P \quad \begin{array}{c} \\ M \\ A \\ S \end{array} \begin{array}{c} \begin{array}{ccc} M & A & S \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \end{array} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

$$P^2 = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{9} & \frac{4}{9} & \frac{4}{9} \end{bmatrix} = \begin{bmatrix} P_{11}^{(2)} & P_{12}^{(2)} & P_{13}^{(2)} \\ P_{21}^{(2)} & P_{22}^{(2)} & P_{23}^{(2)} \\ P_{31}^{(2)} & P_{32}^{(2)} & P_{33}^{(2)} \end{bmatrix}$$

$$P^3 = P^1 \cdot P^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{9} & \frac{4}{9} & \frac{4}{9} \end{bmatrix}$$

$$P^3 = \begin{array}{c|ccc} & M & A & S \\ \hline M & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ A & \frac{1}{9} & \frac{4}{9} & \frac{4}{9} \\ S & \frac{4}{27} & \frac{7}{27} & \frac{16}{27} \end{array} = \begin{bmatrix} P_{11}^{(3)} & P_{12}^{(3)} & P_{13}^{(3)} \\ P_{21}^{(3)} & P_{22}^{(3)} & P_{23}^{(3)} \\ P_{31}^{(3)} & P_{32}^{(3)} & P_{33}^{(3)} \end{bmatrix}$$

In 2000 - 2002

(i) Santro - Santro $= P_{33}^{(2)} = \frac{4}{9}$

(ii) Santro - Maruti $= P_{31}^{(2)} = \frac{1}{9}$

In 2000 - 2003.

(iii) Santro - Ambassador

$P_{32}^{(3)} = \frac{7}{27}$

(iv) Santro - Santro

$P_{33}^{(3)} = \frac{16}{27}$

In the long run, of probability of having santro is.

$P(S) = \frac{3}{6} = \frac{1}{2} = 0.5$

4 a) The joint probability distribution of 2 random variables x and y is

x \ y	-3	2	4
1	0.1	0.2	0.2
3	0.3	0.1	0.1

(i) Are x and y are independent
(ii) Evaluate $P(y \le 2)$
(iii) Evaluate $P[x + y \le 2]$

⇒ marginal Distribution of x:

$X = x_i$	$x_1 = -3$	$x_2 = 2$	$x_3 = 4$
$f(x_i)$	$f(x_1) = 0.4$	$f(x_2) = 0.3$	$f(x_3 = 0.3$

$f(x_1) = 0.1 + 0.3 = 0.4$
$f(x_2) = 0.2 + 0.1 = 0.3$
$f(x_3) = 0.2 + 0.1 = 0.3$.

marginal Distribution of y.

$y = y_j$	$y_1 = 1$	$y_2 = 3$
$g(y_j)$	$g(y_1) = 0.5$	$g(y_2) = 0.5$

$g(y_1) = 0.1 + 0.2 + 0.2 = 0.5$
$g(y_2) = 0.3 + 0.1 + 0.1 = 0.5$.

(i) If x and y are independent random variables, then
$f(x_i) \cdot g(y_j) = J_{ij}$
Let us assume x_1 and y_2
$f(x_1) = 0.4, \quad g(y_2) = 0.5$
$f(x_1) \cdot g(y_2) = 0.4(0.5) = 0.2$
$J_{12} = 0.3$ Here, $f(x_1) \cdot g(y_2) \ne J_{ij}$

So, X and y are dependent random variables.

(iii) $P(y \leq 2)$

$X = \{-3, 2, 4\} = (x_1, x_2, x_3)$

$y = \{1\} = \{y_1\}$

$P(y \leq 2) = J_{11} + J_{21} + J_{31}$

$= 0.1 + 0.2 + 0.2$

$\boxed{P(y \leq 2) = 0.5}$

(iv) $P(x+y \leq 2)$

$x = \{-3, 2, 4\} = \{x_1, x_2, x_3\}$

$y = \{1, 3\} = \{y_1, y_2\}$

$P(x+y \leq 2) = J_{11} + J_{21}$

$= 0.1 + 0.2$

$\boxed{P(x+y \leq 2) = 0.3}$

Q 6) Define probability vectors, stochastic matrices, Regular stochastic matrix, Stationary distribution and absorbing state of markov chain?

⇒ 1) Probability vector:-

A vector $v = (v_1, v_2, \ldots, v_n)$ is called a probability vector if each one of its component are non-negative and their sum is equal to unity.

