II Jai Sri Gurudev II
Sri Adichunchanagiri Shikshana Trust (R)
SJC Institute of Technology
www.sjcit.ac.in Chickballapur - 562 101, Karnataka Estd: 1986

DEPARTMENT OF MATHEMATICS

MATHEMATICS-3 FOR COMPUTER SCIENCE STREAM (BCS301)

MODEL QUESTION PAPER SOLUTIONS

Prepared by:
Purushotham P
Assistant Professor
SJC Institute of Technology
Email id: purushotham@sjcit.ac.in
Phone: 7338005244

Model Paper – 1 Solutions

Module – 1

1.

a) A shipment of 8 similar microcomputers to retail outlet contains 3 that are defective. If 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.

Sol^n:

Given 8-3=5 working micro computers

So , there are $8_{C2} = 28$ ways to select 2 computers.

There are $3_{C0} \times 5_{C2} = 10$ ways to select zero defective and 2 working computers.

The probability that to select zero defective computers $P(0) = \frac{10}{28} = 0.3571$

There are $3_{C1} \times 5_{C1} = 15$ ways to select ONE defective and ONE working computers.

The probability that to select one defective computers $P(1) = \frac{15}{28} = 0.5357$

There are $3_{C2} \times 5_{C0} = 3$ ways to select 2 defective and zero working computers.

The probability that to select two defective computers $P(2) = \frac{3}{28} = 0.1071$

No. of defective computers	0	1	2
Probability P(x)	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

x	$P(x)$	$xP(x)$	x^2	$x^2P(x)$
0	0.3571	0	0	0
1	0.5357	0.5357	1	0.5357
2	0.1071	0.2142	4	0.4284
$\sum$	-	0.75	-	0.9641

Mean $\mu = E(X) = \sum xP(x) = 0.75$

Variance $\sigma^2 = E(X^2) - \mu^2 = \sum x^2 P(x) - \mu^2 = 0.9641 - (0.75)^2$

$\Rightarrow \sigma^2 = 0.9641 - 0.5625 == 0.4016$

$\Rightarrow S.D. = \sigma = \sqrt{0.4016} = 0.6337$

b) In a certain factory turning out razor blades there is a small probability of $\frac{1}{500}$ for any blade to be defective. The blades are supplied in a packet of 10. Use poisson distribution to calculate approximate number of packets containing,
i)No defective
ii)2 defective
iii)3 defective
in the consignment of 10000 packets.

Sol^n:

Let X be the poisson variant follows the blades to be defective of the poisson distribution.

The probability mass function of the poisson distribution is $P(X=x)=\frac{e^{-\lambda}\lambda^x}{x!}$

Given, p=$\frac{1}{500}=0.002$, n=10, $\mu=np=0.002\times 10=0.02=\lambda$

$$\therefore P(X=x)=\frac{e^{-0.02}(0.02)^x}{x!}$$

i) No blades are defective out 10000 packets $=10000\times P(x=0)$

$$=10000\times\frac{e^{-0.02}(0.02)^0}{0!}$$
$$=10000\times 0.9802$$
$$=9802$$

$\therefore$9802 packet blades are not defective out of 10000 packets.

ii) 2 defective blades out of 10000 packets $=10000\times P(x=2)$

$$=10000\times\frac{e^{-0.02}(0.02)^2}{2!}$$
$$=10000\times 0.0002$$
$$=2$$

$\therefore$2 packets blades are 2 defectives out of 10000 packets.

iii) 3 defective blades out of 10000 packets $=10000\times P(x=3)$

$$=10000\times\frac{e^{-0.02}(0.02)^3}{3!}$$
$$=10000\times 0.0000$$
$$=0$$

$\therefore$No packets blades are 3 defectives out of 10000 packets.

c) If the mileage (in thousands of miles) of a certain radial tyre is a random variable with exponential distribution with mean 40000 miles. Determine the probability that the tyre will last
i) At least 20000 km
ii) At most 30000 km

Sol^n:

Given X be a continuous random variable of an exponential distribution is,

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

and given the mean of exponential distribution is 40000.

$\Rightarrow \mu=40000\Rightarrow\frac{1}{\alpha}=40000\Rightarrow\alpha=\frac{1}{40000}$

$$\therefore f(x)=\begin{cases}\frac{1}{40000}e^{-\frac{x}{40000}} & ,x\geq 0\\ 0 & ,otherwise\end{cases}$$

i) The probability that the radial tyre will last at least 20000 km is,

$$P(x\geq 20000)=\int_{20000}^{\infty}f(x)dx=\int_{20000}^{\infty}\frac{1}{40000}e^{-\frac{x}{40000}}dx$$
$$=\frac{1}{40000}\int_{20000}^{\infty}e^{-\frac{x}{40000}}dx$$

$$= -\left[e^{-\frac{x}{40000}}\right]_{20000}^{\infty} = -\left[0 - e^{-\frac{20000}{40000}}\right] = \frac{1}{\sqrt{e}}$$

ii) The probability that the radial tyre will last at most 30000 km is,

$$P(x \le 30000) = \int_0^{30000} f(x)dx = \int_0^{30000} \frac{1}{40000} e^{-\frac{x}{40000}} dx$$
$$= \frac{1}{40000}\int_0^{30000} e^{-\frac{x}{40000}} dx$$
$$= -\left[e^{-\frac{x}{40000}}\right]_0^{30000} = -\left[e^{-\frac{30000}{40000}} - e^0\right] = 1 - \frac{1}{e^{\frac{3}{4}}}$$

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

$$f(x) = \begin{cases} k\sqrt{x}, & 0 < x < 1 \\ 0, & elsewhere \end{cases},$$

i) Find k

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

Sol^n:

Given probability function, $f(x) = \begin{cases} k\sqrt{x}, & 0 < x < 1 \\ 0, & elsewhere \end{cases}$

$$WKT, \int_{-\infty}^{\infty} f(x)dx = 1$$
$$\Rightarrow i) \int_{-\infty}^{0} f(x)dx + \int_0^1 f(x)dx + \int_1^{\infty} f(x)dx = 1$$
$$\Rightarrow 0 + \int_0^1 k\sqrt{x}\, dx + 0 = 1$$
$$\Rightarrow k\int_0^1 \sqrt{x}\, dx = 1$$
$$\Rightarrow \frac{2k}{3}\left[x^{\frac{3}{2}}\right]_0^1 = 1$$
$$\Rightarrow \frac{2k}{3} = 1$$
$$\Rightarrow 2k = 3$$
$$\Rightarrow k = \frac{3}{2}$$

WKT, the cdf of the pdf is,

$$F(x) = \int_{-\infty}^{x} f(u)du$$
$$= \int_{-\infty}^{0} f(u)du + \int_0^x f(u)du$$
$$= 0 + \int_0^x k\sqrt{u}\, du$$
$$= \frac{3}{2}\times\frac{2}{3}\left[u^{\frac{3}{2}}\right]_0^x$$
$$= x^{\frac{3}{2}} = x\sqrt{x}$$

$$P[0.3 < X < 0.6] = F(0.6) - F(0.3)$$
$$= 0.6(\sqrt{0.6}) - 0.3(\sqrt{0.3})$$
$$= 0.4647 - 0.1643$$

$$= 0.3004$$

b) Find the mean and variance of Binomial Distribution.

Sol^n:

Let X be a discrete random variable, 'p' be the probability of success and let 'q' be the probability of failure, then the probability mass function of the binomial distribution can be defined as,

$$P(X = x) = b(n, p, x) = \begin{cases} n_{c_x} p^x q^{n-x} & , x \geq 0 \\ 0 & , Otherwise \end{cases}$$

WKT, the probability mass function of the binomial distribution is,

$$P(X = x) = f(x) = \begin{cases} n_{c_x} p^x q^{n-x}, & x \geq 0 \\ 0 \quad , & otherwise \end{cases}$$

i) Mean:

$$\mu = E(x) = \sum_{x=0}^{n} xP(X = x)$$

$$= \sum_{x=0}^{n} x\, n_{c_x} p^x q^{n-x}$$

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

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

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

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

$$= np \sum_{x=1}^{n} (n-1)_{c_{(x-1)}} p^{x-1} q^{(n-1)-(x-1)}$$

$$= np(1)$$

$$\mu = E(x) = np$$

i) Variance:

$$\sigma^2 = E(x^2) - [E(x)]^2 \; ----(1)$$

$$\Rightarrow E(x^2) = E(x(x-1) + x)$$

$$\Rightarrow E(x^2) = E\big(x(x-1)\big) + E(x) ----(2)$$

$$\therefore E(x(x-1)) = \sum_{x=0}^{n} x(x-1)p(x)$$

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

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

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

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

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

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

$$= n(n-1)p^2 \sum_{x=2}^{n} (n-2)_{c_{(x-2)}} p^{x-2} q^{(n-2)-(x-2)}$$

$$= n(n-1)p^2(1)$$

$$\therefore E(x(x-1)) = n(n-1)p^2$$

$$(2) \Rightarrow E(x^2) = E\big(x(x-1)\big) + E(x)$$

$$\Rightarrow E(x^2) = n(n-1)p^2 + np$$

$$(1) \Rightarrow \sigma^2 = E(x^2) - [E(x)]^2$$

$$\Rightarrow \sigma^2 = n(n-1)p^2 + np - [np]^2$$

$$\Rightarrow \sigma^2 = n^2p^2 - np^2 + np - n^2p^2$$

$$\Rightarrow \sigma^2 = np - np^2$$

$$\Rightarrow \sigma^2 = np(1-p)$$

$$but \quad 1 - p = q$$

$$\therefore \sigma^2 = npq$$

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 of 2040 hours and SD of 60 hours. Estimate the number of bulbs likely to burn for...
i) More than 2150 hours
ii) Less than 1950 hours
iii) Between 1920 and 2160 hours.

Sol^n:

Let X be the continuous random variable
Given,
Mean of the Normal distribution $\mu = 2040$
Standard deviation of the Normal distribution $\sigma = 60$
$\therefore$The standard normal variate $z = \frac{x-\mu}{\sigma} \Rightarrow z = \frac{x-2040}{60}$

When x =2150 then $z = \frac{2150-2040}{60} = 1.83$
When x =1950 then $z = \frac{1950-2040}{60} = -1.5$
When x =1920 then $z = \frac{1920-2040}{60} = -2$
When x =2160 then $z = \frac{2160-2040}{60} = 2$

i) The probability that the number of bulbs likely to burn of more than 2150 hours:

$$\begin{aligned} P(x > 2150) &= P(z > 1.83) \\ &= 0.5 - A(1.83) \\ &= 0.5 - 0.4664 = 0.0336 \end{aligned}$$

The number of bulbs likely to burn of more than 2150 hours out of 2000 bulbs=2000x0.0336
=67.2=67

ii) The probability that the number of bulbs likely to burn of less than 1950 hours:

$$P(x < 1950) = P(z < -1.5)$$
$$= 0.5 - A(1.5)$$
$$= 0.5 - 0.4332 = 0.0668$$

The number of bulbs likely to burn of less than 1950 hours out of 2000 bulbs=2000x0.0668
=133.6=137

iii) The probability that the number of bulbs likely to burn between 1920 and 2160 hours:

$$P(1920 < x < 2160) = P(-2 < z < 2)$$
$$= 2P(0 < z < 2)$$
$$= 2A(2)$$
$$= 2 \times 0.4772$$
$$= 0.9544$$

The number of bulbs likely to burn between 1920 and 2160 hours out of 2000 bulbs=2000x0.9544
=1908.8
=1909

Module – 2

3.
a) The joint distribution of two random variables X and Y are as follows:

X \ Y	-4	2	7
1	$^1/_8$	$^1/_4$	$^1/_8$
5	$^1/_4$	$^1/_8$	$^1/_8$

Compute the following,
i) E(X) and E(Y)
ii) E(XY)
iii) σ_X & σ_Y
iv) $\rho(X,Y)$

Sol^n: Given,

$$x_1 = 1,\ x_2 = 5,\ y_1 = -4,\ y_2 = 2,\ y_3 = 7$$

And the probabilities are

$$p_{11} = \frac{1}{8},\ p_{12} = \frac{1}{4},\ p_{13} = \frac{1}{8},\ p_{21} = \frac{1}{4},\ p_{22} = \frac{1}{8},\ p_{23} = \frac{1}{8}$$

Given the joint probability distribution is follows as

X \ Y	-4	2	7	$f(x_i)$
1	$^1/_8$	$^1/_4$	$^1/_8$	$^1/_2$
5	$^1/_4$	$^1/_8$	$^1/_8$	$^1/_2$
$g(y_i)$	$^3/_8$	$^3/_8$	$^1/_4$	1

The marginal distribution of X and Y are

x_i	1	5
$f(x_i)$	$^1/_2$	$^1/_2$

y_i	-4	2	7
$g(y_i)$	$3/8$	$3/8$	$1/4$

i) $\mu_X = E(X) = \sum_{i=1}^{2} x_i f(x_i) = \left(1 \times \frac{1}{2}\right) + \left(5 \times \frac{1}{2}\right) = 3$

$\mu_Y = E(Y) = \sum_{j=1}^{3} y_j g(y_j) = \left(-4 \times \frac{3}{8}\right) + \left(2 \times \frac{3}{8}\right) + \left(7 \times \frac{1}{4}\right) = 1$

ii) $E(XY) = \sum_{i=1}^{2} \sum_{j=1}^{3} x_i y_j f(x_i, y_j)$

$$= \left(1 \times (-4) \times \frac{1}{8}\right) + \left(1 \times 2 \times \frac{1}{4}\right) + \left(1 \times 7 \times \frac{1}{8}\right) + \left(5 \times (-4) \times \frac{1}{4}\right) + \left(5 \times 2 \times \frac{1}{8}\right) + \left(5 \times 7 \times \frac{1}{8}\right)$$
$$= \frac{3}{2}$$

iii) $\sigma_X{}^2 = E(X^2) - \mu_X{}^2 = \sum_{i=1}^{2} x_i{}^2 f(x_i) - \mu_X{}^2 = \left(1^2 \times \frac{1}{2}\right) + \left(5^2 \times \frac{1}{2}\right) - 9 = 13 - 9 = 4 \Rightarrow \sigma_X = 2$

$\sigma_Y{}^2 = E(Y^2) - \mu_Y{}^2 = \sum_{j=1}^{3} y_j{}^2 g(y_j) - \mu_Y{}^2 = \left((-4)^2 \times \frac{3}{8}\right) + \left(2^2 \times \frac{3}{8}\right) + \left(7^2 \times \frac{1}{4}\right) - 1^2 = \frac{75}{4} \Rightarrow \sigma_Y = 4.33$

iv) $COV(X, Y) = E(XY) - \mu_X \mu_Y = \frac{3}{2} - (3)(1) = -\frac{3}{2}$

v) $\rho(X, Y) = \frac{COV(X,Y)}{\sigma_X \sigma_Y} = \frac{-3/2}{2 \times 4.33} = -0.1732$

Hence the given random variables are not independent

b) Find the unique fixed probability vector of $\mathbf{P} = \begin{bmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

Sol^n:

Given, $P = \begin{bmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{bmatrix}$

Since the given matrix P is of order 3x3, the required fixed probability vector Q must be also order of 3x3.

