Deterministic & Stochastic Statistical Methods
(20AOE9925)

Lecture Notes

III –BTECH

Prepared by

Dr.B.Aruna Kumari

Department of Humanities & Basic Sciences

ANNAMACHARYA INSTITUTE OF TECHNOLOGY AND SCIENCES

Approved by AICTE, New Delhi & Permanent Affiliation to JNTUA, Anantapuramu
Venkatapuram(Village), Renigunta(Mandal), Tirupati, Chittoor District,
Andhra Pradesh-517520
Phone: +91-0877-2285609, Fax: +91-0877-2285608,
Mobile: +91-9948149222, +919948661276

Course Code	Deterministic & Stochastic Statistical Methods	L	T	P	C
20AOE9925		3	0	0	3

Course Objectives

Study of various Mathematical Methods and Statistical Methods which is needed for Artificial Intelligence, Machine Learning, and Data Science and also for Computer Science and engineering problems.

Course outcomes (CO) : After completion of the course, the student can able to

CO-1: Apply logical thinking to problem-solving in context.

CO-2: Employ methods related to these concepts in a variety of data science applications.

CO-3: Use appropriate technology to aid problem-solving and data analysis.

CO-4: The Bayesian process of inference in probabilistic reasoning system.

CO-5: Demonstrate skills in unconstrained optimization.

Syllabus

UNIT - I- Data Representation

Distance measures, Projections, Notion of hyper planes, half-planes. Principal Component Analysis- Population Principal Components, sample principal coefficients, covariance, matrix of data set, Dimensionality reduction, Singular value decomposition, Gram Schmidt process.

UNIT - II - Single Variable Distribution

Random variables (discrete and continuous), probability density functions, properties, mathematical expectation Probability distribution - Binomial, Poisson approximation to the binomial distribution and normal distribution their properties-Uniform distribution-exponential distribution.

UNIT III- Stochastic Processes And Markov Chains:

Introduction to Stochastic processes- Markov process. Transition Probability, Transition Probability Matrix, First order and Higher order Markov process, step transition probabilities, Markov chain, Steady state condition, Markov analysis.

UNIT IV- Multivariate Distribution Theory

Multivariate Normal distribution Properties, Distributions of linear combinations, independence, marginal distributions, conditional distributions, Partial and Multiple correlation coefficient. Moment generating function. BAYESIAN INFERENCE AND ITS APPLICATIONS: Statistical tests and Bayesian model comparison, Bit, Surprisal, Entropy, Source coding theorem, Joint entropy, Conditional entropy, Kullback- Leibler divergence.

UNIT V- Optimization

Unconstrained optimization, Necessary and sufficiency conditions for optima, Gradient descent methods, Constrained optimization, KKT conditions, Introduction to non-gradient techniques, Introduction to least squares optimization, Optimization view of machine learning. Data Science Methods: Linear regression as an exemplar function approximation problem, linear classification problems.

Textbooks:
1. Mathematics for Machine Learning by A. Aldo Faisal, Cheng Soon Ong, and Marc Peter Deisenroth 2. Dr.B.S Grewal, Higher Engineering Mathematics, 45th Edition, Khanna Publishers. 3. Operations Research, S.D. Sharma
Reference Books:
1. Operations Research, An Introduction, Hamdy A. Taha, Pearson publishers. 2. A Probabilistic Theory of Pattern Recognition by Luc Devroye,. Laszlo Gyorfi, Gabor Lugosi.

①

Distance Measures

Many algorithms whether supervised (or) unsupervised make use of distance Measures

These measures such as Euclidean distance (or) Cosine Similarity Can often be found in algorithms Such as K-NN, UMAP, HDBSCAN etc.

Understanding the field of distance measure is more importance than you might realize.

Distance measures plays an important role in machine learning

They Provide the foundation for many Popular and effective machine learning algorithms like K- nearest neighbours for supervised learning and K- means clustering for unsupervised learning.

"Knowing when to use which distance Can help you go from a Poor classifier to an accurate model"

There are many distance measures which ~~we~~ explore how and when they best Can be used.

Some of the main distance measures are follows below.

(1) Euclidean distance

(2) Manhattan distance

(3) Minkowski distance

(4) Cosine Index (or) Cosine Similarity

(5) Hamming distance.

(6) Chebyshev distance.

(7) Jaccard Index

2

(1) Euclidean distance :-

Euclidean distance is the distance between two points (or) the straight line distance.

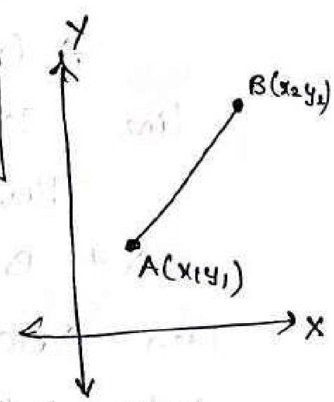
To find the two points on a plane, the length of a segment connecting the two points is measured.

we derive the Euclidean distance formula by using the pythagoras theorem.

Euclidean distance formula :-

Let us assume that (x_1, y_1) & (x_2, y_2) are the two points in a two-dimensional plane. Then the Euclidean distance formula is

$$\boxed{d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}}$$



where

(x_1, y_1) are Co-ordinates of One point

(x_2, y_2) are Co-ordinates of other point

d is the distance between $(x_1 y_1)$ & (x_2, y_2)

(1) What is Euclidean distance formula ?

(A) The Euclidean distance formula is used to find the distance between two points on a plane.

This formula says the distance between the two points (x_1, y_1) & (x_2, y_2) is

$$\boxed{d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}}$$

(2) How to derive Euclidean distance formula ?

(A) To derive the Euclidean distance formula Consider the two points

$A(x_1, y_1)$ & $B(x_2, y_2)$ and join them by a line segment.

Then draw horizontal & vertical lines from A to B to meet at C.

Then ABC is a Right angled Δle and hence we can apply pythagoras theorem to it

Then we get $AB^2 = AC^2 + BC^2$

$$d^2 = (x_2-x_1)^2 + (y_2-y_1)^2$$

Taking square root on both sides

$$\boxed{d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}}$$

3

(3) What are the applications of Euclidean distance formula ?

A) The Euclidean distance formula is used to find the length of a line segment given two points on a plane.

Finding distance helps in proving the given vertices form a Square, Rectangle, etc (or) Proving given vertices form an equilateral Δle Right angled Δle etc.