Example:- $u = (1, 0)$

$v = \left(\frac{1}{2}, \frac{1}{2}\right)$

$w = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$

2) **Stochastic matrices :-**

A stochastic matrices P is said to be, A square matrix $P = P_{ij}$ having every row in the form of probability vector is called a stochastic matrix.

Example :

$$v = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$$

$$W = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

3) **Regular stochastic matrix :-**

A stochastic matrix P is said to be regular stochastic matrix if all the elements of entries of some power P^n are positive.

Example :- $A = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

$$A^2 = A \cdot A = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$$

A is regular stochastic matrix and we have $\boxed{n = 2}$

4) **Stationary Distribution :-**

As $n \to \infty$ $P_j^{(n)} = V_j$ where $j = 1, 2, 3, \ldots, m$ is called the stationary Distribution of a markov chain.

$p^{(n)} = p^{(0)} \cdot p^{(n)}$ where $p^{(n)} = [P_1^{(n)}, P_2^{(n)}, \ldots P_m^{(n)}]$

5) **Absorbing state of a Markov chain :-**

In a markov chain, the process reaches to a certain state after which it continues to remain in the same state, Such a state is called an Absorbing state of a markov chain.

i.e; $P_{ij} = 1$ for $i = j$ and 0 otherwise

Example
$$P = \begin{array}{c} a_1 \\ a_2 \\ a_3 \end{array} \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$
The state is absorbing state of Markov chain.

c) A Salesman's territory consists of 3 cities A, B and C. He never sells in the same city on successive days. If he sells in city A, then the next day he sells in city B. However, if he sells in either B or C, then the next day he is twice as likely to sell in city A as in other city. In long run, how often does he sells in each of the cities?

$$TPM = P = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{bmatrix} \end{array}$$

To find unique fixed probability vector, WKT, let (x, y, z) be UFPV

$\Rightarrow VP = V$ and $x+y+z=1$

$$(x \;\; y \;\; z) \begin{bmatrix} 0 & 1 & 0 \\ 2/3 & 0 & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{bmatrix} = [x \;\; y \;\; z]$$

$$\begin{bmatrix} 0 + \frac{2y}{3} + \frac{2z}{3} & , x + \frac{3}{3}, & \frac{y}{3} \end{bmatrix} = [x \quad y \quad z]$$

$$\begin{bmatrix} \frac{2y}{3} + \frac{2z}{3} & , \frac{x+z}{3} & , \frac{y}{3} \end{bmatrix} = [x \quad y \quad z]$$

Equating

$\frac{2y}{3} + \frac{2z}{3} = x \qquad x + \frac{z}{3} = y \quad , \frac{y}{3} = z$

$2y + 2z = 3x \qquad y = 3z$

$2(3z) + 2z = 3x$

$6z + 2z = 3x$

$8z = 3x$

$x = \frac{8z}{3}$

WKT; $x+y+z = 1$

$\frac{8z}{3} + 3z + z = 1 \Rightarrow 8z + 9z + 3z = 3$

$20z = 3$

$\boxed{z = \frac{3}{20}}$, $\boxed{y = \frac{9}{20}}$, $\boxed{x = \frac{8}{20}}$

$$\boxed{\therefore v(x, y, z) = \frac{8}{20}, \frac{9}{20}, \frac{3}{20} = (A, B, C)}$$

In the long run of probability,

In city A, he sells 8/20,

In city B, he sells 9/20,

In city C, he sells 3/20,

MODULE - 3

5 a) Define Null Hypothesis, significance level, critical region, Type-I and Type-II errors in a statistical test ?

⇒ 1) Null Hypothesis :- It predicts there is no relationship between two variables. It is denoted by H_0. It is followed by '=' sign.

Example :- New drug does not reduce the number of days to recover from a disease compare to a standard deviation.