Let $Q = [x \quad y \quad z] For every x \geq 0, y \geq 0, z \geq 0 \& x + y + z = 1$

Also, QP=Q

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

$$\Rightarrow QP = \left[\frac{1}{2}y \quad \frac{3}{4}x + \frac{1}{2}y + z \quad \frac{1}{4}x\right]$$

WKT

$QP = Q$

$\Rightarrow \left[\frac{1}{2}y \quad \frac{3}{4}x + \frac{1}{2}y + z \quad \frac{1}{4}x\right] = [x \quad y \quad z]$

$\Rightarrow x = \frac{1}{2}y, y = \frac{3}{4}x + \frac{1}{2}y + z, z = \frac{1}{4}x$

After solving the above equations. we get,

$\Rightarrow 2x + y = 0, x + 6y = 4, 3x + 4y = -4$ where z=1-x-y

On solving above equations. we get,

$\Rightarrow x = \frac{-20}{7}, y = \frac{8}{7}, z = \frac{19}{7}$

$\therefore Q[x \quad y \quad z] = \left[\frac{-20}{7} \quad \frac{8}{7} \quad \frac{19}{7}\right]$

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 an Ambassador, he trades it for Santro. However, if he had a Santro, he is just as likely to trade it for a

Maruthi or an Ambassador. In 2000 he bought his first car which was a Santro. Find the probability that he has,

i) 2002 Santro

ii) 2002 Maruthi

iii) 2003 Ambassador

iv) 2003 Santro

Sol^n:

Given a man trades his car for a new car with the probabilities as below,

$$P = \begin{array}{c} M \\ A \\ S \end{array}\begin{bmatrix} p_{MM}{}^{(1)} & p_{MA}{}^{(1)} & p_{MS}{}^{(1)} \\ p_{AM}{}^{(1)} & p_{AA}{}^{(1)} & p_{AS}{}^{(1)} \\ p_{SM}{}^{(1)} & p_{SA}{}^{(1)} & p_{SS}{}^{(1)} \end{bmatrix} = \begin{array}{c} M \\ A \\ S \end{array}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

Also given, he has bought his first car in 2000 was Santro.

$\therefore$ The initial probability vector $p^{(0)} = [p_M^{(0)} \quad p_A^{(0)} \quad p_S^{(0)}] = [0 \quad 0 \quad 1]$

$$\therefore P^2 = P.P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix}.\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

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

$$\therefore p^{(2)} = p^{(0)}P^2 = [0 \quad 0 \quad 1].\begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{bmatrix} = [0 \quad 1/2 \quad 1/2] = [p_M^{(2)} \quad p_A^{(2)} \quad p_S^{(2)}]$$

$$\therefore p^{(3)} = p^{(0)}P^3 = [0 \quad 0 \quad 1].\begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{bmatrix} = [1/4 \quad 1/4 \quad 1/2] = [p_M^{(3)} \quad p_A^{(3)} \quad p_S^{(3)}]$$

i) $\therefore$ The probability to have a Santro car in the year 2002, $p_S^{(2)} = 1/2 = 50\%$

ii) $\therefore$ The probability to have a Maruthi car in the year 2002,$p_M^{(2)} = 0 = 0\%$

iii) $\therefore$ The probability to have an Ambassador car in the year 2003,$p_A^{(3)} = 1/4 = 25\%$

$\therefore$ The probability to have a Santro car in the year 2003,$p_S^{(3)} = 1/2 = 50\%$

4.

a) The joint probability distribution of two 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 independent?

ii) Evaluate $P[Y \leq 2]$

iii) Evaluate $P[X + Y \leq 2]$

Sol^n:

Given,

$x_1 = 1,\ x_2 = 3,\ y_1 = -3,\ y_2 = 2,\ y_3 = 4$

And the probabilities are

$p_{11} = 0.1,\ p_{12} = 0.2,\ p_{13} = 0.2,\ p_{21} = 0.3,\ p_{22} = 0.1,\ p_{23} = 0.1$

Given the joint probability distribution is follows as

Y \ X	-3	2	4	$f(x_i)$
1	0.1	0.2	0.2	0.5
3	0.3	0.1	0.1	0.5
$g(y_i)$	0.4	0.3	0.3	1

The marginal distribution of X and Y are

x_i	1	3
$f(x_i)$	0.5	0.5

y_i	-3	2	4
$g(y_i)$	0.4	0.3	0.3

$$\mu_X = E(X) = \sum_{i=1}^{2} x_i f(x_i) = (1 \times 0.5) + (3 \times 0.5) = 2$$
$$\mu_Y = E(Y) = \sum_{j=1}^{3} y_j g(y_j) = (-3 \times 0.4) + (2 \times 0.3) + (4 \times 0.3) = 0.6$$

$$E(XY) = \sum_{i=1}^{2} \sum_{j=1}^{3} x_i y_j f(x_i, y_j)$$
$$= (1 \times (-3) \times 0.1) + (1 \times 2 \times 0.2) + (1 \times 4 \times 0.2) + (3 \times (-3) \times 0.3) + (3 \times 2 \times 0.1) + (3 \times 4 \times 0.1)$$
$$= -0.3 + 0.4 + 0.8 - 2.7 + 0.6 + 1.2 = 0$$

$$COV(X,Y) = E(XY) - \mu_X \mu_Y = 0 - (2)(0.6) = 0 - 1.2 = -1.2$$

Hence the given random variables are not independent

$$\therefore P(Y \le 2) = p_{11} + p_{12} + p_{21} + p_{22} = 0.1 + 0.2 + 0.3 + 0.1 = 0.7$$

$$\therefore P(X + Y \le 2) = p_{11} + p_{21} = 0.1 + 0.3 = 0.4$$

b) Define probability vectors, Stochastic matrices, Regular Stochastic matrix, Stationary Distribution and Absorbing state of Markov Chain.

Sol^n:

Probability Vector

A vector $V = [v_1, v_2, v_3, \ldots\ldots, v_n]$ is called the probability vector if each one of its components are non-negative and their sum is equal to unity or 1.

Ex: $= [0.1, 0.6, 0.3]$, $V = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right]$, etc…

Stochastic Matrix

A square matrix P is called a stochastic matrix if all the entries of P are non-negative and the sum of all the entries of any row is 1

(or)

A square matrix P is called a stochastic matrix where each row is in the form of the probability vector.

Ex: $= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$, $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$

Regular Stochastic Matrix

A matrix P is said to be a Regular Stochastic Matrix, if all the entries of some power (P^n) are positive. The Regular Stochastic Matrix P has a unique probability vector Q such that QP=Q and all the sum of the probabilities of a fixed vector matrix should be equal to 1.

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2n} \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & p_{n3} & \cdots & p_{nn} \end{bmatrix}$$

c) A Salesman's territory consists of three cities A, B, 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. If he sells in B or C then the next day is twice as likely to sell in city A as than other cities. In long run, how often does he sells in each of the city.

Sol^n:

Given a salesman can move to the cities A, B, C with the probabilities as below,

$$P = \begin{array}{c} A \\ B \\ C \end{array}\begin{bmatrix} p_{AA}{}^{(1)} & p_{AB}{}^{(1)} & p_{AC}{}^{(1)} \\ p_{BA}{}^{(1)} & p_{BB}{}^{(1)} & p_{BC}{}^{(1)} \\ p_{CA}{}^{(1)} & p_{CB}{}^{(1)} & p_{CC}{}^{(1)} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 2/3 & 0 & 1/3 \\ 2/3 & 1/3 & 0 \end{bmatrix}$$

Let $Q = [x \quad y \quad z]$ be the probability vector for which x+y+z=1

$$\therefore QP = Q$$

$$\therefore [x \quad y \quad z].\begin{bmatrix} 0 & 1 & 0 \\ 2/3 & 0 & 1/3 \\ 2/3 & 1/3 & 0 \end{bmatrix} = [x \quad y \quad z]$$

$$\Rightarrow \left[\frac{2y}{3} + \frac{2z}{3} \quad x + \frac{z}{3} \quad \frac{y}{3}\right] = [x \quad y \quad z]$$

$$\Rightarrow \frac{2y}{3} + \frac{2z}{3} = x \;,\; x + \frac{z}{3} = y \;,\; \frac{y}{3} = z$$

$$\Rightarrow 3x - 2y - 2z = 0 \;,\; 3x - 3y + z = 0$$

$$\Rightarrow 3x - 2y - 2(1 - x - y) = 0 \;,\; 3x - 3y + (1 - x - y) = 0$$

$$\Rightarrow 3x - 2y - 2 + 2x + 2y = 0 \;,\; 3x - 3y + 1 - x - y = 0$$

$$\Rightarrow 5x = 2 \;,\; 2x - 4y = -1$$

$$\Rightarrow x = \frac{2}{5}$$

$$\Rightarrow 4y = \frac{9}{5} \Rightarrow y = \frac{9}{20}$$

$$\Rightarrow z = 1 - x - y \Rightarrow z = 1 - \frac{2}{5} - \frac{9}{20} \Rightarrow z = \frac{3}{20}$$

$$\therefore Q = [x \quad y \quad z] = \left[\frac{2}{5} \quad \frac{9}{20} \quad \frac{3}{20}\right]$$

Thus, the salesman in the long run sells,

$\frac{2}{5}$ $in\ city\ A = 40\%$, $\frac{9}{20}$ $in\ city\ B = 45\%$, $\frac{3}{20}$ $in\ city\ C = 15\%$

Module – 3

5.

a) Define Null Hypothesis, Significance Level, Critical Region, Type-I and Type-II errors in a statistical test.

Sol^n:

Null Hypothesis:

The **null hypothesis** is a general statement or default position that there is no relationship between two measured phenomena or no association among groups.

Example: Given the test scores of two random samples, one of men and one of women, does one group differ from the other? A possible null hypothesis is that the mean male score is the same as the mean female score:

H_0: $\mu_1 = \mu_2$

where

H_0 = the null hypothesis,

μ_1 = the mean of population 1, and

μ_2 = the mean of population 2.

A stronger null hypothesis is that the two samples are drawn from the same population, such that the variances and shapes of the distributions are also equal.

Significance levels (α):

The significance level of an event (such as a statistical test) is the probability that the event could have occurred by chance. If the level is quite low, that is, the probability of occurring by chance is quite small, we say the event is significant.

The level of significance is the measurement of the statistical significance. It defines whether the null hypothesis is assumed to be accepted or rejected. It is expected to identify if the result is statistically significant for the null hypothesis to be false or rejected.

$$\alpha = 5\% \quad \alpha = 1\% \quad \alpha = 0.27\%$$

Example: A level of significance of p=0.05 means that there is a 95% probability that the results found in the study are the result of a true relationship/difference between groups being compared. It also means that there is a 5% chance that the results were found by chance alone and no true relationship exists between groups.

Critical Region:

A critical region, also known as the rejection region, is a set of values for the test statistic for which the null hypothesis is rejected. i.e. if the observed test statistic is in the critical region then we reject the null hypothesis and accept the alternative hypothesis.

Type I and Type II Errors:

When we test a statistic at specified confidence level, there are chances of taking wrong decisions due to small sample size or sampling fluctuations etc.

Type I error is the incorrect rejection of a true null hypothesis, i.e. we reject H_0, when itis true, whereas Type II error is the incorrect acceptance of a false null hypothesis, i.e. we accept H_0 when it is false.

b) A coin is tossed 400 times and it turns up head 216 times. Discuss whether the coin may be regarded as unbiased one.