(4) What is the difference between Euclidean distance formula & Manhattan distance formula.

Sol:- For any two points (x_1, y_1) & (x_2, y_2) on a plane

→ The Euclidean distance formula says, the distance between the above points is $\boxed{d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}}$

→ The Manhattan distance formula says, the distance between the above points is $\boxed{d = |x_2 - x_1| + |y_2 - y_1|}$

4

$$d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$$

Hence the Euclidean distance formula is derived.

Problems.

(1) Find the distance between points $P(3,2)$ & $Q(4,1)$

Sol : given $P(\underset{x_1}{3},\underset{y_1}{2})$ $Q(\underset{x_2}{4},\underset{y_2}{1})$

Using Euclidean distance formula we have

$$d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$$

$$PQ = \sqrt{(4-3)^2 + (1-2)^2}$$

$$= \sqrt{(1)^2 + (-1)^2}$$

$$PQ = \sqrt{2} \text{ units}$$

$\therefore$ The Euclidean distance between points $A(3,2)$ $B(4,1)$ is $\sqrt{2}$ units.

(2) Prove that points $A(0,4)$ $B(6,2)$ & $C(9,1)$ are Collinear

Sol : To Prove the given three points to be Collinear it is sufficient to prove that the sum of the distances between two pairs of points is equal to the distance between the third pair.

now we will find the distance between every pair of points Using the Euclidean distance formula.

$$AB = \sqrt{(6-0)^2 + (2-4)^2}$$
$$= \sqrt{(6)^2 + (-2)^2}$$
$$= \sqrt{36+4} = \sqrt{40} = \underline{\underline{2\sqrt{10}}}.$$

$$BC = \sqrt{(9-6)^2 + (1-2)^2}$$
$$= \sqrt{(3)^2 + (-1)^2} = \sqrt{9+1} = \underline{\underline{\sqrt{10}}}.$$

$$CA = \sqrt{(0-9)^2 + (4-1)^2}$$
$$= \sqrt{(9)^2 + (3)^2}$$
$$= \sqrt{81+9} = \sqrt{90} = \underline{\underline{3\sqrt{10}}}.$$

Here we can see that

$AB + BC = CA$.

$\underline{\underline{2\sqrt{10} + \sqrt{10} = 3\sqrt{10}}}$.

$\therefore$ we proved that $\underline{\underline{A, B, C}}$ are collinear.

(3) Check that points $A(\sqrt{3}, 1)$, $B(0,0)$, $C(2,0)$ are the vertices of an Equilateral Δ^{le}.

Sol :- Three vertices A, B & C are vertices of an equilateral Δ^{le} $\Longleftrightarrow$ if $AB = BC = CA$.

given $A(\sqrt{3}, 1)$ $B(0,0)$ $C(2,0)$
$\quad\quad\; x_1 \; y_1 \quad\quad x_2 \; y_2 \quad\quad x_3 \; y_3$

Using Euclidean distance formula.

5

$$AB = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2}$$
$$= \sqrt{(0-\sqrt{3})^2+(0-1)^2}$$
$$= \sqrt{3+1} = \sqrt{4} = \underline{\underline{2}}.$$

$$BC = \sqrt{(x_3-x_2)^2+(y_3-y_2)^2}$$
$$= \sqrt{(2-0)^2+(0-0)^2}$$
$$= \sqrt{4+0} = \sqrt{4} = \underline{\underline{2}}.$$

$$CA = \sqrt{(x_3-x_1)^2+(y_3-y_1)^2}$$
$$= \sqrt{(2-\sqrt{3})^2+(0-1)^2}$$
$$= \sqrt{(9+25)}$$
$$= \sqrt{34}$$

Here $AB = BC \neq CA$.

$\therefore$ A, B & C are not the vertices of an equilateral $\triangle$le.

***

4) Difference between Euclidean Distance formula and Manhatton Distance formula. ?

For any two points $(x_1 y_1)$ & $(x_2 y_2)$ on a plane

(1) The Euclidean distance formula says, the distance between the above Points is $d = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2}$

(2) The Manhattan distance formula says, the distance between the above Points is $d = |x_2-x_1| + |y_2-y_1|$

5) Calculate the Euclidean distance between ~~Skoopsin~~ ~~A1B2A~~ ~~where~~ the points (1,1,0) & (4,5,0) A in XY plane.

<u>Sol</u>: distance between points

$(1,1,0)$ $(4,5,0)$
$x_1 y_1 z_1$ $x_2 y_2 z_2$

$$d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$$

$$= \sqrt{(4-1)^2 + (5-1)^2}$$

$$= \sqrt{(3)^2 + (4)^2} = \sqrt{9+16} = \sqrt{25} = \underline{\underline{5 \text{ Units}}}$$

6) Calculate the distance between the two points

$A(-5,2,4)$ & $B(-2,2,0)$
$x_1 y_1 z_1$ $x_2 y_2 z_2$.

$$d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}$$

$$= \sqrt{(-2+5)^2 + (2-2)^2 + (0-4)^2}$$

$$= \sqrt{(3)^2 + (0)^2 + (-4)^2}$$

$$= \sqrt{9+16} = \sqrt{25} = \underline{5 \text{ Units}}.$$

(6)

7) The distance between $(1,2,3)$ & $(4,5,6)$ will be

Sol :- $(1,2,3)$ $(4,5,6)$

$x_1\, y_1\, z_1 \quad x_2\, y_2\, z_2$

$$D = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}$$

$$= \sqrt{(4-1)^2+(5-2)^2+(6-3)^2}$$

$$= \sqrt{(3)^2+(3)^2+(3)^2}$$

$$= \sqrt{9+9+9} = \sqrt{27} \text{ Units}$$

(8) The distance between $P(-1,2,3)$ & $(4,0,-3)$ is

Sol :- $(-1,2,3)$ $(4,0,-3)$

$x_1\, y_1\, z_1 \quad x_2\, y_2\, z_2$

$$D = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}$$

$$= \sqrt{(4+1)^2+(0-2)^2+(-3-3)^2}$$

$$= \sqrt{(5)^2+(-2)^2+(-6)^2}$$

$$= \sqrt{25+4+36} = \sqrt{65} \text{ units}$$

(9) The distance of the point $(5,0,12)$ from the Origin $(0,0,0)$ is

Sol :- $(5,0,12)$ $(0,0,0)$

$x_1\, y_1\, z_1 \quad x_2\, y_2\, z_2$

$$D = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2} = \sqrt{\cancel{(5-0)^2}}$$