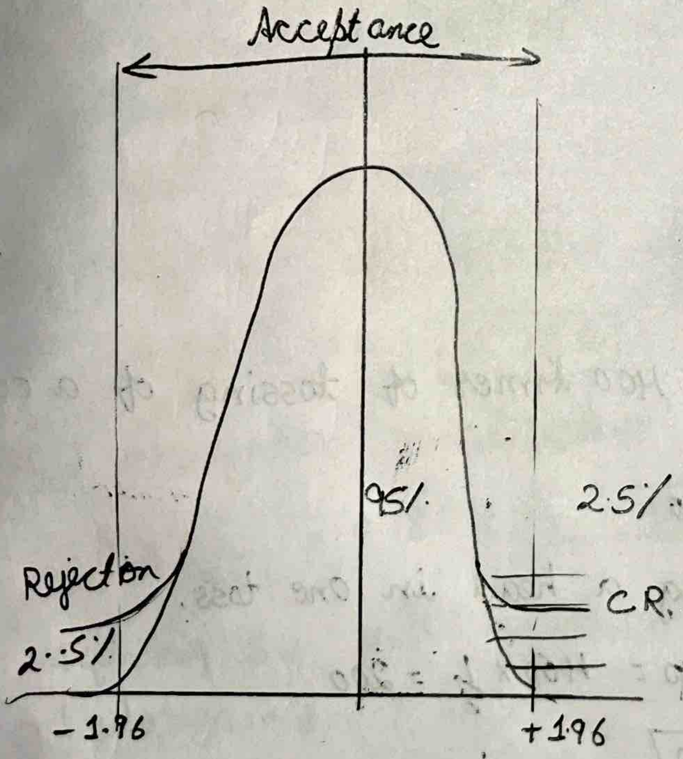
2) Significance level :- Significance level are also known as alpha (α) level (or) level of significance.

It is a parameter used in hypothesis testing to determine the threshold at which the null hypothesis is rejected.

It is denoted by α.

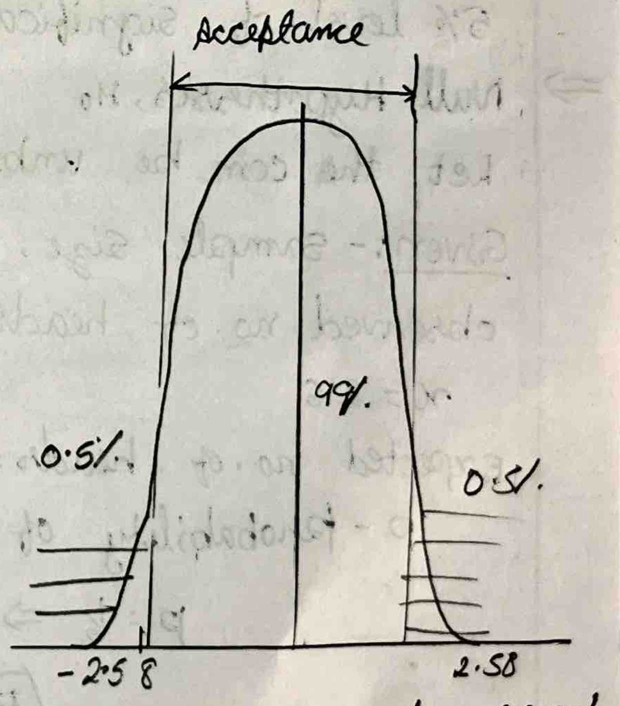
The significance level represents the probability of rejecting (or) accepting the null hypothesis.

commonly used significance levels are 0.05 and 0.01



95% sure of correct decision

5% chance of wrong decision



99% of confidence of correct decision

1% chance of wrong decision

3) Critical region :- A region which amounts to the rejection of null hypothesis is called critical region (or) region of rejection.

4) Type-I error :- A Type-I error occurs if a null hypothesis is rejected that is actually true is the population.
H_0 is rejected instead of accepting.
Example :- There is no relationship between eating chocolate and Back pain.

5) Type-II error :- A Type-II error occurs if a null hypothesis is not rejected that is actually false in the population.
H_0 is accepted instead of rejecting.
Example :- There is no relationship between eating chocolate and toothache.

b) A coin was tossed 400 times and head turned up 216 times. Test the hypothesis that the coin is unbiased at 5% level of significance?

⇒ Null Hypothesis, H_0
Let the coin be unbiased
Given :- sample size, $n = 400$
observed no. of heads on 400 times of tossing of a coin.
$x = 216$
Expected no. of heads $= \mu = np$
p - probability of getting a head in one toss.

$$p = \frac{1}{2} \Rightarrow \mu = np = 400 \times \frac{1}{2} = 200$$

$$\boxed{\mu = 200}$$

$$p + q = 1$$

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

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

Test of significance $\Rightarrow$ $z = \frac{x-\mu}{\sigma} = \frac{x-np}{\sqrt{npq}}$

$$z = \frac{216-200}{\sqrt{200 \times \frac{1}{2}}} = \frac{16}{10} = 1.6$$

$$\boxed{z = 1.6}$$

$|z| < z_\alpha$

$1.6 < 1.96$

[$z_\alpha = 1.96$ for 5% level of significance.