Sol^n:

Set the null hypothesis $H_0; P = \frac{1}{2}$

Set the Alternative hypothesis $H_1: P \neq \frac{1}{2}$

The level of significance $\alpha = 0.05$ (5%)

$\therefore$ The test statistic $Z = \frac{p-P}{\sqrt{\frac{PQ}{n}}}$, where P+Q=1 => Q=1-P

Given, the coin is tossed and it turns up in the equal proportion

$$P = \frac{1}{2} \Rightarrow Q = 1 - P$$

$$\Rightarrow Q = 1 - \frac{1}{2} = \frac{1}{2}$$

And the coin turns up head 216 times when it tossed $n = 400$times

$$\therefore p = \frac{216}{400} = 0.54$$

$$\therefore Z = \frac{0.54 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{400}}}$$

$$\Rightarrow Z = \frac{0.04}{\sqrt{0.000625}} = 1.6$$

At 5% level, the tabulated value of $Z_\alpha$is 1.96

Since $|Z| = 1.6 < 1.96$

Hence, the null hypothesis is accepted at 5% level of significance and the coin may be regarded as unbiased.

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

Sol^n:

Set the null hypothesis $H_0; P_1 = P_2$

Set the Alternative hypothesis $H_1: P_1 \neq P_2$

The level of significance $\alpha = 0.05$ (5%)

Given

$$n_1 = 900, n_2 = 1600$$

$x_1 = 20\%$of random sample of 900=0.2x900=180

$x_2 = 18.5\%$of random sample of 1600=0.185x1600=296

$$\therefore p_1 = 20\% = \frac{20}{100} = 0.2, p_2 = 18.5\% = \frac{18.5}{100} = 0.185$$

We know that $P = \frac{x_1+x_2}{n_1+n_2}$

$$\Rightarrow P = \frac{180 + 296}{900 + 1600}$$
$$\Rightarrow P = \frac{476}{2500}$$
$$\Rightarrow P = 0.1904 \Rightarrow Q = 1 - P = 1 - 0.1904 = 0.8096$$
$$\therefore Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$
$$\Rightarrow Z = \frac{0.2 - 0.185}{\sqrt{\left((0.1904 \times 0.8096)\left(\frac{1}{900} + \frac{1}{1600}\right)\right)}}$$
$$\Rightarrow Z = \frac{0.015}{\sqrt{(0.1541)(0.00173)}}$$
$$\Rightarrow Z = \frac{0.015}{\sqrt{0.00026}}$$
$$\Rightarrow Z = \frac{0.015}{0.01612}$$
$$\Rightarrow Z = 0.9305$$

At 5% level, the tabulated value of Z_α is 1.96

Since $|Z| = 0.9305 < 1.96$

Hence, the null hypothesis H_0 is accepted at 5% level of significance and hence there is no significant difference.

6.

a) Explain the following terms:

i) Standard Error

ii) Statistical Hypothesis

iii) Critical Region of a statistical test

iv) Test of Significance

Sol^n:

Standard Error:

The standard deviation of the sampling distribution of a statistic is Known as Standard Error (S.E.).

Null Hypothesis (or) Statistical Hypothesis:

The **null hypothesis** is a general statement or default position that there is no relationship between two measured phenomena or no association among groups.

Example: Given the test scores of two random samples, one of men and one of women, does one group differ from the other? A possible null hypothesis is that the mean male score is the same as the mean female score:

H_0: $\mu_1 = \mu_2$

where

H_0 = the null hypothesis,

μ_1 = the mean of population 1, and

μ_2 = the mean of population 2.

A stronger null hypothesis is that the two samples are drawn from the same population, such that the variances and shapes of the distributions are also equal.

Critical Region:

A critical region, also known as the rejection region, is a set of values for the test statistic for which the null hypothesis is rejected. i.e. if the observed test statistic is in the critical region then we reject the null hypothesis and accept the alternative hypothesis.

Test of Significance (or) Test of Hypothesis:

Let x be the observed number of successes in a sample size of n and $\mu = np$ be the expected number of successes .Then the standard normal variate Z is defined as

$$Z = \frac{x-\mu}{\sigma} = \frac{x-np}{\sqrt{npq}}$$

b) A die is thrown 9000 times and a throw of 3 or 4 was observed 3240 times. Show that the die Can not be regarded as an unbiased one.

Sol^n:

The probability of getting 3 or 4 in a single through is $p = {}^2/_6 = {}^1/_3$

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

$\therefore$ Expected number of success $= \frac{1}{3} \times 9000 = 3000$

$\therefore$ The difference =3240-3000=240

$$Z = \frac{x - np}{\sqrt{npq}}$$

Consider $\Rightarrow Z = \frac{(3240) - \left(9000 \times \frac{1}{3}\right)}{\sqrt{9000 \times \frac{1}{3} \times \frac{2}{3}}}$

$$\Rightarrow Z = \frac{240}{\sqrt{2000}}$$

$$\Rightarrow Z = 5.37$$

Since Z=5.37>2.58 ,

We conclude that 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.

(Warning: Smoking is injurious to health, causes cancer, Tabaco causes painful death)

Sol^n:

Set the null hypothesis $H_0; P_1 = P_2$

Set the Alternative hypothesis $H_1: P_1 \neq P_2$

The level of significance $\alpha = 0.05\ (5\%)$

Given

$$n_1 = 600, n_2 = 900 \ \& \ x_1 = 450, x_2 = 450$$

$$\therefore p_1 = \frac{450}{600} = 0.75, p_2 = \frac{450}{900} = 0.5$$

We know that $P = \frac{x_1 + x_2}{n_1 + n_2}$

$$\Rightarrow P = \frac{450 + 450}{600 + 900}$$

$$\Rightarrow P = \frac{900}{1500} = 0.6$$

$$\Rightarrow P = 0.6 \Rightarrow Q = 1 - P = 1 - 0.6 = 0.4$$

$$\therefore Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

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

$$\Rightarrow Z = \frac{0.25}{\sqrt{(0.24)(0.00277)}}$$

$$\Rightarrow Z = \frac{0.25}{\sqrt{0.0006648}}$$

$$\Rightarrow Z = \frac{0.25}{0.02578}$$

$$\Rightarrow Z = 9.69$$

At 5% level, the tabulated value of Zα is 1.645.

Since |Z| = 9.69 > 1.645

Hence Null Hypothesis H_0 is rejected at 5% level of significance.

Module – 4

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 μ = 53 and variance σ^2 = 400.

Sol^n:

The central limit theorem states that the sample mean $\bar{x}$ follows approximately the normal distribution with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$ (is also called Standard error), i.e., $\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$, where μ , σ are mean and standard deviation of the population from where the sample.

Given,
Sample size n=100
Mean of the population $\mu = 53$
Variance of the population $\sigma^2 = 400 \Rightarrow \sigma = \sqrt{400} = 20$

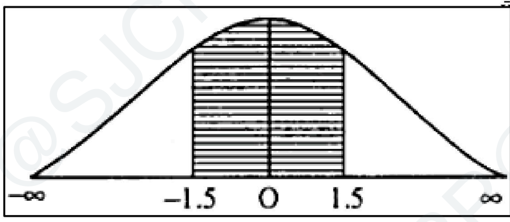
$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

$$\Rightarrow \bar{X} \sim N\left(53, \frac{20}{\sqrt{100}}\right)$$

$$\Rightarrow \bar{X} \sim N(53,2)$$

$$\therefore \text{ we know that } Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$

$$\Rightarrow Z = \frac{\bar{X}-53}{20/\sqrt{100}}$$



$$\Rightarrow Z = \frac{\bar{X}-53}{2}$$

$\therefore$ At $\bar{X}$=50 $\Rightarrow Z = \frac{50-53}{2} = -\frac{3}{2} = -1.5 = z_1$

At $\bar{X}$=56 $\Rightarrow Z = \frac{56-53}{2} = \frac{3}{2} = 1.5 = z_2$

$$\therefore P(50 < \bar{X} < 56) = P(-1.5 < z < 1.5)$$
$$= 2P(0 < z < 1.5)$$
$$= 2A(1.5)$$
$$= 2 \times 0.4332$$

$$\therefore P(50 < \bar{X} < 56) = 0.8664$$

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

Sol^n:

Given the sample size n=25
Mean of sample $\bar{X} = 78.3$
Standard deviation $\sigma = 2$
We know, Confidence level of 99%, the corresponding z value is 2.58. This is determined from the normal distribution table.

Confidence interval $C.I. = \mu = Mean \pm Z(S\tan d\,ardDeviation/\sqrt{SampleSize} = \bar{X} \pm Z\frac{\sigma}{\sqrt{n}}$

$$\therefore C.I. = \mu = 78.3 \pm \left(2.58 \times \frac{2}{\sqrt{25}}\right)$$
$$\Rightarrow \mu = 78.3 \pm 1.032$$
$$\Rightarrow C.I \Rightarrow (78.3 - 1.032\,, 78.3 + 1.032) = (77.268\,, 79.332)$$

c) A survey of 320 families with 5 children each revealed 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?

Sol^n:

Given,
Number of families selected for the survey = 320
The probability of female and male birth is equal, $p = \frac{1}{2} = 0.5 \Rightarrow q = 1 - p = 1 - 0.5 = 0.5$
Number of children in the selected families, n = 5

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

The statistical hypothesis is,
H_0: The probability of female and male birth is equal.
H_1: The probability of female and male birth is not equal.
Here Chi square distribution is used to test the hypothesis.
Therefore, by the Binomial distribution.
We have,

$$P(x) = nC_x p^x q^{n-x}$$
$$\therefore P(x) = 5C_x (0.5)^x (0.5)^{5-x}$$
$$\Rightarrow P(x) = 5C_x (0.5)^5$$

The expected frequencies can be calculated for 320 families as

$$E(x) = 320 \times P(x) = 320 \times 5C_x (0.5)^5$$
$$\therefore E(0) = 320 \times P(0) = 320 \times 5C_0 (0.5)^5 = 320 \times (0.5)^5 = 10 = E_0$$
$$E(1) = 320 \times P(1) = 320 \times 5C_1 (0.5)^5 = 320 \times 5 \times (0.5)^5 = 50 = E_{\backslash 1}$$
$$E(2) = 320 \times P(2) = 320 \times 5C_2 (0.5)^5 = 320 \times 5C_2 \times (0.5)^5 = 100 = E_{\backslash 2}$$
$$E(3) = 320 \times P(3) = 320 \times 5C_3 (0.5)^5 = 320 \times 5C_3 \times (0.5)^5 = 100 = E_{\backslash 3}$$
$$E(4) = 320 \times P(4) = 320 \times 5C_4 (0.5)^5 = 320 \times 5C_4 \times (0.5)^5 = 50 = E_{\backslash 4}$$
$$E(5) = 320 \times P(5) = 320 \times 5C_5 (0.5)^5 = 320 \times 5C_5 \times (0.5)^5 = 10 = E_{\backslash 5}$$

No. of Boys	No. of Girls	Total Observed Frequencies (O_i)	Expected Frequencies(E_i)	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
5	0	14	10	4	16	1.6
4	1	56	50	6	36	0.72
3	2	110	100	10	100	1
2	3	88	100	-12	144	1.44
1	4	40	50	-10	100	2
0	5	12	10	2	4	0.4

We have the Table value of χ^2 for 5 degrees of freedom at level of significance 5% from the chi-square table is 11.07.

$$\therefore \chi^2 = \sum_i \left[\frac{(O_i - E_i)^2}{E_i}\right] \Rightarrow \chi^2 = 7.16 < 11.02$$

Since the calculated χ^2 value is less than tabulated χ^2 value then the decision is fail to reject the H_0 (Accept H_0) that means both the male and female birth is equal.

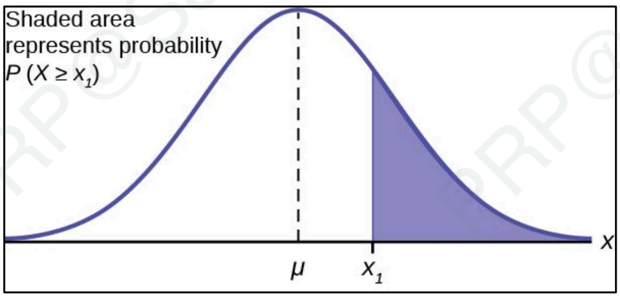
8.

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.

Sol^n:

Given,
Sample size n=64
Mean of the population $\mu = 112$
Variance of the population $\Rightarrow \sigma^2 = 144 \Rightarrow \sigma = 12$

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$
$$\Rightarrow \bar{X} \sim N\left(112, \frac{12}{\sqrt{64}}\right)$$
$$\Rightarrow \bar{X} \sim N(90, 1.5)$$
$$\therefore \text{we know that } Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$
$$\Rightarrow Z = \frac{\bar{X} - 112}{12/\sqrt{64}}$$
$$\Rightarrow Z = \frac{\bar{X} - 112}{1.5}$$
$$\therefore \text{At } \bar{X} = 114.5 \Rightarrow z = \frac{114.5 - 112}{1.5} = 1.66$$



$$\therefore P(\bar{X} > 114.5) = P(z > 1.66)$$
$$\Rightarrow P(z > 1.66) = 0.5 - P(0 < z < 1.66)$$
$$= 0.5 - 0.4515$$
$$\Rightarrow P(z > 1.66) = 0.0489$$

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 95% confidence intervals for μ.