$(5, 0, 12)$. $(0\ 0\ 0)$
$x_1\ y_1\ z_1 \quad x_2\ y_2\ z_2$

$$D = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}$$

$$= \sqrt{(0-5)^2 + (0-0)^2 + (0-12)^2}$$

$$= \sqrt{(-5)^2 + (0)^2 + (-12)^2}$$

$$= \sqrt{25+144} = \sqrt{169} = \underline{13 \text{ units}}$$

$$\begin{array}{r} 13 \\ 13 \\ \hline 39 \\ 13\times \\ \hline 169 \end{array}$$

10) The distance between $A(2,-1)$ & $B(2,3)$ is

Sol:- $(2,-1)\ (2,3)$
$x_1\ y_1 \quad x_2\ y_2$

$$D = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$$

$$= \sqrt{(2-2)^2 + (3+1)^2}$$

$$= \sqrt{(0)^2 + (4)^2} = \sqrt{16} = \underline{4 \text{ units}}$$

<u>Manhattan</u> <u>distance</u> :- Formula is.

For 2 vectors $\boxed{d = |x_1 - x_2| + |y_1 - y_2|}$.

For 3 vectors $\boxed{d = |x_1 - x_2| + |y_1 - y_2| + |z_1 - z_2|}$

7

(d) Cosine - Correlation distance :-

$$\boxed{\cos(A,B) = \frac{A \cdot B}{\|A\| \cdot \|B\|}}$$

$A \cdot B = (1)(2) + (0)(1) + (2)(0) + (5)(3) + (3)(-1)$

$A \cdot B = \underline{14}$.

$\|A\| = \sqrt{1^2 + 0^2 + 2^2 + 5^2 + 3^2} = \underline{6.24}$

$\|B\| = \sqrt{2^2 + 1^2 + 0^2 + 3^2 + (-1)^2} = \underline{3.87}$.

$\therefore \cos(A,B) = \frac{A \cdot B}{\|A\| \cdot \|B\|} = \frac{14}{(6.24)(3.87)} = \underline{0.57}$.

~~**Distance Measure :-**~~

~~Distance measures play an important role in machine learning. They Provide the Foundation for many Popular and effective machine learning algorithms like K-nearest neighbours for supervised learning and K-Means Clustering for unsupervised learning.~~

---

(4) Cosine - Correlation distance :-

$$\boxed{Cos(A,B) = \frac{A \cdot B}{||A|| \cdot ||B||}}$$

$A = (1, 0, 2, 5, 3)$

$B = (2, 1, 0, 3, -1)$

$A \cdot B = (1)(2) + (0)(1) + (2)(0) + (5)(3) + (3)(-1) = \underline{\underline{14}}.$

$||A|| = \sqrt{1^2 + 0^2 + 2^2 + 5^2 + 3^2} = \underline{\underline{6.24}}$

$||B|| = \sqrt{2^2 + 1^2 + 0^2 + 3^2 + (-1)^2} = \underline{\underline{3.87}}.$

$\therefore Cos(A,B) = \frac{14}{6.24 * 3.87} = \underline{\underline{0.57}}.$

→ Cosine distance measure for clustering determines the Cosine of the angle between two vectors given by the formula $Cos(A,B) = \frac{A \cdot B}{||A|| \cdot ||B||}$

8

1) $X_1 = (1, 2, 2)$ with $x_1\ y_1\ z_1$ $\quad X_2 = (2, 5, 3)$ with $x_2\ y_2\ z_2$

Manhattan (L_1) :-

$$L_1 = |1-2| + |2-5| + |2-3|$$

$$= |-1| + |-3| + |-1|$$

$$= 1 + 3 + 1 = \underline{\underline{5}}.$$

Euclidean :- (L_2) :-

$$L_2 = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}$$

$$= \sqrt{(2-1)^2 + (5-2)^2 + (3-2)^2}$$

$$= \sqrt{(1)^2 + (3)^2 + (1)^2}$$

$$= \sqrt{1+9+1}$$

$$= \underline{\underline{\sqrt{11}}} \text{ Units}.$$

(1) Given 5 Dimensional Samples

$A = (1, 0, 2, 5, 3)$

$B = (2, 1, 0, 3, -1)$ then Find.

(1) Euclidean distance between Points :- $\boxed{\sqrt{(A_K - B_K)^2}}$

$$= \sqrt{(1-2)^2 + (0-1)^2 + (2-0)^2 + (5-3)^2 + (3+1)^2}$$

$$= \underline{5.09}.$$

(2) City block / Manhattan distance :-

$$\boxed{\sum_{K=1}^{n} |x_{iK} - x_{jK}|}$$

$$d_{AB} = |1-2| + |0-1| + |2-0| + |5-3| + |3+1|$$

$$= |-1| + |-1| + |2| + |2| + |4|$$

$$= 1 + 1 + 2 + 2 + 4$$

$$= \underline{10}$$

(3) Minkowski distance :-

given external variable $\boxed{P = 3}$

$$= \left[\sum |A_K - B_K|^3 \right]^{1/3}$$

$$= \left[|1-2|^3 + |0-1|^3 + |2-0|^3 + |5-3|^3 + (3+1)^3 \right]^{1/3}$$

$$= \underline{4.34}.$$

$$\boxed{\left[\sum |A_K - B_K|^P \right]^{1/P}}$$

(9)

(2) Manhattan distance :-

This determines the absolute difference among the Pair of the Coordinates.

Suppose we have two points P and Q to determine the distance between these Points we simply have to Calculate the Perpendicular distance of the points from X-axis & Y-axis In a plane with P at Coordinate $(x_1\ y_1)$ and Q at $(x_2\ y_2)$

Manhattan distance between P & Q is.

$$\boxed{d = |x_2 - x_1| + |y_2 - y_1|}$$

The Manhattan distance, often Called as "Taxicab distance" (or) City Block distance. Calculates the distance between real-Valued vectors.

Imagine vectors that describe objects on a uniform grid such as a Chessboard.

Manhattan distance then refers to the distance between two vectors If they Could only move right angles. There is no diagonal movement involved in Calculating the distance.