$\therefore$ Accept the Hypothesis. H_0

$\therefore$ The coin is unbiased.

---

c) In a city A 20% of random sample of 900 school boys had a certain slight physical defect. In another city B, 18.5% of a random sample of 1600 school boys had the same defect. Is the difference between the proportions significant at 5% significance level.

$\Rightarrow$ $n_1 = 900$, $n_2 = 1600$.

P_1 = proportion of defective of eyesight in city A.

$P_1 = \frac{20}{100} = 0.2$, $\boxed{P_1 = 0.2}$

P_2 = proportion of defective of eyesight in city B.

$P_2 = \frac{185}{100} = 0.185$, $\boxed{P_2 = 0.185}$

defining the Null Hypothesis, $H_0 : P_1 = P_2$

calculation:- $z = \frac{P_1 - P_2}{\sqrt{PQ\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

$$P = \frac{(900 \times 0.2) + (1600 \times 0.185)}{900 + 1600}$$

$$\boxed{P = 0.190}$$

$q = 1 - p$

$q = 1 - 0.190$

$$\boxed{q = 0.810}$$

$$Z = \frac{0.2 - 0.185}{\sqrt{(0.19)(0.81)\left(\frac{1}{900} + \frac{1}{1600}\right)}}$$

$$\boxed{Z = 0.918}$$

$|Z| < Z_{\alpha}\ (0.05)$

$0.918 < 1.96$

$\therefore$ Accept the Null Hypothesis, H_0.

OR

6 a) Explain the following terms:-

i) Standard error

ii) statistical hypothesis

iii) Critical region of a statistical test.

iv) Test of significance.

$\Rightarrow$ i) Standard error:- Standard error of a statistic is the approximate standard deviation of a statistical sample population.

ii) Statistical hypothesis:- It is a statement that can be tested by scientific research. If we want to test a relationship between 2 or more things, we need to write hypothesis before we start experiment (or) data collection.

Example: consumption of apple everyday leads to visit the fewer times.

3) Critical region of a statistical Test :- A region which amount to the rejection of Null Hypothesis is called critical region (or) region of rejection.

4) Test of significance :- The process which helps to decide about the acceptance (or) rejection of the hypothesis is called the Test of significance.

∴ i,e; to reach to any decision about the population, it is essential to make certain assumption, such an assumption is called as statistical Hypothesis.

6 b) A die was thrown 9000 times and a throw of 5 or 6 was obtained 3240 times. On the assumption of random throwing do the data indicate an unbiased die at 1% level of significance?

⇒ Let us assume Null Hypothesis is, H_0 - die cannot be regarded as an unbiased one.

Given:- n - sample size = 9000 times.

observed no. of success, $x = 3240$

Expected no. of success, $\mu = np$

p - probability of getting 5 or 6 is $\frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$

$q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3}$

Test of significance, $Z = \frac{x - \mu}{\sigma} = \frac{3240 - 9000\left(\frac{1}{3}\right)}{\sqrt{9000\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)}}$

$Z = 5.3666$

$\boxed{Z = 5.37}$

$|Z| > Z_\alpha$ for 1% level of significance

∴ $5.37 > 2.58$ [$Z_\alpha = 2.58$ for 1%]

∴ Reject the Hypothesis, H_0

i.e; The die is biased.

c) In a sample of 600 men from a certain city, 450 are found smokers. In another sample of 900 men from another city 450 are smokers. Do the indicate, that the cities are significantly different with respect to the habit of smoking among men. Test at 5% significance level?

⇒ Let us assume Null Hypothesis, H_0 -

There is a difference with respect to the habit of smoking among men.

Given:- $n_1 = 600 \quad n_2 = 900$

$$P_1 = \frac{450}{600}, \quad P_2 = \frac{450}{900}$$

We have, $$Z = \frac{P_1 - P_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$p = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

$$p = \frac{600\left(\frac{450}{600}\right) + 900\left(\frac{450}{900}\right)}{600 + 900}$$

$$q = 1 - p = 1 - 0.6 = 0.4$$

$$Z = \frac{0.75 - 0.5}{\sqrt{(0.6)(0.4)\cdot\left(\frac{1}{600} + \frac{1}{900}\right)}}$$

$$\boxed{Z = 9.682}$$

For 5% level of significance, $Z_\alpha = 1.96$

$|Z| > Z_\alpha$

$9.682 > 1.96$

∴ Hypothesis is rejected.