Sol^n:

Given the sample size n=20
Mean of sample $\bar{X} = 81.2$
Variance $\sigma^2 = 80 \Rightarrow \sigma = \sqrt{80} = 8.9442$
We know, Confidence level of 95%, 90% the corresponding z values are 1.96 , 1.645. This is determined from the normal distribution table.

Confidence interval $C.I. = \mu = Mean \pm Z(S\tan d\, ardDeviation/\sqrt{SampleSize} = \bar{X} \pm Z\frac{\sigma}{\sqrt{n}}$

For 95%:

$$\therefore C.I. = \mu = 81.2 \pm \left(1.96 \times \frac{8.9442}{\sqrt{20}}\right)$$
$$\Rightarrow \mu = 81.2 \pm 3.92$$
$$\Rightarrow C.I = (81.2 - 3.92, 81.2 + 3.92) = (77.28, 85.12)$$

For 90%:

$$\therefore C.I. = \mu = 81.2 \pm \left(1.645 \times \frac{8.9442}{\sqrt{20}}\right)$$
$$\Rightarrow \mu = 81.2 \pm 3.29$$
$$\Rightarrow C.I = (81.2 - 3.29, 81.2 + 3.29) = (77.91, 84.49)$$

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?

Sol^n:

Given sample values: 45, 47, 50, 52, 48, 47, 49, 53, 51
Therefore, sample size n=9
Population Mean $\mu = 47.50$
$\therefore$Sample mean $\bar{x} = \frac{1}{n}\sum x = \frac{442}{9} = 49.11$
Variance, $s^2 = \frac{1}{n-1}\sum(x - \bar{x})^2$

$$\Rightarrow s^2 = \frac{1}{8}\begin{Bmatrix} (45-49.11)^2 + (47-49.11)^2 + (50-49.11)^2 + (52-49.11)^2 + (48-49.11)^2 \\ +(47-49.11)^2 + (49-49.11)^2 + (53-49.11)^2 + (51-49.11)^2 \end{Bmatrix}$$

$$\Rightarrow s^2 = \frac{54.9}{8} = 6.8625 \Rightarrow s = \sqrt{6.8625} = 2.6196$$

$\therefore$Thee Null hypothesis $H_0: \mu = 47.5$

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$
$$\Rightarrow t = \frac{49.11 - 47.5}{\left(2.6196/\sqrt{9}\right)}$$
$$\Rightarrow t = \frac{1.61}{0.8732}$$
$$\Rightarrow t = 1.8437$$

$\therefore$ Level of significance =5%
Critical value at 5 % level of significance for v=9-1=8 degrees of freedom is 2.3060.
Since the calculated value 1.8437 is less than the tabulated value 2.3060.
Hence the Null hypothesis is accepted.

Module – 5

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**

Sol^n:

To carry out the analysis of variance, we form the following tables

							Total	Squares
F1	8	12	19	8	6	11	T_1=64	T^2_1=4096
F2	4	5	4	6	9	7	T_2=35	T^2_2=1225
F3	11	8	7	13	7	9	T_3=55	T^2_3=3025
Total T							154	-

The squares are as follows

							Sum of Squares
F1	64	144	361	64	36	121	790
F2	16	25	16	36	81	49	223
F3	121	64	49	169	49	81	533
Grand Total - $\sum_i \sum_j {x_{ij}}^2$							1546

Set the null hypotheses $H_0: \mu_1 = \mu_2 = \mu_3$

Correction Factor $CF = \frac{T^2}{N} = \frac{(154)^2}{18} = \frac{23716}{18} = 1317.55$

Therefore Total sum of squares $TSS = \sum_i \sum_j {x_{ij}}^2 - CF$

$$\Rightarrow TSS = 1546 - 1317.55$$
$$\Rightarrow TSS = 228.45$$

Sum of the squares of between the treatments $SST = \sum_i \frac{{T_i}^2}{n_i} - CF$

$$SST = \frac{4096}{6} + \frac{1225}{6} + \frac{3025}{6} - 1317.55$$
$$\Rightarrow SST = 682.66 + 204.166 + 504.166 - 1317.55$$
$$\Rightarrow SST = 1391 - 1317.55$$
$$\Rightarrow SST = 73.45$$

Therefore, sum of squares due to error SEE=TSS-SST

$$\Rightarrow SSE = 228.45 - 73.45 \Rightarrow SSE = 155$$

Sources variation	**d.f.**	**SS**	**MSS**	**F Ratio**
Between treatments	3-1=2	SST=73.45	$MST = \frac{73.45}{2} = 36.725$	$F = \frac{36.725}{10.33} = 3.55$

Error	18-3=15	SSE=155	$MSE = \frac{155}{15} = 10.33$	
Total	18-1=17	-	-	

Since evaluated value 3.55 <3.68 for F(2,15) at 5% level of significance

Hence the null hypothesis is accepted , there is no significance between the three process.

b) Analyze and interpret the following statistics concerning output of wheat for field obtained as result of experiment conducted to test for Four varieties of wheat viz. A, B, C and D under Laton 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**

Sol^n:

Given observations are

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

Null hypothesis H_o : There is no significant difference between rows, columns and treatment

Code the data by subtracting 20 from each value, we get

					T	T^2
	C 5	B 3	A 0	D 0	8	64
	A -1	D -1	C 1	B -2	-3	9
	B -1	A -6	D -3	C 0	-10	100
	D -3	C 0	B 1	A -5	-7	49
P	0	-4	-1	-7	=- 12	
P^2	0	16	1	49	-	-

The squares are as follows

C 25	B 9	A 0	D 0	
A 1	D 1	C 1	B 4	
B 1	A 36	D 9	C 0	
D 9	C 0	B 1	A 25	
36	46	11	29	$\sum_i \sum_j {x_{ij}}^2$ =122

Correction Factor $CF = \frac{T^2}{N} = \frac{(-12)^2}{16} = \frac{144}{16} = 9$

Therefore, Total sum of squares $TSS = \sum_i \sum_j {x_{ij}}^2 - CF$

$$\Rightarrow TSS = 122 - 9$$
$$\Rightarrow TSS = 113$$

Sum of the row squares $SSR = \sum_i \frac{{T_i}^2}{n_i} - CF$

$$SSR = \frac{64}{4} + \frac{9}{4} + \frac{100}{4} + \frac{49}{4} - 9$$
$$\Rightarrow SSR = 16 + 2.25 + 25 + 12.25 - 9$$
$$\Rightarrow SSR = 55.5 - 9$$
$$\Rightarrow SSR = 4$$

Sum of the column squares $SSC = \sum_i \frac{{P_i}^2}{n_i} - CF$

$$SSC = 0 + \frac{16}{4} + \frac{1}{4} + \frac{49}{4} - 9$$
$$\Rightarrow SSC = 4 + 0.25 + 12.25 - 9$$
$$\Rightarrow SSC = 16.5 - 9$$
$$\Rightarrow SSC = 7.5$$

To find the sum of the treatments

Observations					$Q = \sum(Observations)$	Q^2
A	0	-1	-6	-5	-12	144
B	3	-2	-1	1	1	1
C	5	1	0	0	6	36
D	0	-1	-3	-3	-7	49

Sum of the squares of treatments $SST = \sum_i \frac{{Q_i}^2}{n_i} - CF$

$$SST = \frac{144}{4} + \frac{1}{4} + \frac{36}{4} + \frac{49}{4} - 9$$
$$\Rightarrow SST = 36 + 0.25 + 9 + 12.25 - 9$$
$$\Rightarrow SST = 57.50 - 9$$
$$\Rightarrow SST = 48.50$$

$\therefore SSE = TSS - SSR - SSC - SST \Rightarrow SSE = 113 - 46.5 - 7.5 - 48.50 = 10.5$, We know that $F(3,6) = 4.76$

Sources variation	d.f	SS	MSS	F Ratio	Conclusion
Rows	4-1=3	SSR=46.5	$MSR = \frac{46.5}{3} = 15.5$	$F_r = \frac{15.5}{1.75} = 8.85$	$F_r > F(3,6)$ H_0-Rejected
Columns	4-1=3	SSC=7.5	$MSC = \frac{7.5}{3} = 2.5$	$F_c = \frac{2.5}{1.75} = 1.428$	$F_c < F(3,6)$ H_0-Accepted
Treatments	4-1=3	SST=48.5	$MST = \frac{48.5}{3} = 16.16$	$F_T = \frac{16.16}{1.75} = 9.23$	$F_T > F(3,6)$ H_0-Rejected
Error	3x2=6	SSE=10.5	$MSE = \frac{10.5}{6} = 1.75$	-	-
Total	25-1=24	-	-	-	-

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 it the variety differences are significant at 5% significant level.

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

Sol^n:

To carry out the analysis of variance, we form the following tables

Per acre production data				T	T^2
Plot of land	Variety				
	A	B	C		
1	6	5	5	16	256
2	7	5	4	16	256
3	3	3	3	9	81
4	8	7	4	19	361
P	24	20	16	=60	-
P^2	576	400	256		

The squares are as follows

Variety			
A	B	C	
36	25	25	
49	25	16	
9	9	9	
64	49	16	Grand Total - $\sum_i \sum_j x_{ij}^2$=332

Set the null hypotheses $H_0: \mu_1 = \mu_2 = \mu_3$, N=12

Correction Factor $CF = \frac{T^2}{N} = \frac{(60)^2}{12} = \frac{3600}{12} = 300$

Therefore, Total sum of squares $TSS = \sum_i \sum_j x_{ij}^2 - CF$

$$\Rightarrow TSS = 332 - 300$$
$$\Rightarrow TSS = 32$$

Sum of the row squares $SSR = \sum_i \frac{{T_i}^2}{n_i} - CF$

$$SSR = \frac{256}{3} + \frac{256}{3} + \frac{81}{3} + \frac{361}{3} - 300$$
$$\Rightarrow SSR = 85.33 + 85.33 + 27 + 120.33 - 300$$
$$\Rightarrow SSR = 318 - 300$$
$$\Rightarrow SSR = 18$$

Sum of the column squares $SSC = \sum_i \frac{{P_i}^2}{n_i} - CF$

$$SSC = \frac{576}{4} + \frac{400}{4} + \frac{256}{4} - 300$$
$$\Rightarrow SSC = 144 + 100 + 64 - 300$$
$$\Rightarrow SSC = 308 - 300$$
$$\Rightarrow SSC = 8$$

Therefore SSE=TSS-SSR-SSC

SSE=32-18-8=6

Sources variation	d.f.	SS	MSS	F Ratio
Rows	4-1=3	SSR=18	$MSR = \frac{18}{3} = 6$	$F_r = \frac{6}{1} = 6$
Columns	3-1=2	SSC=8	$MSC = \frac{8}{2} = 4$	
Error	3X2=6	SSE=6	$MSE = \frac{6}{6}$=1	$F_c = \frac{4}{1} = 4$
Total	12-1=11	-	-	

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

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

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

Group 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**

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

Sol^n:

Given observations from different people (A, B, C) to the different drugs (X, Y, Z) are as

Group of people	Drug			T	T^2
	X	Y	Z		
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	-	-

Where N=6+6+6=18

Correction Factor $CF = \frac{T^2}{N} = \frac{(187)^2}{18} = \frac{34969}{18} = 1942.722$

The squares are as follows

Group of people	Drug			Sum of Squares	
	X	Y	Z		
A	196 225	100 81	121 121	844	
B	144 121	49 64	100 121	599	
C	100 121	121 121	64 49	576	$\sum_i \sum_j x_{ij}^2$ =2019

Therefore, Total sum of squares $TSS = \sum_i \sum_j x_{ij}^2 - CF$

$$\Rightarrow TSS = 2019 - 1942.722$$

$$\Rightarrow TSS = 76.28$$

Sum of the row squares $SSR = \sum_i \frac{T_i^2}{n_i} - CF$

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

$$\Rightarrow SSR = 816.67 + 580.16 + 560.67 - 1942.722$$

$$\Rightarrow SSR = 14.78$$

Sum of the column squares $SSC = \sum_i \frac{P_i^2}{n_i} - CF$

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

$$\Rightarrow SSC = 888.16 + 522.66 + 560.67 - 1942.722$$

$$\Rightarrow SSC = 28.77$$

SS within samples (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

Therefore,

SSE=TSS-SSR-SSC-SST

$$\Rightarrow \text{SSE} = 76.28 - 14.78 - 28.77 - 3.5$$

$$\Rightarrow \text{SSE} = 29.23$$

We have $F_{(2,9)}$=4.26 , F(4,9)=3.63

Sources variation	d.f.	SS	MSS	F Ratio	Conclusion
Rows	3-1=2	SSR=14.78	$MSR = \frac{14.78}{2} = 7.39$	$F_r = \frac{7.39}{0.389} = 19$	$F_r > F(2,9)$ H_0-Rejected
Columns	3-1=2	SSC=28.77	$MSC = \frac{28.77}{2} = 14.385$	$F_c = \frac{14.385}{0.389} = 37$	$F_r > F(2,9)$ H_0-Rejected
Treatments	9	SST=3.5	$MST = \frac{3.5}{9} = 0.389$	$F_T = \frac{7.33}{0.389} = 18.84$	$F_T > F(4,9)$ H_0-Rejected
Error	4	SSE=29.33	$MSE = \frac{29.33}{4}$=7.33	-	-

Model Paper – 2 Solutions

Module – 1

1.

a) A random variable X has a probability function for various values of x. Find i) k, ii) $P(x < 6)$, iii) $P(x \geq 6)$ and $P(3 < x \leq 6)$. Also find the probability distribution and distribution function of x.

x	0	1	2	3	4	5	6	7
$P(x)$	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

Sol[n]**:**

Let X be the random variable for the random values,

$x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3, x_5 = 4, x_6 = 5, x_7 = 6, x_8 = 7$

and the given probabilities are,

$P(X = x_1) = P(0) = 0$

$P(X = x_2) = P(1) = k$

$P(X = x_3) = P(2) = 2k$

$P(X = x_4) = P(3) = 2k$

$P(X = x_5) = P(4) = 3k$

$P(X = x_6) = P(5) = k^2$

$P(X = x_7) = P(6) = 2k^2$

$P(X = x_7) = P(7) = 7k^2 + k$

i) We know that,

$\sum_{i=1}^{7} P(X = x_i) = 1$

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

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

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

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

$$\Rightarrow 10k - 1 = 0\,, k + 1 = 0$$

$$\Rightarrow k = \frac{1}{10}\,, k \neq -1$$

x	0	1	2	3	4	5	6	7
$P(x)$	O	0.1	0.2	0.2	0.3	0.01	0.02	0.17

$$i)\ P(x<6)=1-P(x\geq 6)=1-\{P(6)+P(7)\}=1-\{0.02+0.17\}=0.81$$

$$ii)\ P(x\geq 6)=P(6)+P(7)=0.02+0.17=0.19$$

$$iii)\ P(3<x\leq 6)=P(4)+P(5)+P(6)=0.3+0.01+0.02=0.33$$

b) Find the mean and variance of Poisson Distribution.