The Manhattan distance between two points $(x_1\ y_1)$ & $(x_2\ y_2)$ is given as

$$|x_1 - x_2| + |y_1 - y_2| \quad \text{(or)} \quad |x_2 - x_1| + |y_2 - y_1|$$

(1) Find the Manhattan distance between the points given below

(1) $\underset{x_1\ y_1}{(1, 2)}\ \underset{x_2\ y_2}{(3, 4)}$

$\Rightarrow |3-1| + |4-2|$

$\Rightarrow 2+2 = \underline{\underline{4}}$

(2) $\underset{x_1\ y_1}{(-4, 6)}\ \underset{x_2\ y_2}{(3, -4)}$

Manhattan.

$\Rightarrow |3+4| + |-4-6|$

$\Rightarrow |7| + |-10|$

$\Rightarrow 7+10 = \underline{\underline{17}}$.

$\Rightarrow$ Manhattan distance is the most Preferable for high dimensional applications.

Thus Manhattan distance is Preferred over the Euclidean distance metric as the dimension of the data increases.

$\Rightarrow$ If we need to calculate the distance between two data points in a grid-like path we use Manhattan.

(10)

(3) Calculate the Manhattan distance from the Points given below

$X_1 = (1, 2, 3, 4, 5, 6)$

$X_2 = (10, 20, 30, 1, 2, 3)$

$\Rightarrow |10-1| + |20-2| + |30-3| + |1-4| + |2-5| + |3-6|$

$\Rightarrow 9 + 18 + 27 + 3 + 3 + 3$

$\Rightarrow \underline{\underline{63}}$.

(3) <u>Minkowski distance</u> :-

Minkowski distance is a distance measured between two points in N- dimensional space.

It is basically a generalization of the Euclidean distance and Manhattan distance:

It is widely used in field of machine Learning especially in the Concept to find the optimal Correlation or Classification of data

Minkowski distance is used in Certain algorithms like K- Nearest Neighbors, LVQ. (Learning Vector Quantization), SOM (self Organizing Map) and K- Means clustering

→ Let us Consider 2-dimensional space having three Points

$$P_1(x_1 y_1),\ P_2(x_2 y_2)\ \ P_3(x_3 y_3)$$

The Minkowski distance is given by

$$\left(|x_1-y_1|^p + |x_2-y_2|^p + |x_3-y_3|^p \right)^{1/p}$$

(or)

The formula for <u>Minkowski distance</u> is given as

$$D(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}}$$

Most interestingly about this distance measure is use of Parameter P.

we can use this Parameter to manipulate the distance metrics to closely resemble Others.

<u>Common values of p are</u> :-

(1) $P = 1 \Longrightarrow$ Manhattan distance.

(2) $P = 2 \Longrightarrow$ Euclidean distance

(3) $P = \infty \Longrightarrow$ Chebyshev distance.

11

Problem

(1) Given 5 dimensional samples

$A = (1, 0, 2, 5, 3)$
$B = (2, 1, 0, 3, -1)$ & external variable $P = 3$.

(or) parameter $P = 3$

Sol :- $D(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^{P} \right)^{\frac{1}{P}}$

$= \left[|1-2|^3 + |0-1|^3 + |2-0|^3 + |5-3|^3 + |3+1|^3 \right]^{1/3}$

Here $P = 3$

$= \underline{4.34}$

(2) we are given two vectors vector A & vector B as
$A = (4, 2, 6, 8)$
$B = (5, 1, 7, 9)$ Find Minkowski distance for $\boxed{P = 2}$

Sol :- $\left[|4-5|^2 + |2-1|^2 + |6-7|^2 + |8-9|^2 \right]^{1/2}$

$= \underline{2}$.

(3) Calculate the Minkowski distance between two vectors using a power of $\boxed{P=3}$

$A = (2, 4, 4, 6)$

$B = (5, 5, 7, 8)$

Ans :- 3.979057

(4) $A = (2, 4, 4, 6)$

$B = (5, 5, 7, 8)$

$C = (9, 9, 9, 8)$

$D = (4, 2, 3, 3)$

Note :- Each vector in the matrix should be the same length.

The Minkowski distance between A & B is 3.98.

The Minkowski distance between A & C is 8.43

The Minkowski distance between A & D is 3.33

The Minkowski distance between

B & C is 5.14

B & D is 6.54

C & D is 10.61.

- **Problem :-**

1) Given two Objects represented by the tuples $(22, 1, 42, 10)$ and $(20, 0, 36, 8)$

   a) Compute the Euclidean distance between two Objects

   b) Compute the Manhattan distance between two Objects

   c) Compute the Minkowski distance between the two Objects using $p=3$

Sol :- a) **Euclidean distance :-**

$(22, 1, 42, 10)$ $(20, 0, 36, 8)$

$$= \sqrt{|22-20|^2 + |1-0|^2 + |42-36|^2 + |10-8|^2}$$

$$= \underline{6.71}.$$

b) **Manhattan distance :-**

$(22, 1, 42, 10)$ $(20, 0, 36, 8)$

$$= |22-20| + |1-0| + |42-36| + |10-8|$$

$$= \underline{11}.$$

(c) **Minkowski distance :-**

$$\left(|22-20|^3 + |1-0|^3 + |42-36|^3 + |10-8|^3\right)^{1/3}$$

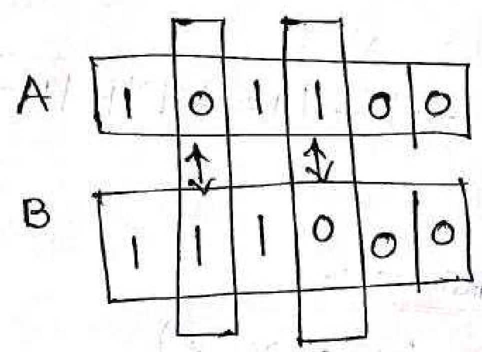
$$= \underline{6.15}$$

(4) Hamming distance :-

Hamming distance is the number of values that are different between two vectors.

It is typically used to Compare two binary strings of equal length

It can also be used for strings to Compare how Similar they are to each other by Calculating the number of Characters that are different from each other.



(13)

- (1) Find the Hamming distance between the Code words of
$C = \{(0000), (0101) (1011) (0111)\}$

Sol :- Let

$x = 0\ 0\ 0\ 0$

$y = 0\ 1\ 0\ 1$

$z = 1\ 0\ 1\ 1$

$w = 0\ 1\ 1\ 1$

$d(x, y) = \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{matrix} = \underline{\underline{2}}$

$d(x, z) = \begin{matrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{matrix} = \underline{\underline{3}}$

$d(x, w) = \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{matrix} = \underline{\underline{3}}$

$d(y, z) = \begin{matrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{matrix} = \underline{\underline{3}}$

$d(y, w) = \begin{matrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{matrix} = \underline{\underline{1}}$

$d(z, w) = \begin{matrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{matrix} = \underline{\underline{2}}$

(5) Chebyshev distance :-

Chebyshev distance is defined as the greatest of difference between two vectors along any coordinate dimension.