∴ There is no difference with respect to the habit of smoking among the men

7 a) state central limit theorem. use the theorem to evaluate $P[50 < \bar{X} < 56]$ where $\bar{x}$ represents the mean of a random sample of size 100 from an infinite population with mean $\mu = 53$ and variance $\sigma^2 = 400$

⇒ The central limit theorem states that the sample mean $\bar{x}$ follows approx the normal distribution with μ and std $\frac{\sigma}{\sqrt{n}}$ i.e., $\bar{x} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$ where, μ, σ. are mean and standard deviation.

sample size $n = 100$

$\mu = 53$

$\sigma^2 = 400$

$\sigma = \sqrt{400} = 20$

$$z = \frac{\bar{x} - \mu}{\sigma \sqrt{N}}$$

$\bar{x} = 50$

$$z = \frac{50 - 53}{2} = -\frac{3}{2} = -1.5 = z_1$$

$\bar{x} = 56$

$$z = \frac{56 - 53}{2} = \frac{3}{2} = \frac{3}{2} = 1.5 = z_2$$

$P(50 < \bar{x} < 56) = 0.8664$.

7 b) A random sample of size 25 from a normal distribution $N(\mu, \sigma^2 = 4)$ yields, sample mean $\bar{x} = 78.3$. obtain a 99% confidence interval for μ.

$\Rightarrow$ $n = 25$

$\bar{x} = 78.3$

$\sigma = 2$

confidence level of 99%, Z value is 2.58.

confidence interval $C.I = \mu = \text{Mean} \pm Z \times \frac{\sigma}{\sqrt{n}}$

$$\mu = 78.3 \pm \left(2.58 \times \frac{2}{\sqrt{25}}\right.$$

$$\mu = 78.3 \pm 1.032$$

$$C.I = (78.3 - 1.032, 78.3 + 1.032)$$

$$C.I = (77.268, 79.332)$$

7 c) A survey of 320 families with 5 children each reveal the following distribution.

No. of boys	5	4	3	2	1	0
No. of girls	0	1	2	3	4	5
No. of families	14	56	110	88	40	12

Is the result consistent with the hypothesis that male and female births are equally probable at 5% level of significance.

$\Rightarrow$ $n = 320$

$p = \frac{1}{2} = 0.5$

$q = \frac{1}{2} = 0.5$

$n = 5$

$H_0 \rightarrow$ probability of female and male birth is equ
By Binomial distribution

8

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

$$P(x) = 5C_x\,(0.5)^5$$

$$F(x) = 320 \times P(x) = 320 \times 5C_x\,(0.5)^5.$$

$x = 0$

$F(0) = 10$
$F(1) = 50$
$F(2) = 100$
$F(3) = 100$
$F(4) = 50$
$F(5) = 10$

O_i	14	56	110	88	40	12.
E_i	10	50	100	100	50	10

$$\chi^2 = \Sigma\left[\frac{(O_i - E_i)^2}{E_i}\right]$$

$$\chi^2 = 1.6 + 0.72 + 1 + 1.44 + 2 + 0.4.$$

$$\chi^2 = 7.16.$$

Table value $= 11.02$

$7.16 < 11.02$

$\therefore$ Null Hypothesis accepted.

8 a) A random sample of size 64 is taken from an infinite population having mean 112 and variance 144. Using central limit theorem, find the probability of getting the sample mean $\bar{x}$ greater than 114.5.

⇒ $n = 64$

$\mu = 112$

$\sigma^2 = 144$

$\sigma = 12$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$z = \frac{\bar{x} - 112}{12\ \sqrt{64}}.$$

$$z = \frac{\bar{x} - 1.12}{1.5}.$$

$\bar{x} = 114.5$

$$z = \frac{114.5 - 112}{1.5}$$

$z = 1.66$

$P(z > 1.66)$

$0.5 - P(0 < z < 1.66)$

$P(z > 1.66)$

$= 0.0489$

b) Let the observed value of the mean $\bar{x}$ of a random sample of size 20 from a normal distribution with mean μ and variance $\sigma^2 = 80$ be 81.2. Find a 90% and a 95% confidence intervals for μ.