Sol^n:

Let X be the discrete random variable for any real value λ , such that the probability mass function of poisson distribution can be defined as,

$$P(X=x)=P(x)=\begin{cases}\frac{e^{-\lambda}\lambda^x}{x!} & ,x\geq 0\\ 0, & otherwise\end{cases}$$

WKT, the probability mass function of the poisson distribution is,

$$P(X=x)=P(x)=\begin{cases}\frac{e^{-\lambda}\lambda^x}{x!} & , & x\geq 0\\ 0 & , & otherwise\end{cases}$$

i) Mean:

$$\mu=E(x)=\sum_{x=0}^{\infty}xP(x)$$

$$=\sum_{x=0}^{\infty}x\,\frac{e^{-\lambda}\lambda^x}{x!}$$

$$=\sum_{x=0}^{\infty}x\,\frac{e^{-\lambda}\lambda^{x-1}\lambda}{x(x-1)!}$$

$$=\lambda\sum_{x=1}^{\infty}\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}$$

$$=\lambda(1)$$

$$\mu=\lambda$$

ii) Variance:

$$\sigma^2=E(x^2)-\mu^2\ ----(1)$$

$$=E(x(x-1)+x)-\mu^2$$

$$=E\big(x(x-1)\big)+E(x)-\mu^2----(2)$$

$$\therefore E(x(x-1))=\sum_{x=0}^{\infty}x(x-1)P(x)$$

$$=\sum_{x=0}^{\infty}x(x-1)\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} x(x-1)\ \frac{e^{-\lambda}\lambda^{x-2}\lambda^2}{x(x-1)(x-2)!}$$

$$= \lambda^2 \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!}$$

$$= \lambda^2(1)$$

$$E(x(x-1)) = \lambda^2(1)$$

$$(2) \Rightarrow \sigma^2 = \lambda^2 + \lambda - \lambda^2$$

$$\Rightarrow \sigma^2 = \lambda$$

$$S.D = \sigma = \sqrt{\lambda}$$

$$Mean = \lambda = np$$

c) In a certain town the duration of shower has mean 5 minutes, what is the probability that shower will last for,
i)10 minutes and more
ii)Less than 10 minutes
iii)Between 10 & 12 minutes.

Sol^n:

Given X be a continuous random variable of an exponential distribution is,

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

and given the mean of exponential distribution is 5.

$\Rightarrow \mu = 5 \Rightarrow \frac{1}{\alpha} = 5 \Rightarrow \alpha = \frac{1}{5}$

$$\therefore f(x) = \begin{cases} \frac{1}{5} e^{-\frac{x}{5}} & ,x \ge 0 \\ 0 & ,otherwise \end{cases}$$

iii) The probability that the shower will last 10 minutes and more is,

$$P(x \ge 10) = \int_{10}^{\infty} f(x)dx = \int_{10}^{\infty} \frac{1}{5} e^{-\frac{x}{5}} dx = \frac{1}{5}\int_{10}^{\infty} e^{-\frac{x}{5}} dx = -\left[e^{-\frac{x}{5}}\right]_{10}^{\infty} = -[0 - e^{-2}] = \frac{1}{e^2}$$

iv) The probability that the shower will last less than 10 minutes is,

$$P(x < 10) = \int_{-\infty}^{10} f(x)dx = \int_{-\infty}^{0} f(x)dx + \int_{0}^{10} f(x)dx$$

$$\Rightarrow P(x < 10) = 0 + \int_{0}^{10} \frac{1}{5} e^{-\frac{x}{5}} dx = \frac{1}{5}\int_{0}^{10} e^{-\frac{x}{5}} dx = -\left[e^{-\frac{x}{5}}\right]_{0}^{10} = -[e^{-2} - 1] = 1 - \frac{1}{e^2}$$

v) The probability that the shower will last between 10 & 12 minutes is,

$$P(10 < x < 12) = \int_{10}^{12} f(x)dx = \int_{10}^{12} \frac{1}{5} e^{-\frac{x}{5}} dx = \frac{1}{5}\int_{10}^{12} e^{-\frac{x}{5}} dx$$

$$\Rightarrow P(10 < x < 12) = -\left[e^{-\frac{x}{5}}\right]_{10}^{12} = -\left[e^{\frac{-12}{5}} - e^{-2}\right] = \frac{1}{e^{\frac{12}{5}}} - \frac{1}{e^2}$$

2.

a) Determine the value of k, so that the function $f(x) = k(x^2+4)$, for $x = 0,1,2,3$ can serve as a probability distribution of the discrete random variable X:

Also find,

i) $P[0 < X \leq 2]$

ii) $P[X \geq 1]$

Sol^n:

Given the probability function for the discrete random variable $x = f(x) = k(x^2+4) \; \forall \; x = 0,1,2,3$

x	0	1	2	3
$P(x)$	4k	5k	8k	13k

WKT,

$$\sum_{i=1}^{4} P(X = x_i) = 1$$
$$\Rightarrow 4k + 5k + 8k + 13k = 1$$
$$\Rightarrow 30k = 1$$
$$\Rightarrow k = \frac{1}{30}$$

i) $P(0 < X \leq 2) = P(1) + P(2) = 5k + 8k = 13k = \frac{13}{30}$

ii) $P(X \geq 1) = P(1) + P(2) + P(3) = 5k + 8k + 13k = 26k = \frac{26}{30}$

b) In 800 families with 5 children each, how many family would be expected to have,

i)3 boys

ii)5 girls

iii)Atmost 2 girls

iv)Either 2 or 3 boys

by assuming probability for boys and girls to be equal.

Sol^n:

The total number of families given is 800 and number of children per family is, n=5
Given the probability of boy or girl to born, p=0.5
then q = 1-p = 1-0.5 = 0.5

The pmf of binomial distribution is, $P(X = x) = P(x) = n_{C_x} p^x q^{n-x} = 5_{C_x}(0.5)^x (0.5)^{5-x}$
$= 5_{C_x}(0.5)^5 = 5_{C_x}(0.03125)$

i) The probability to have exactly 3 boys, $P(3) = 5_{C_3}(0.03125)$
$= (10)(0.03125)$
$= 0.3125$

$\therefore$ The total number of families may have exactly 3 boys $= 800 \times 0.3125 = 250.$

ii) The probability to have exactly 5 girls, $P(5) = 5_{C_5}(0.03125)$
$= (1)(0.03125)$
$= 0.03125$

$\therefore$ The total number of families may have exactly 5 girls, $= 800 \times 0.03125 = 25.$

iii) The probability to have atmost two girls, $P(x \leq 2) = P(0) + P(1) + P(2)$
$= 5_{C_0}(0.03125) + 5_{C_1}(0.03125) + 5_{C_2}(0.03125)$
$= 0.03125 + 0.15625 + 0.3125$
$= 0.5$

$\therefore$ The total number of families may have atmost two girls, $= 800 \times 0.5 = 400.$

iv) The probability to have either 2 or 3 boys, $P(2 \leq x \leq 3) = P(2) + P(3)$
$= 5_{C_2}(0.03125) + 5_{C_3}(0.03125)$
$= 0.03125 + 0.03125$
$= 0.0625$

$\therefore$ The total number of families may have either 2 or 3 boys, $= 800 \times 0.0625 = 500.$

c) In a normal distribution, 31% of items are under 45 and 8% of the items are 0ver 64.Find the mean and standard deviation of the distribution.

Sol^n:

Let X be the continuous random variable

Given

Let μ andσ be the Mean and Standard deviation of the distribution

$\therefore$The standard normal variate $z = \frac{x-\mu}{\sigma}$ --------------(1)

When x=35 the standard normal variate $z = \frac{45-\mu}{\sigma} = z_1(Say)$

When x=64 the standard normal variate $z = \frac{64-\mu}{\sigma} = z_2(Say)$

Given

$$P(x < 35) = P(z < z_1) = 0.31$$
$$\Rightarrow P(z < z_1) = P(-\infty < z < 0) - P(0 < z < z_1) = 0.31$$
$$\Rightarrow 0.5 - A(z_1) = 0.31$$
$$\Rightarrow A(z_1) = 0.5 - 0.31 = 0.19$$

$$\Rightarrow A(z_1) = A(0.5)$$
$$\Rightarrow z_1 = 0.5$$
$$\Rightarrow \frac{45-\mu}{\sigma} = 0.5$$
$$\Rightarrow \mu + 0.5\sigma = 45 - - - - - - -(2)$$

And $P(x > 64) = P(z > z_2) = 0.08$

$$\Rightarrow P(z > z_2) = 0.5 - P(0 < z < z_2) = 0.08$$
$$\Rightarrow 0.5 - A(z_2) = 0.08$$
$$\Rightarrow A(z_2) = 0.08 - 0.5 = -0.42$$
$$\Rightarrow A(z_2) = A(1.4)$$
$$\Rightarrow z_2 = 1.4$$
$$\Rightarrow \frac{64-\mu}{\sigma} = 1.4$$
$$\Rightarrow \mu + 1..4\sigma = 64 - - - - - - - - -(3)$$

Solving eq(2) and (3)

we get ,

$$\mu = 50$$
$$\sigma = 10$$

Module – 2

3.

a) If the joint probability distribution of X and Y is given by

$$f(x,y) = \frac{x+y}{30}, for\ x = 0,1,2,3; y = 0,1,2$$

Find,

i) $P[X \le 2, Y = 1]$

ii) $P[X > 2, Y \le 1]$

iii) $P[X > Y]$

Sol^n:

Given X, Y are the two random variables follows,

$f(x,y) = P(x,y) = \frac{x+y}{30}$ $\quad\forall\ x = 0,1,2,3\ \&\ y = 0,1,2$

The joint probability distribution can be defined as,

Y \ X	0	1	2
0	0	$\frac{1}{30}$	$\frac{2}{30}$
1	$\frac{1}{30}$	$\frac{2}{30}$	$\frac{3}{30}$
2	$\frac{2}{30}$	$\frac{3}{30}$	$\frac{4}{30}$
3	$\frac{3}{30}$	$\frac{4}{30}$	$\frac{5}{30}$

i) $P(X \leq 2, Y = 1) = P(0,1) + P(1,1) + P(2,1) = \frac{1}{30} + \frac{2}{30} + \frac{3}{30} = \frac{6}{30} = \frac{1}{5}$

ii) $P(X > 2, Y \leq 1) = P(3,0) + P(3,1) = \frac{3}{30} + \frac{4}{30} = \frac{7}{30}$

iii) $P(X > Y) = P(1,0) + P(2,0) + P(2,1) + P(3,0) + P(3,1) + P(3,2)$
$= \frac{1}{30} + \frac{2}{30} + \frac{3}{30} + \frac{3}{30} + \frac{4}{30} + \frac{5}{30} = \frac{18}{30} = \frac{3}{5}$

b) Find the fixed probability vector of the regular stochastic matrix P = $\begin{bmatrix} \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{\frac{1}{6}} & \mathbf{\frac{1}{2}} & \mathbf{\frac{1}{3}} \\ \mathbf{0} & \mathbf{\frac{2}{3}} & \mathbf{\frac{1}{3}} \end{bmatrix}$**.**

Sol^n:

Given, $P = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$

Since the given matrix P is of order 3x3, the required fixed probability vector Q must be also order of 3x3.