In Otherwords, it is Simply the maximum distance along One axis

Due to its nature, it is often refered as Chessboard distance since the minimum number of moves needed by a King to go from one square to another is equal to Chebyshev distance.

$$\boxed{D(x,y) = \max_i \left(|x_i - y_i| \right)}$$

→ Consider two Points P_1 & P_2 with Coordinates as follows

$$P_1 = (P_1, P_2, P_3 \cdots PN)$$
$$P_2 = (q_1, q_2, q_3 \cdots q_N)$$

Then the Chebyshev distance between the two Points P_1 & P_2 is

$$\text{Chebyshev distance} = \max(|P_i - q_i|)$$

14

- 1) The Point A has Coordinate (0, 3, 4, 5) and Point B has Coordinate (7, 6, 3, -1)

The Chebyshev distance between Point A & B is

$$d_{AB} = Max\{|0-7|, |3-6|, |4-3|, |5+1|\}$$

$$= Max\{7, 3, 1, 6\} = \underline{\underline{7}}.$$

2) $distance\ (A,B) = Max(|x_A - x_B|, |y_A - y_B|)$

$distance\ (A,B) = Max\ (|70-330|), |40-220|)$

$distance\ (A,B) = max\ (|-260|, |-188|)$

$distance\ (A,B) = max\ (260, 188)$

$$\boxed{distance\ (A,B) = 260.}$$

(6) Jaccard Index :-

The Jaccard distance measures the Similarity of the two data set items as the Intersection of those items divided by the Union of the data items

$$\boxed{J(A,B) = \frac{|A \cap B|}{|A \cup B|}}$$

where, J = Jaccard distance
A = Set-1
B = Set-2

→ To Calculate the Jaccard distance we Simply subtract the Jaccard index from '1'

$$\boxed{D(x,y) = 1 - \frac{|A \cap B|}{|A \cup B|}}$$

(15)

Hyperplane, Subspace & Halfspace

(1) Hyperplane :-

Geometrically, a hyperplane is a geometric entity whose dimension is one less than that of its ambient space.

what does it mean ?

It means the following

For example ;

If you take the 3D space then hyperplane is a geometric entity that is '1' dimensionless so its going to be 2 dimensions and a 2 dimensional entity in a 3D space would be a plane.

Now If you take 2 dimensions, then '1' dimensionless would be a single-dimensional geometric entity, which would be a line and so on.

(1) The hyperplane is usually described by an equation as follows

$$\boxed{X^T n + b = 0}$$

(2) If we expand this out for 'n' Variables we will get Something like this

$$x_1 n_1 + x_2 n_2 + x_3 n_3 + \cdots + x_n n_n + b = 0$$

(3) In just two dimensions we will get Something like this which is nothing but an equation of a line

$$\boxed{x_1 n_1 + x_2 n_2 + b = 0.}$$

Ex:- Let us Consider a 2D geometry with

$$n = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \& \quad b = 4$$

Though its a 2D geometry the Value of X will be $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

So according to the equation of hyperplane it Can be Solved as

$$X^T n + b = 0$$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 4 = 0$$

$$\boxed{x_1 + 3x_2 + 4 = 0}$$

So as you can see from the Solution the hyperplane is the equation of a line.

16

(2) <u>Subspace</u> :-

- Hyper-planes, in general, are not subspaces. However, If we have hyper-planes of the form

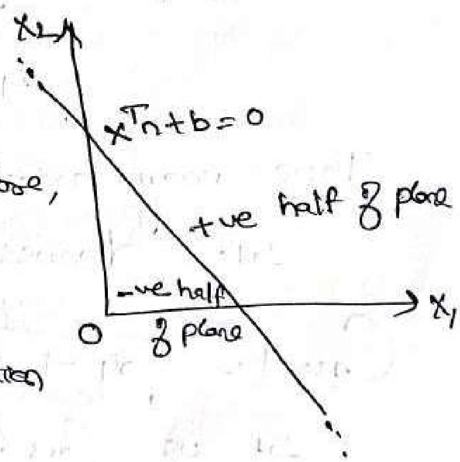
$$\boxed{x^T n = 0}$$

That is if the plane goes through the Origin then a <u>hyperplane</u> also becomes a <u>Subspace</u>.

(3) <u>Half-space</u> :-

Consider this 2-dimensional Picture given below

So here we have a 2-dimensional space in X_1 & X_2 and as we have discussed before, an equation in two dimensions would be a line which would be a hyperplane. So the equation to the line is written as



$$\boxed{x^T n + b = 0.}$$

So, for this two dimensions, we could write this line as we discussed Previously

$$\boxed{X_1 n_1 + X_2 n_2 + b = 0.}$$

You can notice from the above graph that this whole two-dimensional space is broken into two spaces.

One on this side (+ve half of plane) of a line and the other one on this side (-ve half of the plane) of a line. Now these two spaces are called as Half-spaces.

Example :- Let us consider the same example that we have taken in hyperplane case.

So by solving, we got the equation as

$$\boxed{x_1 + 3x_2 + 4 = 0}$$

There may arise 3 cases.

Let's discuss each case with an example.

Case-1 :- $\boxed{x_1 + 3x_2 + 4 = 0} \rightarrow$ On the line.

Let us consider two points $(-1,-1)$, when we put this value on the equation of line we got '0' so we can say that this point is on the hyperplane of the line.

(17)

- **Case - 2** :- Ily $\boxed{x_1 + 3x_2 + 4 > 0} \rightarrow$ Positive Half space.

Consider two points $(1, -1)$, when we put this value on the equation of line we got '2' which is greater than '0' so we can say that this point is on the Positive Half space.

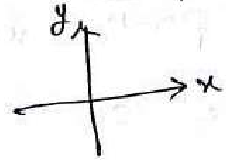
Case - 3 :- Ily $\boxed{x_1 + 3x_2 + 4 < 0} \rightarrow$ Negative Half-space

Consider two points $(1, -2)$, when we put this value on the equation of line we got -1 which is less than '0' so we can say that this point is on the Negative Half space.