$\Rightarrow$ $n = 20$

$\bar{x} = 81.2$.

$\sigma^2 = 80$

$\sigma = 8.9442$.

confidence level of 95%, 90%, the corresponding z values are 1.96, 1.645.

$$C.I = \mu = \text{mean} \pm z \left(\frac{\text{standard} \times \text{Standard Deviation}}{\sqrt{\text{sample size}}} \right)$$

$$\mu = \bar{x} + z \frac{(\bar{V})}{\sqrt{n}}$$

$$\mu = 81.2 \pm \left(1.96 \times \frac{8.9442}{\sqrt{20}}\right)$$

95% C.I $= (81.2 - 3.92 , 81.2 + 3.92)$

C.I $= (77.28 , 85.12)$.

90% C.I $= \mu = 81.2 \pm \left(1.645 \times \frac{8.9442}{\sqrt{20}}\right)$

$\mu = 81.2 \pm 3.29$

C.I $= (81.2 - 3.29 , 81.2 + 3.29)$

C.I $= (77.91 , 84.49)$

8 c) The nine items of a sample have the following values: 45, 47, 50, 52, 48, 47, 49, 53, 51. Does the mean of these differ significantly from the assumed mean of 47.5 at 5% significance level?

⇒ $n = 9$

$\mu = 47.50$

$\bar{x} = \frac{1}{n} \Sigma x = \frac{442}{9} = 49.11$

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

$$S^2 = \frac{1}{8} \left\{ \begin{array}{l} (45-49.11)^2 + (47-49.11)^2 + (50-49.11)^2 + (52-49.11)^2 \\ (48-49.11)^2 + (47-49.11)^2 + (49-49.11)^2 + (53-49.11)^2 \\ (51-49.11)^2 \end{array} \right.$$

$S^2 = \frac{54.9}{8} = 6.8625 \Rightarrow S = \sqrt{6.8625} = 26196$

$H_0 = \mu = 47.5$

$$t = \frac{49.11 - 47.5}{(2.6196/\sqrt{9})}$$

$$t = \frac{1.61}{0.8732}$$

$t = 1.8437$

Level of significance = 5%

d.f = 8 is 2.3060

$1.8437 < 2.3060$

Null Hypothesis accepted

9 a) Three different kinds of food are tested on three groups of rats for 5 weeks. The objective is to check the difference in mean weight (in grams) of the rats per week, Apply one-way ANOVA using a 0.05 significance level to the following data:

Food 1	8	12	19	8	6	11
Food 2	4	5	4	6	9	7
Food 3	11	8	7	13	7	9

$\Rightarrow$ $F1 = 8 + 12 + 19 + 8 + 6 + 11$

Total $= T_1 = 64$, $T_1^2 = 4096$

$F2 = 4 + 5 + 4 + 6 + 9 + 7$

Total $= T_2 = 35$ $T_2^2 = 1225$

$F3 = 11 + 8 + 7 + 13 + 7 + 9$

Total $= T3 = 55$ $T_3^2 = 3025$

Total $= 64 + 35 + 55 = 154$

$F_1 = 64 + 144 + 361 + 64 + 36 + 121$

Sum of square $= 790$

$F2 = 16 + 25 + 16 + 36 + 81 + 49$.

Sum of square $= 223$

$F3 = 121 + 64 + 49 + 169 + 49 + 81$

sum of square $= 533$.

$F1 + F2 + F_3$

$790 + 223 + 533 = 1546$

$$C.F = \frac{T^2}{N} = \frac{(154)^2}{18} = \frac{23716}{18} = 1317.55$$

Total sum of square.

$= 1546 - 1317.55$

$= 228.45$

$$SST = \sum_i \frac{T_i^2}{n_i} - CF$$

$$SST = \frac{4096}{6} + \frac{1225}{6} + \frac{3025}{6} - 1317.55$$

$$SST = 73.45$$

$$SSE = TSS - SST$$

$$SSE = 228.45 - 73.45$$

$$SSE = 155.$$

sources of variation	d.f	SS	MSS	F. Ratio
B/W Treatments	3 − 1 = 2	SST 73.45	MST 36.725	$F = \frac{36.725}{10.33} = 3.55$
Error	18 − 3 = 15	SSE 155	MSE 10.33	
Total				

3.55 < 3.68 F (2, 15) → 5% significance.

null hypothesis accepted.