Let $Q = [x \quad y \quad z] For every x \geq 0, y \geq 0, z \geq 0 \& x + y + z = 1$

Also, QP=Q

$$\therefore QP = [x \quad y \quad z]\begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

$$\Rightarrow QP = \begin{bmatrix} \frac{1}{6}y & x + \frac{1}{2}y + \frac{2}{3}z & \frac{1}{3}y + \frac{1}{3}z \end{bmatrix}$$

$WKT, \ QP = Q$

$\Rightarrow \begin{bmatrix} \frac{1}{6}y & x + \frac{1}{2}y + \frac{2}{3}z & \frac{1}{3}y + \frac{1}{3}z \end{bmatrix} = [x \quad y \quad z]$

$\Rightarrow x = \frac{1}{6}y, y = x + \frac{1}{2}y + \frac{2}{3}z, z = \frac{1}{3}y + \frac{1}{3}z$

$\Rightarrow 6x - y = 0, 2x - 7y = -4 \ldots (1) 2x + 3y = 2 \ldots\ldots (2)$

By solving eq (1) & (2)

$$\Rightarrow x = \frac{1}{10}, y = \frac{6}{10}, z = 1 - \frac{1}{10} - \frac{6}{10} = 1 - \frac{7}{10} = \frac{3}{10}$$

$$\therefore Q = [x \quad y \quad z] = \begin{bmatrix} \frac{1}{10} & \frac{6}{10} & \frac{3}{10} \end{bmatrix}$$

c) A gambler's luck follows a pattern: if he wins a game, the probability of winning next game is 0.6. However, he loses the game, the probability of losing the next game is 0.7. There is an even chance of gambler winning the first game if so,

i) What is the probability of winning second game.

ii) What is the probability of winning the third game.

iii) In the long run, how often he will win.

Sol^n:

Let W = Win the game , L = Lose the game

The transition probability matrix is given,

$$P = \begin{bmatrix} p_{ww} & p_{wl} \\ p_{lw} & p_{ll} \end{bmatrix} = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix}$$

And we know that the probability of winning and losing have the equal priority.

$\therefore$ The initial probability vector $p^{(0)} = [p_w{}^{(0)} \quad p_l{}^{(0)}] = [0.5 \quad 0.5]$

$\therefore p^{(1)} = p^{(0)}P = [0.5 \quad 0.5].\begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix} = [0.45 \quad 0.55]$

$\therefore p^{(2)} = p^{(1)}P = [0.45 \quad 0.55].\begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix} = [0.435 \quad 0.565]$

$\therefore p^{(3)} = p^{(2)}P = [0.435 \quad 0.565].\begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix} = [0.4305 \quad 0.5605]$

i) $\therefore p_w{}^{(2)} = 0.435 = 43.5\%$

ii) $\therefore p_w{}^{(3)} = 0.4305 = 43.05\%$

iii) Let Q=[x y] be the probability vector for which x+y=1

$\therefore QP = Q$

$\therefore [x \quad y].\begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix} = [x \quad y]$

$\Rightarrow 0.6x + 0.3y = x \quad , \quad 0.4x + 0.7y = y$

$\Rightarrow 0.4x - 0.3y = 0$

$\Rightarrow 0.4x - 0.3(1 - x) = 0$

$\Rightarrow 0.4x - 0.3 + 0.3x = 0$

$\Rightarrow 0.7x = 0.3$

$\Rightarrow x = \frac{3}{7}$

$\Rightarrow y = 1 - x \Rightarrow y = 1 - \frac{3}{7} \Rightarrow y = \frac{4}{7}$

$\therefore Q = [p_w \quad p_l] = \begin{bmatrix} \frac{3}{7} & \frac{4}{7} \end{bmatrix}$

4.

a) Determine the value of k so that the function

$\boldsymbol{f(x,y) = k|x-y|, for\ x = -2, 0, 2; y = -2, 3}$

represents joint probability distribution of the random variables X and Y. Also determine $Cov(X,Y)$.

Sol^n:

Given X, Y are the two random variables follows,

$f(x,y) = P(x,y) = k|X - Y| \qquad \forall\ x = -2, 0, 2\ \&\ y = -2, 3$

The joint probability distribution can be defined as,

X \ Y	-2	3	$f(x)$
-2	0	$5k$	$5k$
0	$2k$	$3k$	$5k$
2	$4k$	k	$5k$
$g(y)$	$6k$	$9k$	$15k$

$\therefore WKT, \quad \sum_i \sum_j p(x_i, y_j) = 1$

$\Rightarrow 15k = 1$

$\Rightarrow k = \frac{1}{15}$

$\therefore$ The marginal distribution of X and Y are,

x	-2	0	2
$f(x)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

y	-2	3
$g(y)$	$\frac{2}{5}$	$\frac{3}{5}$

$$\mu_X = E(X) = \sum_{i=1}^{2} x_i f(x_i) = \left((-2)\times\frac{1}{3}\right) + \left(0\times\frac{1}{3}\right) + \left(2\times\frac{1}{3}\right) = -\frac{2}{3}+\frac{2}{3} = 0$$

$$\mu_Y = E(Y) = \sum_{j=1}^{3} y_j g(y_j) = \left(-2\times\frac{2}{5}\right) + \left(3\times\frac{3}{5}\right) = -\frac{4}{5}+\frac{9}{5} = 1$$

$$E(XY) = \sum_{i=1}^{2}\sum_{j=1}^{3} x_i y_j f(x_i, y_j)$$
$$= 0 - 30k - 16k + 6k = -46k + 6k = -40k = \frac{-40}{15} = \frac{-8}{3}$$

$$COV(X,Y) = E(XY) - \mu_X\mu_Y = \frac{-8}{3} - (0)(1) = \frac{-8}{3} = -2.67$$

b) Show that the matrix $P = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}$ is a regular stochastic matrix.

Sol^n:

Given matrix A, each element is non-negative and the sum of the elements in each row is equal to 1.
$\therefore$ P is a stochastic matrix

$$\text{Let } P^2 = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

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

Hence, all entries in P^3 are non-negative and sum of each row = 1
$\therefore$ The given matrix P is a regular stochastic matrix.

c) Three boys A, B, C are throwing ball to each other. "A" always throws the ball to "B" and "B" always throws ball to "C". "C" is just as likely to throw the ball to "B" as to "A". If, "C" was the first person to throw the ball, find the probabilities that after three throws.
i) A has the ball
ii) B has the ball
iii) C has the ball

Sol^n:

Given three boys A, B, C are throwing a ball associated with the transition probability matrix of the Markov chain as below,

$$P = \begin{bmatrix} p_{AA}^{(1)} & p_{AB}^{(1)} & p_{AC}^{(1)} \\ p_{BA}^{(1)} & p_{BB}^{(1)} & p_{BC}^{(1)} \\ p_{CA}^{(1)} & p_{CB}^{(1)} & p_{CC}^{(1)} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

$$\Rightarrow P^2 = P.P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix}.\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

$$\Rightarrow P^2 = \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

$$\therefore P^3 = P^2.P = \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{bmatrix} . \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

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

Initially if C has the ball, associated with the initial probability vector is given by $p^{(0)} = [0 \quad 0 \quad 1]$

$$\therefore p^{(3)} = p^{(0)}P^3 = [0 \quad 0 \quad 1].\left[\begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}\right] = [1/4 \quad 1/4 \quad 1/2]$$

$$\therefore p^{(3)} = [p_A{}^{(3)} \quad p_B{}^{(3)} \quad p_C{}^{(3)}] = [1/4 \quad 1/4 \quad 1/2]$$

Thus, after three throws, the probability that the ball is with A is $p_A{}^{(3)} = \frac{1}{4}$, with B is $p_B{}^{(3)} = \frac{1}{4}$ and with C is $p_C{}^{(3)} = \frac{1}{2}$

Module – 3

5.

a) Explain the following terms:

i) Standard Error

ii) Statistical Hypothesis

iii) Critical Region of a statistical test

iv) Test of Significance

Sol^n:

Standard Error:

The standard deviation of the sampling distribution of a statistic is Known as Standard Error (S.E.).

Null Hypothesis (or) Statistical Hypothesis:

The **null hypothesis** is a general statement or default position that there is no relationship between two measured phenomena or no association among groups.

Example: Given the test scores of two random samples, one of men and one of women, does one group differ from the other? A possible null hypothesis is that the mean male score is the same as the mean female score:

$H_0: \mu_1 = \mu_2$

where

H_0 = the null hypothesis,

μ_1 = the mean of population 1, and

μ_2 = the mean of population 2.

A stronger null hypothesis is that the two samples are drawn from the same population, such that the variances and shapes of the distributions are also equal.

Critical Region:

A critical region, also known as the rejection region, is a set of values for the test statistic for which the null hypothesis is rejected. i.e. if the observed test statistic is in the critical region then we reject the null hypothesis and accept the alternative hypothesis.

Test of Significance (or) Test of Hypothesis:

Let x be the observed number of successes in a sample size of n and $\mu = np$ be the expected number of successes .Then the standard normal variate $Z$is defined as

$$Z = \frac{x-\mu}{\sigma} = \frac{x-np}{\sqrt{npq}}$$

b) In 324 throws of a six faced 'die', an odd number turned up 181 times. Is it possible to think that the 'die' is an unbiased one at 5% level of significance?

Sol^n:

Let us suppose that the die is unbiased.
and let p= the probability of the turn up of an odd number is=3/6=1/2=0.5
Since p+q=1, q=1-p=1/2=0.5
Expected number of successes=np=324x0.5=162, npq=81
∴ The difference is $x - \mu$ =181-162=19
∴ Consider $Z = \frac{x-\mu}{\sigma} = \frac{x-np}{\sqrt{npq}}$
$\Rightarrow Z = \frac{19}{\sqrt{81}} = \frac{19}{9} 2.11 < 2.58$
Thus, we can that the die is unbiased.

c) In an examination given to students at a large number of different schools the mean grade was 74.5 and S.D grade was 8. At one particular school where 200 students took the examination the mean grade was 75.9. Discuss the significance of this result at both 5% and 1% level of significance.

Sol^n:

The level of significance $\alpha = 0.05$ (5%)$\Rightarrow Z_{0.05}$ =1.96
The level of significance $\alpha = 0.01$ (1%)$\Rightarrow Z_{0.01}$=1.64
Given

$$n = 200$$
$$\sigma = 8$$
$$\mu = 74.5 and \bar{x} = 75.9$$

We calculate Z through Test Statistic,

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$
$$\Rightarrow Z = \frac{75.9 - 74.5}{8/\sqrt{200}}$$
$$\Rightarrow Z = \frac{1.4}{8/14.1421}$$
$$\Rightarrow Z = \frac{1.4 \times 14.1421}{8}$$
$$\Rightarrow Z = 2.4748$$

i) Thus At 5% level, the tabulated value of Zα is 1.645.
Since |Z| = 2.4748 > 1.96
Hence Null Hypothesis H_0 is rejected at 5% level of significance.
ii) Thus At 1% level, the tabulated value of Zα is 1.645.
Since |Z| = 2.4748 > 1.645
Hence Null Hypothesis H_0 is rejected at 1% level of significance.

6.
a) Define:
i) Alternative Hypothesis
ii) A Statistic
iii) Level of Significance
iv) Two-tailed Test

Sol^n:

Alternative hypothesis:

An alternative hypothesis is an opposing theory to the null hypothesis. For example, if the null hypothesis predicts something to be true, the alternative hypothesis predicts it to be false. The alternative hypothesis often is the statement you test when attempting to disprove the null hypothesis.

Statistic:

A statistic (singular) or sample statistic is any quantity computed from values in a sample which is considered for a statistical purpose. Statistical purposes include estimating a population parameter, describing a sample, or evaluating a hypothesis. The average (or mean) of sample values is a statistic.

Level of significance:

The level of significance is the measurement of the statistical significance. It defines whether the null hypothesis is assumed to be accepted or rejected. It is expected to identify if the result is statistically significant for the null hypothesis to be false or rejected.

Two-tailed test:

A two-tailed test, in statistics, is a method in which the critical area of a distribution is two-sided and tests whether a sample is greater than or less than a certain range of values. It is used in null-hypothesis testing and testing for statistical significance.

b) A coin is tossed 1000 times and head turns up 540 times. Decide on the hypothesis that the coin is unbiased at 1 % level of significance.

Sol^n:

Let us suppose that the coin is unbiased.
and let p= the probability of getting a head in one toss=1/2=0.5
Since p+q=1, q=1-p=1/2=0.5
Expected number of heads in 1000 tosses=np=1000x0.5=500 , npq=250
∴ The difference is $x - \mu$ =540-500=40
∴ Consider $Z = \frac{x-\mu}{\sigma} = \frac{x-np}{\sqrt{npq}}$
$\Rightarrow Z = \frac{40}{\sqrt{250}} = 2.53 < 2.58$
1% level of significance = 99% confidence level.
Therefore accept the hypothesis that the coin is unbiased.

c) One type of air craft is found to develop engine trouble in 5 flights out of a total of 100 and another type in 7 flights out of a total of 200 flights. Is there a significance difference in the two types of air craft's so far as engine defects are concerned? Test at 5% significance level.

Sol^n:

Set the null hypothesis $H_0; P_1 = P_2$
Set the Alternative hypothesis $H_1: P_1 \neq P_2$
The level of significance $\alpha = 0.05\ (5\%)$
Given

$n_1 = 100, n_2 = 200\ \ \&\ x_1 = 5, x_2 = 7$

$\therefore p_1 = \frac{5}{100} = 0.05, p_2 = \frac{7}{200} = 0.35$

We know that $P = \frac{x_1+x_2}{n_1+n_2}$

$\Rightarrow P = \frac{5+7}{100+200}$

$\Rightarrow P = \frac{12}{300}$

$\Rightarrow P = 0.04 \Rightarrow Q = 1 - P = 1 - 0.04 = 0.96$

$\therefore Z = \frac{p_1-p_2}{\sqrt{PQ\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$

$\Rightarrow Z = \frac{0.05-0.35}{\sqrt{\left(0.04\times0.96\right)\left(\frac{1}{100}+\frac{1}{200}\right)}}$

$\Rightarrow Z = -\frac{0.3}{\sqrt{(0.384)(0.015)}}$

$\Rightarrow Z = -\frac{0.3}{\sqrt{0.00576}}$

$\Rightarrow Z = -\frac{0.3}{0.07589}$

$\Rightarrow Z = -3.953$

At 5% level, the tabulated value of Zα is 1.645.
Since |Z| = 3.953 > 1.645
Hence Null Hypothesis H_0 is rejected at 5% level of significance.