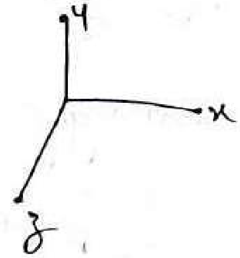
(18)

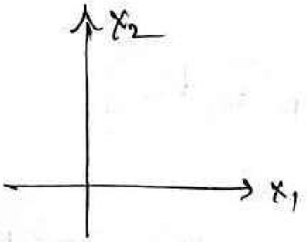
Hyper plane :-

2D space



3D space





2D.

x_1 & x_2 hyperplane is $\boxed{c_1x_1 + c_2x_2 = K}$ (x_1) (x_2) value set or Collection of hyperpl

x_1 & x_2, x_3 hyper plane is $\boxed{c_1x_1 + c_2x_2 + c_3x_3 = K}$ x_1 x_2 x_3

satisfying x_1 x_2 x_3 value set or Collection is called as hyper plane.

Hyperplane :-

In R^n (ie:- n-dimensional space) the set of Points $x = (x_1 x_2 \ldots x_n)$ satisfying the equation

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = K \longrightarrow (1)$$

(not all $c_i = 0$)

is Called a hyper plane for given value of c's

→ As Particular case in R^3 (ie - 3-dimensional space) the set of Points $x = (x_1 x_2 x_3)$ satisfying the equation $\boxed{c_1x_1 + c_2x_2 + c_3x_3 = K}$.

→ As Particular Case in $\mathbb{R}^4$ (ie: 4 dimensional space)

the set of points $x = (x_1, x_2, x_3, x_4)$

Satisfying the equation

$$\boxed{c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 = K}$$

Notes :- In a Linear Programming Problem the Objective function and the Constraints equations represents the hyperplanes

Notes :- If $\boxed{K=0}$ then the hyperplane is said to passes through the Origin, and then its equation can be written in the form

$$\boxed{c_1x_1 + c_2x_2 + c_3x_3 + \cdots + c_nx_n = 0}$$

Notes :- In matrix notation the equation of the hyperplane ① can be written as $\boxed{CX = K}$

where C is row vector

$$C = [c_1 \, c_2 \, \cdots \, c_n]$$

and X is Column vector

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1 \, x_2 \, \cdots \, x_n]$$

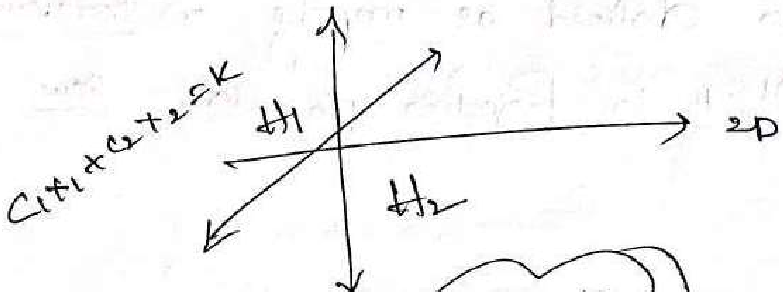
and K is a Constant matrix

(19)

→ If the hyperplane passes through the Origin then its equation is $\boxed{CX = 0}$

→ If a hyperplane divides R^n into two half space which can be denoted by

$$H_1 = \{ X / CX \geq K \}$$
$$H_2 = \{ X / CX \leq K \}$$



$CX = K$
$CX \geq K$
$CX \leq K$

H_1 is the Half space

ie:- that portion of R^n that contain the vectors X for which $\boxed{CX \geq K}$ and

H_2 is the Halfspace

ie:- that contain the vector X for which $\boxed{CX \leq K}$

Projection :-

Representing n- dimensional Object into (n-1) dimension is Known as Projection.

→ It is the Process of Converting a 3D Object into a 2D Object

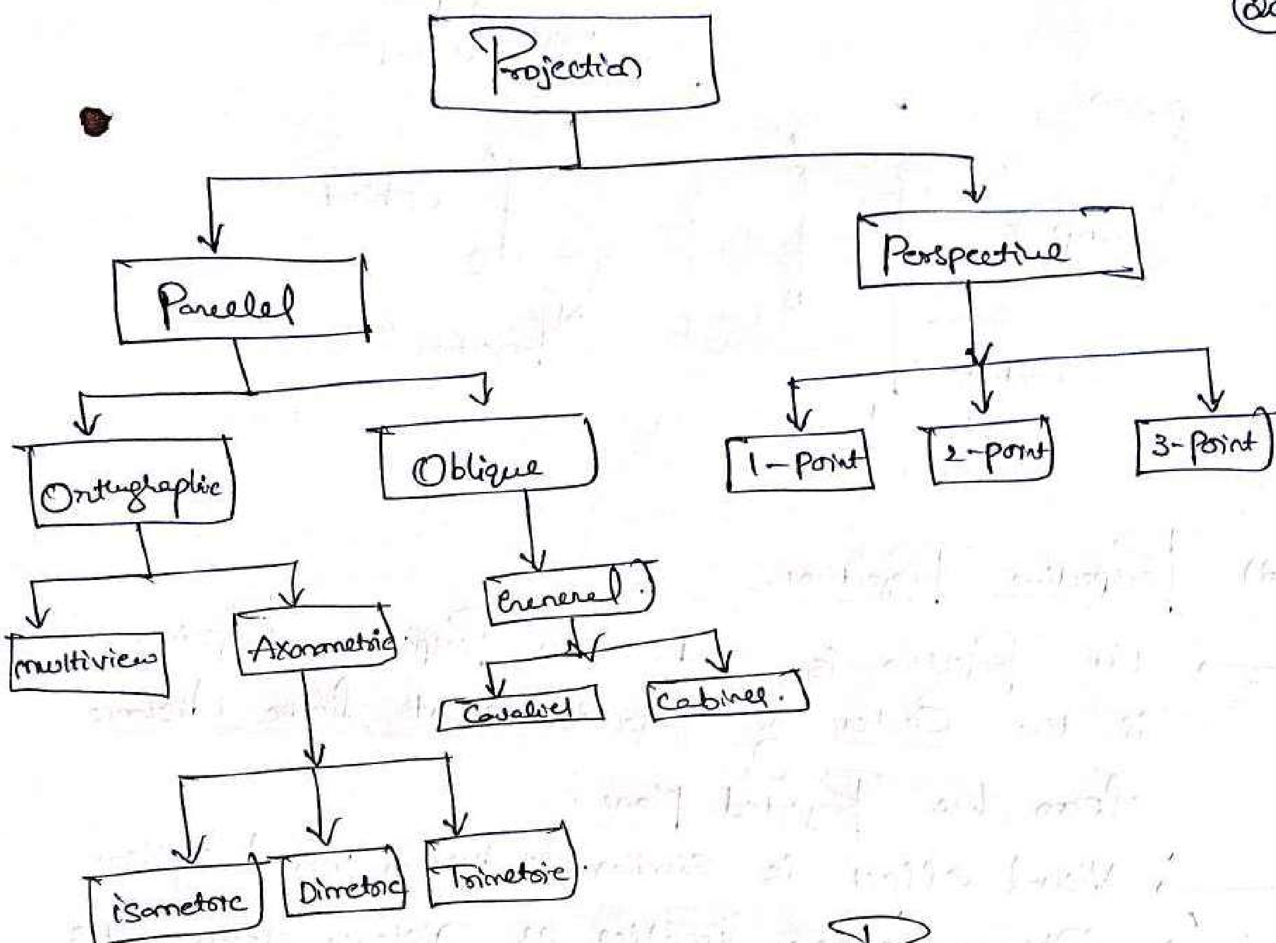
→ It is also defined as mapping (or) transformation of the Object in Projection plane (or) View plane.