9 b) Analyze and interpret the following statistics concerning output of wheat per field obtained as a result of experiment conducted to test four varieties of wheat viz. A, B, C and D under a Latin-Square design.

C 25	B 23	A 20	D 20
A 19	D 19	C 21	B 18
B 19	A 14	D 17	C 20
D 17	C 20	B 21	A 15

⇒ $H_0 \to$ No significant difference between rows, col, treatment

				T
C 5	B 3	A 0	D 0	8
A -1	D -1	C 1	B. -2	-3
B -1	A -6	D -3	C 0	-10
D -3	C 0	B 1	A -5	-7
0	-4	-1	-7	= -12

$n = 16$

$c = 4$

$r = 4$

$L = 4$

$T = -12$

$$SSR = \frac{\Sigma T_i^2}{n_i} - \frac{T^2}{n} = \frac{8^2 + (-3)^2 + (-10)^2 + (-7)^2}{4} - \frac{(-12)^2}{16}$$

$$SSR = 46.5$$

$$SSC = \frac{\Sigma T_j^2}{n_j} - \frac{T^2}{n} = \frac{0^2 + (-4)^2 + (-1)^2 + (-7)^2}{4} = \frac{(-12)^2}{16}$$

$$SSC = 7.5$$

$$SST = \Sigma x^2 - \frac{T^2}{n}$$

$$= 122 - \frac{(-12)^2}{16}$$

$$SST = 113$$

	25	9	0	0
	1	1	1	4
	1	36	9	0
	9	0	1	25
$\Sigma x^2 =$	36	46	11	29

$$SSL = \frac{(\Sigma A_i)^2}{L^3} - \frac{T^2}{n}$$

$$SSL = \left.\begin{matrix} ((0 + (-1) + (-6) + (-5))^2 + \\ ((3 + (-2) + (-1) + (1)^2) + \\ (5 + 1 + 0 + 0)^2 + \\ ((0 + (-1)(-3) + (-3))^2 \end{matrix}\right\} - 4 - \frac{(-2^2)}{16}$$

$$SSL = \frac{61}{2} - \frac{(-12)^2}{16}$$

$$SSL = 48.5$$

$$SSE = SST - (SSR + SSC + SSL)$$

$$= 113 - (46.5 + 7.5 + 48.5)$$

$$SSE = 10.5$$

$$MSR = \frac{SSR}{r-1} = \frac{46.5}{4-1} = 15.5$$

$$MSC = \frac{SSC}{C-1} = \frac{48.5}{4-1} = 16.17$$

$$MSE = \frac{SSE}{(C-1)(C-2)} = \frac{10.5}{(4-1)(4-2)} = 1.75.$$

$$F_R = \frac{MSR}{MSE} = \frac{15.5}{1.75} = 8.86.$$

$$F_C = \frac{MSC}{MSE} = \frac{2.5}{1.75} = 1.43$$

$$F_L = \frac{MSL}{MSE} = \frac{16.17}{1.75} = 9.24$$

Source of Variation	Sum of Squares	d.f	mean Square	F. Ratio.
B/W Row.	SSR= 46.5	r-1= 3	MSR = 15.5	F_R = 8.86 > 4.76 at F(3,6)
B/W column	SSC = 7.5	c-1= 3	MSC = 2.5	F_C = 1.43 ≯ 4.76 at F(3,6)
B/W letters	SSL = 48.5	l-1= 3	MSL = 16.17	F_L = 9.24 > 4.76 at F(3.6)
Errors	SSE = 10.5	(C-1)(C-2) = 6	MSE = 1.75.	
Total	SST = 113			

∴ we accept hypothesis along column and reject hypothesis between rows and letters

10 a) Set up an analysis of variance table for the following per acre production data for three varieties of wheat, each grown on 4 plots and state if the variety differences are significant at 5% significant level.

Plot of land	per acre production data variety of wheat		
	A	B	C
1	6	5	5
2	7	5	4
3	3	3	3
4	8	7	4

A	B	C	A^2	B^2	C^2	T	T^2
6	5	5	36	25	25	16	256
7	5	4	49	25	16	16	256
3	3	3	9	9	9	9	81
8	7	4	64	49	16	19	361
24	20	16				60	
576	400	256					

$T = 24 + 20 + 16 = 60$.

Grand Total $= 332$.