Module – 4

7.

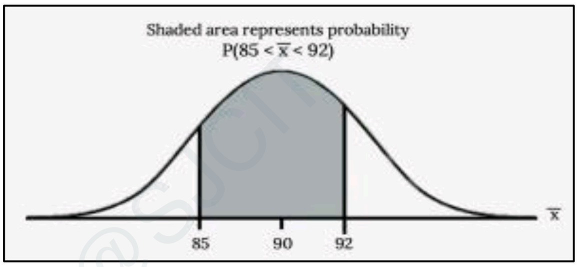
a) An unknown distribution has a mean of 90 and a standard deviation of 15. Samples of size n = 25 are drawn randomly from the population. Find the probability that the sample mean is between 85 and 92.

Sol^n:

Given,
Sample size n=25
Mean of the population $\mu = 90$
Variance of the population $\Rightarrow \sigma = 15$

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

$$\Rightarrow \bar{X} \sim N\left(90, \frac{15}{\sqrt{25}}\right)$$
$$\Rightarrow \bar{X} \sim N(90,3)$$
$$\therefore \text{ we know that } Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$
$$\Rightarrow Z = \frac{\bar{X}-90}{15/\sqrt{25}}$$
$$\Rightarrow Z = \frac{\bar{X}-90}{3}$$



$$\therefore \text{ At } \bar{X} = 85 \Rightarrow z = \frac{85-90}{3} = -\frac{5}{3} = -1.66$$
$$\therefore \text{ At } \bar{X} = 92 \Rightarrow z = \frac{92-90}{3} = \frac{2}{3} = 0.66$$

$$\therefore P(85 < \bar{X} < 92) = P(-1.66 < z < 0.66)$$
$$\Rightarrow P(-1.66 < z < 0.66) = P(0 < z < 1.66) + P(0 < z < 0.66)$$
$$= 0.4515 + 0.2454$$
$$\Rightarrow P(-1.66 < z < 0.66) = 0.6965$$

b) The heights of a random sample of 50 college students showed a mean of 174.5 centimeters and a standard deviation of 6.9 centimeters. Construct a 99% confidence interval for the mean height of all college students.

Sol^n:

Given the sample size n=50
Average height of Students (Mean) $\mu = 174.5c.m.$
Standard deviation of the Students $\sigma = 6.9c.m.$
We know that, Confidence level of 99%, the corresponding z value is 2.576. This is determined from the normal distribution table.

Confidence interval $C.I. == Mean \pm Z(Standard Deviation/\sqrt{SampleSize} = \mu \pm Z\frac{\sigma}{\sqrt{n}}$

$$\therefore C.I. = 174.5 \pm \left(2.576 \times \frac{6.9}{\sqrt{50}}\right)$$
$$\Rightarrow C.I. = 174.5 \pm (2.576 \times 0.9758)$$
$$\Rightarrow C.I. = 174.5 \pm 2.5136$$

The lower end of the confidence interval is $= 174.5 - 2.5136 = 171.9864$
The upper end of the confidence interval is $= 174.5 + 2.5136 = 177.0136$
Therefore, with 99% confidence interval, the mean height of all college students is between 171.9864 timeters and 177.0136 centimeters.

c) A die was thrown 60 times and the following frequency distribution was observed:

Faces	1	2	3	4	5	6
Frequency	15	6	4	7	11	17

Test whether the die is unbiased at 5% significance level.

Sol^n:

The frequencies in the given data are the observed frequencies. Assuming that dice is unbiased, the expected number of frequencies for the numbers 1,2,3,4,5,6 to appear on the face is $\frac{60}{6} = 10$ each.
Now the data is as follows:

x	1	2	3	4	5	6
O_i	15	6	4	7	11	17
E_i	10	10	10	10	10	10

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

$$\Rightarrow \chi^2 = \frac{(15-10)^2}{10} + \frac{(15-6)^2}{10} + \frac{(15-4)^2}{10} + \frac{(15-7)^2}{10} + \frac{(15-11)^2}{10} + \frac{(15-17)^2}{10}$$
$$\Rightarrow \chi^2 = \frac{1}{10}[25 + 81 + 121 + 64 + 16 + 4] = \frac{311}{10}$$
$$\Rightarrow \chi^2 = 31.10$$

8.

a) In a recent study reported on the Flurry Blog, the mean age of tablet users is 34 years. Suppose the standard deviation is 15 years. Take a sample size n = 100. Using central limit theorem, find the probability that the sample mean age is more than 30 years.

Sol^n:

Given,
Sample size n=100
Mean of the population $\mu = 34$
Variance of the population $\Rightarrow \sigma = 15$

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

$$\Rightarrow \bar{X} \sim N\left(34, \frac{15}{\sqrt{100}}\right)$$
$$\Rightarrow \bar{X} \sim N(34, 1.5)$$
$$\therefore \text{ we know that } Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$
$$\Rightarrow Z = \frac{\bar{X}-34}{15/\sqrt{100}}$$
$$\Rightarrow Z = \frac{\bar{X}-34}{1.5}$$
$$\therefore \text{ At } \bar{X} = 30 \Rightarrow z = \frac{30-34}{1.5} = -\frac{4}{1.5} = -2.66$$

$$\therefore P(\bar{X} > 30) = P(z > -2.66)$$
$$\Rightarrow P(z > -2.66) = 0.5 + P(z < 2.66) = 0.5 + \mathrm{A}(2.66) = 0.9961$$

b) Suppose that 10, 12, 16, 19 is a sample taken from a normal population with variance 6.25. Find at 95% confidence interval for the population mean.

Sol^n:

Given samples are 10, 12, 16 and 19
Therefore, sample size n=4
Mean $\bar{X}$=14.25
Variance $\sigma^2 = 6.25 \Rightarrow \sigma = \sqrt{6.25} = 2.5$
We know, Confidence level of 95%, the corresponding z value is 1.96, This is determined from the normal distribution table.

Confidence interval $C.I. = \mu = Mean \pm Z(S\tan d\, ardDeviation/\sqrt{SampleSize} = \bar{X} \pm Z\frac{\sigma}{\sqrt{n}}$

$$\therefore C.I. = \mu = 14.25 \pm \left(1.96 \times \frac{2.5}{\sqrt{4}}\right)$$
$$\Rightarrow \mu = 14.25 \pm 2.45$$
$$\Rightarrow C.I = (14.25 - 2.45, 14.25 + 2.45) = (11.80, 16.70)$$

c) A random sample of 10 boys had the following I.Q: 70, 120, 110, 101, 88, 83, 95, 98, 107, 100. Does this data support the assumption of a population mean I.Q. of 100 at 5% level of Significance? (*Note*:$t_{0.05}$ =2.262 for 9 d.f.).

Sol^n:

Given the I.Q. of 10 boys
x: 70, 120, 110,101,88,83,95,98,107,100
$$\therefore \bar{x} = \frac{1}{n}\sum x = \frac{972}{10} = 97.2$$
$$\text{Variance, } s^2 = \frac{1}{n-1}\sum(x-\bar{x})^2$$
$$\Rightarrow s^2 = \frac{1}{9} \times 1833.6$$
$$\Rightarrow s^2 = 203.73333$$
$$\Rightarrow s = 14.2735$$

Given the mean of population $\mu = 100$

We have,

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$$

$$\Rightarrow t = \frac{97.2-100}{\left(\frac{14.2735}{\sqrt{10}}\right)}$$

$$\Rightarrow t = \frac{-2.8}{4.5136} \approx -0.6203 < 2.262$$

Module – 5

9.

a) Three types of fertilizers are used on three groups of plants for 5 weeks. We want to check if there is a difference in the mean growth of each group. Using the data given below apply a one-way ANOVA test at 0.05 significant level

Fertilizer 1	**6**	**8**	**4**	**5**	**3**	**4**
Fertilizer 2	**8**	**12**	**9**	**11**	**6**	**8**
Fertilizer 3	**13**	**9**	**11**	**8**	**7**	**12**

Sol^n:

To carry out the analysis of variance, we form the following tables

							Total	Squares
F1	6	8	4	5	3	4	T_1=30	T^2_1=900
F2	8	12	9	11	6	8	T_2=54	T^2_2=2916
F3	13	9	11	8	7	12	T_3=60	T^2_3=3600
Total T							144	-

The squares are as follows

							Sum of Squares
F1	36	64	16	25	9	16	166
F2	64	144	81	121	36	64	510
F3	169	81	121	64	49	144	628
Grand Total - $\sum_i \sum_j x_{ij}^2$							1304

Set the null hypotheses $H_0: \mu_1 = \mu_2 = \mu_3$

Correction Factor $CF = \frac{T^2}{N} = \frac{(144)^2}{18} = \frac{20736}{18} = 1152$

Therefore Total sum of squares $TSS = \sum_i \sum_j x_{ij}^2 - CF$

$$\Rightarrow TSS = 1304 - 1152$$
$$\Rightarrow TSS = 152$$

Sum of the squares of between the treatments $SST = \sum_i \frac{T_i^2}{n_i} - CF$

$$SST = \frac{900}{6} + \frac{2916}{6} + \frac{3600}{6} - 1152$$
$$\Rightarrow SST = 150 + 486 + 600 - 1152$$
$$\Rightarrow SST = 1236 - 1152$$
$$\Rightarrow SST = 84$$

Therefore sum of squares due to error SEE=TSS-SST

$$\Rightarrow SSE = 152 - 84 \Rightarrow SSE = 68$$

Sources variation	**d.f.**	**SS**	**MSS**	**F Ratio**
Between treatments	3-1=2	SST=84	$MST = \frac{84}{2} = 42$	$F = \frac{42}{4.533} = 9.2653$
Error	18-3=15	SSE=68	$MSE = \frac{68}{15} = 4.533$	
Total	18-1=17	-	-	

Since evaluated value 9.2653 >3.68 for F(2,15) at 5% level of significance

Hence the null hypothesis is rejected, there is significance between the tree process.

b) Present your conclusions after doing analysis of variance to the following results of the Latin-square design experiment conducted in respect of five fertilizers which were used on plots of different fertility.

A 16	**B 10**	**C 11**	**D 9**	**E 9**
E 10	**C 9**	**A 14**	**B 12**	**D 11**
B 15	**D 8**	**E 8**	**C 10**	**A 18**
D 12	**E 6**	**B 13**	**A 13**	**C 12**
C 13	**A 11**	**D 10**	**E 7**	**B 14**

Sol^n:

Given observations are

A 16	B 10	C 11	D 9	E 9
E 10	C 9	A 14	B 12	D 11
B 15	D 8	E 8	C 10	A 18
D 12	E 6	B 13	A 13	C 12
C 13	A 11	D 10	E 7	B 14

Null hypothesis H_o: There is no significant difference between rows, columns and treatment, Code the data by subtracting 10 from each value. We get,

						T	T^2
	A 6	B 0	C 1	D -1	E -1	5	25
	E 0	C -1	A 4	B 2	D 1	6	36
	B 5	D -2	E -2	C 0	A 8	9	81
	D 2	E -4	B 3	A 3	C 2	6	36
	C 3	A 1	D 0	E -3	B 4	5	25
P	16	-6	6	1	14	= 31	
P^2	256	36	36	1	196	-	-

The squares are as follows:

A 36	B 0	C 1	D 1	E 1	
E 0	C 1	A 16	B 4	D 1	
B 25	D 4	E 4	C 0	A 64	
D 4	E 16	B 9	A 9	C 4	
C 9	A 1	D 0	E 9	B 16	
74	22	30	23	86	$\sum_i \sum_j x_{ij}{}^2 = 235$

Correction Factor $CF = \frac{T^2}{N} = \frac{(31)^2}{25} = \frac{961}{25} = 38.44$

Therefore Total sum of squares $TSS = \sum_i \sum_j x_{ij}{}^2 - CF$

$$\Rightarrow TSS = 235 - 38.44$$
$$\Rightarrow TSS = 196.56$$

Sum of the row squares $SSR = \sum_i \frac{T_i{}^2}{n_i} - CF$

$$SSR = \frac{25}{5} + \frac{36}{5} + \frac{81}{5} + \frac{36}{5} + \frac{25}{5} - 38.44$$
$$\Rightarrow SSR = 5 + 7.2 + 16.2 + 7.2 + 5 - 38.44$$
$$\Rightarrow SSR = 40.60 - 38.44$$
$$\Rightarrow SSR = 2.16$$

Sum of the column squares $SSC = \sum_i \frac{P_i{}^2}{n_i} - CF$

$$SSC = \frac{256}{5} + \frac{36}{5} + \frac{36}{5} + \frac{1}{5} + \frac{196}{5} - 38.44$$
$$\Rightarrow SSC = 51.2 + 7.2 + 7.2 + 0.2 + 39.2 - 38.44$$
$$\Rightarrow SSC = 105 - 38.44 \Rightarrow SSC = 66.56$$