(3D becomes 2D
$3D - 1 = 2D$.
$(n-1)$)

Projection are of two types

(1) Parallel Projection
(2) Perspective Projection

(20)



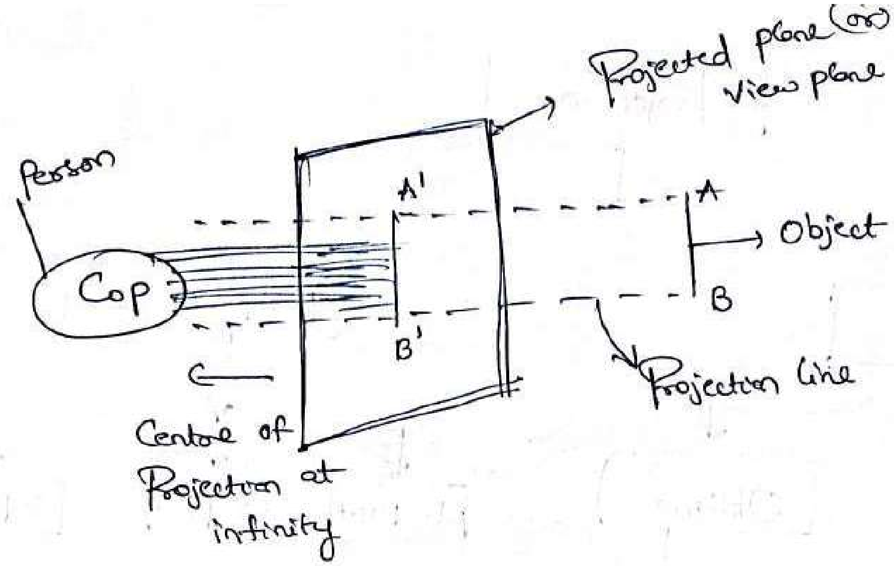
Overview of a Projection

(1) Parallel Projection :-

In this, Coordinate Positions are transformed to the View plane along Parallel lines.

→ A Projection is said to be Parallel if Centre of Projection is at infinite distance from the Projected plane.

→ The Projection lines are Parallel to each Other and extended from the Object and intersect the view plane.



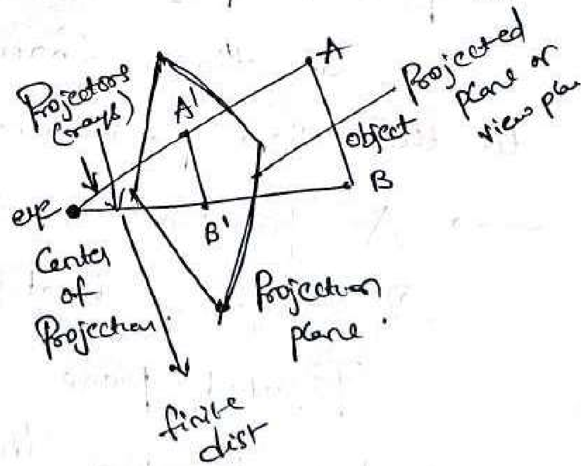
(2) Perspective Projection :-

→ The Projection is said to be Perspective Projection, if the Center of Projection is at finite distance from the Projected plane.

→ Visual effect is similar to human visual system

→ Objects appear smaller as distance from Center of Projection (Cop) (eye of Observer) increases.

→ Difficult to determine exact size and shape of the Object.



Dimension Reduction Techniques :-

The two popular and well-known dimension reduction techniques are

Dimension Reduction Techniques
- Principal Component Analysis (PCA)
- Fisher Linear Discriminant Analysis (LDA)

(1) Principal Component Analysis :- (PCA)

→ Principal Component Analysis is a well-known dimension reduction technique.

→ It transforms the variables into a new set of variables called as Principal Components

→ These Principal Components are Linear Combination of Original variables and are Orthogonal

→ The First Principal Component accounts for most of the possible variation of Original data.

→ The second Principal Component does its best to capture the variance in the data.

→ There can be only two Principal Components for a two-dimensional data set.

PCA Algorithm :-

The steps involved in PCA Algorithm are as follows

Step-1 :- Get data

Step-2 :- Compute the mean vector (μ)

Step-3 :- Subtract mean from the given data

Step-4 :- Calculate the Co-variance matrix.

Step-5 :- Calculate the eigen vectors & eigen values of the Co-variance matrix

Step-6 :- Choosing Components and forming a feature vector

Step-7 :- Deriving the new data set.

Problems

(1) Given data = { 2, 3, 4, 5, 6, 7 ; 1, 5, 3, 6, 7, 8 }
Compute the Principal Component using PCA Algorithm.