$N = 12$

$H_0 : \mu_1 = \mu_2 = \mu_3$

$$C.F = \frac{T^2}{N} = \frac{(60)^2}{12} = 300$$

Sum of squre $=$ Grand Total $- C.F$

$= 332 - 300$

Sum of square $= 32$

B/W Rows : SSR $\Rightarrow \frac{256}{3} + \frac{256}{3} + \frac{81}{3} + \frac{361}{3} - 300$

$SSR = 318 - 300$

$SSR = 18$

B/W columns : $SSC = \frac{576}{4} + \frac{400}{4} + \frac{256}{4} - 300$

$SSC = 308 - 300$

$SSC = 8$

SSE : sum of squares - SSR - SSC

$32 - 18 - 8 = 8$

$SSE = 8$.

Sources variation	d.f	S S	MSS	F-Ratio
Rows	$4-1=3$	SSR=18	MSR=6	$F_r = \frac{6}{1} = 6$
columns	$3-1=2$	SSC=8	MSC=4	$F_c = \frac{4}{1} = 4$.
Error	$3\times2=6$	SSE=6	MSE=1	
Total	$12-1=11$			

$F_r = 6 > F(3,6) = 4.76$

$F_c = 4 < F(6,2) = 19.33$.

10 b) Set up ANOVA table for the following informations relating to three drugs testing to judge the effectiveness in reducing blood pressure for three different groups of people.

Groups of people	Drug		
	X	Y	Z
A	14 15	10 9	11 11
B	12 11	7 8	10 11
C	10 11	11 11	8 7

a) Do the drugs act differently
b) Are the different groups of people affected differently?
c) Is the interaction term significant?
Answer the above questions taking a significant level of 5%.

⇒

Groups of people	Drug X	Drug Y	Drug Z	T	T^2
A	14 15	10 9	11 11	70	4900
B	12 11	7 8	10 11	59	3481
C	10 11	11 11	8 7	58	3364
P	73	56	58	187	
P^2	5329	3136	3364		

$N = 6 + 6 + 6 = 18$

$$C.F = \frac{T^2}{N} = \frac{(187)^2}{18} = \frac{34969}{18} = 1942.722$$

Sum of square.

	X	Y	Z	Sum of squares
A	196 225	100 81	121 121	844
B	144 121	49 64	100 121	599
C	100 121	121 121	64 49	576.
				2019.

TSS = sum of Squares − C.F

TSS = 2019 − 1942.722

TSS = 76.28

$$SSR = \frac{4900}{6} + \frac{3481}{6} + \frac{3364}{6} - 1942.722$$

SSR = 14.78

$$SSC = \frac{5329}{6} + \frac{3136}{6} + \frac{3364}{6} - 1942.722$$

SSC = 28.77

$$SST = \{(14-14.5)^2 + (15-14.5)^2 + (10-9.5)^2 + (9-9.5)^2 + (11-11)^2 + (11-11)^2 +$$
$$(12-11.5)^2 + (11-11.5)^2 + (7-7.5)^2 + (8-7.5)^2 + (10-10.5)^2 +$$
$$(11-10.5)^2 + (10-10.5)^2 + (11-10.5)^2 + (11-11)^2 + (11-11)^2 +$$
$$(8-7.5)^2 + (7-7.5)^2\}$$

$SST = 3.50$

$SSE = TSS - SSR - SSC - SST$

$SSE = 76.28 - 14.78 - 28.77 - 3.50$

$SSE = 29.23$

$F(2,9) = 4.26$

$F(4,9) = 3.63$

Sources of variation.	d.f.	SS	MSS	F-Ratio
Rows.	$3-1=2$	14.78	$MSR = \frac{14.78}{2}$ $MSR = 7.39$	$F_r = \frac{7.39}{0.389}$ $F_r = 19$
columns	$3-1=2$	28.77	$MSC = \frac{28.77}{2}$ $MSC = 14.385$	$F_c = \frac{14.385}{0.389}$ $F_c = 37$
Treatments	9	3.5	$MST = \frac{3.5}{9}$ $MST = 0.389$	$F_T = \frac{7.33}{0.389}$ $F_T = 18.84$
Error	4	29.33	$MSE = \frac{29.33}{4}$ $MSE = 7.33$	

$F_r < F(2,9) \rightarrow$ rejected.

$F_C < F(2,9) \rightarrow$ rejected.

$F_T \leq F(4,9) \rightarrow$ rejected.