To find the sum of the treatments

	Observations					$Q = \sum(Observations)$	Q^2
A	6	4	8	3	1	22	484
B	0	2	5	3	4	14	196
C	1	-1	0	2	3	5	25
D	-1	1	-2	2	0	0	0
E	-1	0	-2	-4	-3	-10	100

Sum of the squares of treatments $SST = \sum_i \frac{{Q_i}^2}{n_i} - CF$

$$SST = \frac{484}{5} + \frac{196}{5} + \frac{25}{5} + \frac{0}{5} + \frac{100}{5} - 38.44$$

$$\Rightarrow SST = 96.8 + 39.2 + 5 + 0 + 20 - 38.44$$

$$\Rightarrow SST = 161 - 38.44$$

$$\Rightarrow SST = 122.56$$

$$\therefore SSE = TSS - SSR - SSC - SST$$

$$\Rightarrow SSE = 196.56 - 2.16 - 66.56 - 122.56$$

$$\Rightarrow SSE = 5.28$$

Sources variation	d.f.	SS	MSS	F Ratio	Conclusion
Rows	5-1=4	SSR=2.16	$MSR = \frac{2.16}{4} = 0.54$	$F_r = \frac{0.54}{0.44} = 1.227$	$F_r < F(4,12)$ H_0-Accepted
Columns	5-1=4	SSC=66.56	$MSC = \frac{66.56}{4} = 16.64$	$F_c = \frac{16.64}{0.44} = 37.81$	$F_c > F(4,12)$ H_0-Rejected
Treatments	5-1=4	SST=122.56	$MST = \frac{122.56}{4} = 30.64$	$F_T = \frac{30.64}{0.44} = 69.63$	$F_T > F(4,12)$ H_0-Rejected
Error	4x3=12	SSE=5.28	$MSE = \frac{5.28}{12}$=0.44	-	-
Total	25-1=24	-	-	-	-

10.

a) A trial was run to check the effects of different diets. Positive numbers indicate weight loss and negative numbers indicate weight gain. Check if there is an average difference in the weight of people following different diets using an ANOVA Table.

Low Fat	Low Calorie	Low protein	Low carbohydrate
8	2	3	2
9	4	5	2
6	3	4	-1
7	5	2	0
3	1	3	3

Sol^n:

To carry out the analysis of variance, we form the following tables

	Low Fat	Low Calorie	Low protein	Low carbohydrate	
	8	2	3	2	
	9	4	5	2	
	6	3	4	-1	
	7	5	2	0	
	3	1	3	3	
T	33	15	17	6	71
T^2	1089	225	289	36	-

The squares are as follows

	Low Fat	Low Calorie	Low protein	Low carbohydrate	
	64	4	9	4	
	81	16	25	4	
	36	9	16	1	
	49	25	4	0	
	9	1	9	9	
$\sum_i \sum_j {x_{ij}}^2$	239	55	63	18	375

Set the null hypotheses $H_0: \mu_1 = \mu_{2=}\mu_3$

Correction Factor $CF = \frac{T^2}{N} = \frac{(71)^2}{20} = \frac{5041}{20} = 252$

Therefore Total sum of squares $TSS = \sum_i \sum_j {x_{ij}}^2 - CF$

$$\Rightarrow TSS = 375 - 252$$
$$\Rightarrow TSS = 123$$

Sum of the squares of between the treatments $SST = \sum_i \frac{{T_i}^2}{n_i} - CF$

$$SST = \frac{1089}{5} + \frac{225}{5} + \frac{289}{5} + \frac{36}{5} - 252$$
$$\Rightarrow SST = 217.8 + 45 + 57.8 + 7.2 - 252$$
$$\Rightarrow SST = 327.8 - 252$$
$$\Rightarrow SST = 75.80$$

Therefore sum of squares due to error SEE=TSS-SST

$$\Rightarrow SSE = 123 - 75.80$$
$$\Rightarrow SSE = 47.2$$

Sources variation	**d.f.**	**SS**	**MSS**	**F Ratio**
Between treatments	4-1=3	SST=75.80	$MST = \frac{75.80}{3} = 25.26$	$F = \frac{25.26}{2.95} = 8.56$
Error	20-4=16	SSE=47.20	$MSE = \frac{47.20}{16} = 2.95$	
Total	20-1=19	-	-	

Since evaluated value 8.56 >3.24 for $F_{(3,16)}$ at 5% level of significance

Hence the null hypothesis is rejected, there is significance between the four process.

b) The following data show the number of worms quarantined from the GI areas of four groups of muskrats in a carbon tetrachloride anthelmintic study. Conduct a two-way ANOVA test.

I	II	III	IV
33	41	12	38
32	38	35	43
26	40	46	25
14	23	22	13
30	21	11	26

Sol^n:

Let A,B,C,D & E are the 5 treatments

Then the Given table as

	I	II	III	IV
A	33	41	12	38
B	32	38	35	43
C	26	40	46	25
D	14	23	22	13
E	30	21	11	26

Subtract 30 from all the observations, we get

	I	II	III	IV
A	3	11	-18	8
B	2	8	5	13
C	-4	10	16	-5
D	-16	-7	-8	-17
E	0	-9	-19	-4

	I	II	III	IV	P	P^2
	3	11	-18	8	4	16
	2	8	5	13	28	784
	-4	10	16	-5	17	289
	-16	-7	-8	-17	-48	2304
	0	-9	-19	-4	-32	1024
T	-15	13	-24	-5	-31	
T^2	225	169	576	25		

The squares are as follows

	I	II	III	IV	
	9	121	324	64	
	4	64	25	169	
	16	100	256	25	
	256	49	64	289	
	0	81	361	16	
$\sum_i \sum_j {x_{ij}}^2$	285	415	1030	563	2293

Set the null hypotheses $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

Correction Factor $CF = \frac{T^2}{N} = \frac{(-31)^2}{20} = \frac{961}{20} = 48$

Therefore, Total sum of squares $TSS = \sum_i \sum_j {x_{ij}}^2 - CF$

$$\Rightarrow TSS = 2293 - 48$$
$$\Rightarrow TSS = 2245$$

Sum of the squares of between the treatments $SSC = \sum_i \frac{{T_i}^2}{n_i} - CF$

$$SSC = \frac{225}{5} + \frac{169}{5} + \frac{576}{5} + \frac{25}{5} - 48$$
$$\Rightarrow SSC = 45 + 33.8 + 115.2 + 5 - 48$$
$$\Rightarrow SSC = 199 - 48$$
$$\Rightarrow SSC = 151$$

Sum of the squares of between the treatments $SSR = \sum_i \frac{{P_i}^2}{n_i} - CF$

$$\Rightarrow SSR = \frac{16}{4} + \frac{784}{4} + \frac{289}{4} + \frac{2304}{4} + \frac{1024}{4} - 48$$
$$\Rightarrow SSR = 4 + 196 + 72.25 + 576 + 256 - 48$$
$$\Rightarrow SSR = 1104.25 - 48$$
$$\Rightarrow SSR = 1056.25$$

Therefore sum of squares due to error SEE=TSS-SSC-SSR

$$\Rightarrow SSE = 2245 - 151 - 1056.25$$
$$\Rightarrow SSE = 1037.75$$

We know that $F(4,12) = 3.26$ & $F(3,12) = 3.49$

Sources variation	d.f.	SS	MSS	F Ratio	Conclusion
Rows	5-1=4	SSR=1056.25	$MSR = \frac{1056.25}{4} = 264.06$	$F_r = \frac{264.06}{86.48} = 3.053$	$F_r < F(4,12)$ $H_0 : Accepted$
Columns	4-1=3	SSC=151	$MSC = \frac{151}{3} = 50.33$	$F_c = \frac{86.48}{50.33} = 1.718$	$F_r < F(3,12)$ $H_0 : Accepted$
Error	4X3=12	SSE=1037.75	$MSE = \frac{1037.75}{12}$=86.48	-	-

***

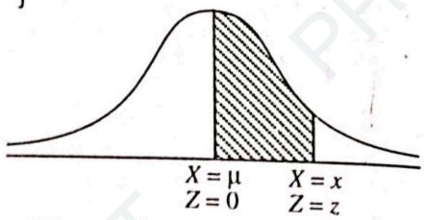
TABLE VI : AREAS UNDER STANDARD NORMAL PROBABILITY CURVE

Normal Probability curve is given by : $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}; -\infty < x < \infty$

and standard normal probability curve is given by :

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right), -\infty < z < \infty$$

where $$Z = \frac{X - E(X)}{\sigma_x} = \frac{X-\mu}{\sigma} \sim N(0, 1)$$



The following table gives the shaded area in the diagram, *viz.* $P(0 < Z < z)$ for different values of z.

AREAS UNDER STANDARD NORMAL PROBABILITY CURVE

z \ z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

NUMERICAL TABLES

T-11

TABLE VIII : CRITICAL VALUES OF STUDENT'S t-DISTRIBUTION

d.f.	LEVEL OF SIGNIFICANCE FOR ONE-TAILED TEST						
	.25	.10	.05	.025	.01	.005	.0005
	LEVEL OF SIGNIFICNCE FOR TWO-TAILED TEST						
υ	.50	.20	.10	.05	.02	.01	.001
1	1.000	3.078	6.314	12.706	31.821	63.657	636.619
2	.816	1.886	2.920	4.303	6.965	9.925	31.599
3	.765	1.638	2.353	3.182	4.541	5.841	12.924
4	.741	1.533	2.132	2.776	3.747	4.604	8.610
5	.727	1.476	2.015	2.571	3.365	4.032	6.869
6	.718	1.440	1.943	2.447	3.143	3.707	5.959
7	.711	1.415	1.895	2.365	2.998	3.499	5.408
8	.706	1.397	1.860	2.306	2.896	3.355	5.041
9	.703	1.383	1.833	2.262	2.821	3.250	4.781
10	.700	1.372	1.812	2.228	2.764	3.169	4.587
11	.697	1.363	1.796	2.201	2.718	3.106	4.437
12	.695	1.356	1.782	2.179	2.681	3.055	4.318
13	.694	1.350	1.771	2.160	2.650	3.012	4.221
14	.692	1.345	1.761	2.145	2.624	2.977	4.140
15	.691	1.341	1.753	2.131	2.602	2.947	4.073
16	.690	1.337	1.746	2.120	2.583	2.921	4.015
17	.689	1.333	1.740	2.110	2.567	2.898	3.965
18	.688	1.330	1.734	2.101	2.552	2.878	3.922
19	.688	1.328	1.729	2.093	2.539	2.861	3.883
20	.687	1.325	1.725	2.086	2.528	2.845	3.850
21	.686	1.323	1.721	2.080	2.518	2.831	3.819
22	.686	1.321	1.717	2.074	2.508	2.819	3.792
23	.685	1.319	1.714	2.069	2.500	2.807	3.768
24	.865	1.318	1.711	2.064	2.492	2.797	3.745
25	.684	1.316	1.708	2.060	2.485	2.787	3.725
26	.684	1.315	1.706	2.056	2.479	2.779	3.707
27	.684	1.314	1.703	2.052	2.473	2.771	3.690
28	.683	1.313	1.701	2.048	2.467	2.763	3.674
29	.683	1.311	1.699	2.045	2.462	2.756	3.659
30	.683	1.310	1.697	2.042	2.457	2.750	3.646
40	.681	1.303	1.684	2.021	2.423	2.704	3.551
60	.679	1.296	1.671	2.000	2.390	2.660	3.460
120	.677	1.289	1.658	1.980	2.358	2.617	3.373
∞	.674	1.282	1.645	1.960	2.326	2.576	3.291

TABLE IX : CRITICAL VALUES OF THE VARIANCE-RATIO F-DISTRIBUTION (RIGHT TAIL AREAS) – 5 PER CENT POINTS

υ_2 \ υ_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39

υ_1 : Degrees of freedom for the numerator.

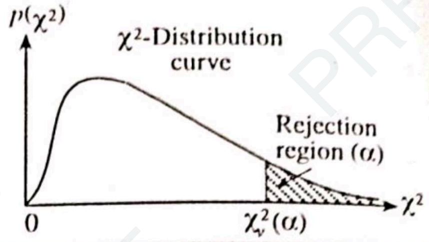
υ_2 : Degrees of freedom for the denominator.

TABLE VII : CRITICAL VALUES $\chi_\upsilon^2(\alpha)$ OF CHI-SQUARE DISTRIBUTION (RIGHT TAIL AREAS) FOR GIVEN PROBABILITY α,

where

$$p\left[\chi^2 > \chi_\upsilon^2(\alpha)\right] = \alpha$$

AND υ IS DEGREES OF FREEDOM (*d.f.*)



Degrees of Freedom (υ)	Probability (α)							
	.99	.975	.95	.90	.10	.05	.025	.01
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086
6	0.872	1.237	1.634	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217
13	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	8.231	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	11.651	27.204	30.144	23.852	36.191
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	11.591	13.240	29.615	23.671	35.479	38.932
22	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	33.196	36.415	39..364	42.980
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642
27	12.879	14.573	16.151	18.114	36.741	40.113	43.194	46.963
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892

For large values of υ, the quantity $\sqrt{2\chi^2} - \sqrt{2\nu - 1}$ may be used as a standard normal variable.