(or)

Consider the two dimensional patterns
(2, 1) (3, 5) (4, 3) (5, 6) (6, 7) (7, 8)
Compute the Principal Component using PCA Algorithm

(or)

Compute the Principal Componenent of following data

Class-1 :- $X = 2, 3, 4$
$Y = 1, 5, 3$

Class-2 :- $X = 5, 6, 7$
$Y = 6, 7, 8$

Sol :- we use the above discussed PCA Algorithm

Step-1 :- Get data

The given feature vectors are

$X_1 = (2, 1)$ $\quad X_4 = (5, 6)$

$X_2 = (3, 5)$ $\quad X_5 = (6, 7)$

$X_3 = (4, 3)$ $\quad X_6 = (7, 8)$

$$\begin{bmatrix}2\\1\end{bmatrix}\begin{bmatrix}3\\5\end{bmatrix}\begin{bmatrix}4\\3\end{bmatrix}\begin{bmatrix}5\\6\end{bmatrix}\begin{bmatrix}6\\7\end{bmatrix}\begin{bmatrix}7\\8\end{bmatrix}$$

Step-2 :- Compute the Mean vector (μ)

Calculate the Mean vector (μ)

Mean vector (μ) =

$$= \frac{(2+3+4+5+6+7)}{6}, \frac{(1+5+3+6+7+8)}{6}$$

$$= (4.5, 5)$$

Thus Mean vector (μ) = $\begin{bmatrix}4.5\\5\end{bmatrix}$

Step-3 :- Subtract mean vector (μ) from the given feature vectors

$$x_1 - \mu = (2-4.5, 1-5) = (-2.5, -4)$$

$$x_2 - \mu = (3-4.5, 5-5) = (-1.5, 0)$$

$$x_3 - \mu = (4-4.5, 3-5) = (-0.5, -2)$$

$$x_4 - \mu = (5-4.5, 6-5) = (0.5, 1)$$

$$x_5 - \mu = (6-4.5, 7-5) = (1.5, 2)$$

$$x_6 - \mu = (7-4.5, 8-5) = (2.5, 3)$$

Feature vectors (x_i) after subtracting mean vector (μ) are

$$\begin{bmatrix}-2.5\\-4\end{bmatrix}\begin{bmatrix}-1.5\\0\end{bmatrix}\begin{bmatrix}-0.5\\-2\end{bmatrix}\begin{bmatrix}0.5\\1\end{bmatrix}\begin{bmatrix}1.5\\2\end{bmatrix}\begin{bmatrix}2.5\\3\end{bmatrix}$$

23

Step-4 :-

Calculate the Covariance matrix

Covariance matrix is given by

$$\boxed{\text{Covariance Matrix} = \frac{\sum (x_i - \mu)(x_i - \mu)^t}{n}}$$

Now

$$m_1 = (x_1 - \mu)(x_1 - \mu)^t = \begin{bmatrix} -2.5 \\ -4 \end{bmatrix} \begin{bmatrix} -2.5 & -4 \end{bmatrix} = \begin{bmatrix} 6.25 & 10 \\ 10 & 16 \end{bmatrix}$$

$$m_2 = (x_2 - \mu)(x_2 - \mu)^t = \begin{bmatrix} -1.5 \\ 0 \end{bmatrix} \begin{bmatrix} -1.5, & 0 \end{bmatrix} = \begin{bmatrix} 2.25 & 0 \\ 0 & 0 \end{bmatrix}$$

$$m_3 = (x_3 - \mu)(x_3 - \mu)^t = \begin{bmatrix} -0.5 \\ -2 \end{bmatrix} \begin{bmatrix} -0.5 & -2 \end{bmatrix} = \begin{bmatrix} 0.25 & 1 \\ 1 & 4 \end{bmatrix}$$

$$m_4 = (x_4 - \mu)(x_4 - \mu)^t = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} \begin{bmatrix} 0.5 & 1 \end{bmatrix} = \begin{bmatrix} 0.25 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

$$m_5 = (x_5 - \mu)(x_5 - \mu)^t = \begin{bmatrix} 1.5 \\ 2 \end{bmatrix} \begin{bmatrix} 1.5 & 2 \end{bmatrix} = \begin{bmatrix} 2.25 & 3 \\ 3 & 4 \end{bmatrix}$$

$$m_6 = (x_6 - \mu)(x_6 - \mu)^t = \begin{bmatrix} 2.5 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 2.5 & 3 \end{bmatrix} = \begin{bmatrix} 6.25 & 7.5 \\ 7.5 & 9 \end{bmatrix}$$

Now Covariance Matrix

$$= \frac{(m_1 + m_2 + m_3 + m_4 + m_5 + m_6)}{6}$$

On adding the above matrices and dividing by 6, we get

$$\text{Covariance Matrix} = \frac{1}{6}\begin{bmatrix} 17.5 & 22 \\ 22 & 34 \end{bmatrix}$$

$$\text{Covariance Matrix} = \begin{bmatrix} 2.92 & 3.67 \\ 3.67 & 5.67 \end{bmatrix}$$

Step-5 :-

Calculate the eigen values and eigen vectors of the Covariance matrix.

λ is an eigen value for a matrix M If it is a solution of Characteristic equation

$$|M - \lambda I| = 0.$$

So, we have

$$\begin{vmatrix} 2.92 & 3.67 \\ 3.67 & 5.67 \end{vmatrix} - \begin{vmatrix} \lambda & 0 \\ 0 & \lambda \end{vmatrix} = 0$$

$$\begin{vmatrix} 2.92-\lambda & 3.67 \\ 3.67 & 5.67-\lambda \end{vmatrix} = 0$$

From here

$$(2.92-\lambda)(5.67-\lambda) - (3.67 \times 3.67) = 0$$

$$16.56 - 2.92\lambda - 5.67\lambda + \lambda^2 - 13.47 = 0$$

$$\lambda^2 - 8.59\lambda + 3.09 = 0$$

On solving this quadratic equation we get

$$\lambda = 8.22, 0.38$$

Thus two eigen values are

$\lambda_1 = 8.22$

$\lambda_2 = 0.38$

Clearly the second eigen value is very small Compared to the first eigen value

So, the second eigen vector can be left out

Eigen vector Corresponding to the greatest eigen value is the Principal Component for the given data set

So we find the eigen vector Corresponding to eigen value λ_1

we use the following equation to find the eigen vector $\boxed{MX = \lambda X}$

where M = Covariance Matrix

X = Eigen vector

λ = Eigen value.

On Substituting the values in the above equation we get

$$\begin{bmatrix} 2.92 & 3.67 \\ 3.67 & 5.67 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 8.22 \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

Solving these we get

$$2.92 X_1 + 3.67 X_2 = 8.22 X_1$$

$$3.67 X_1 + 5.67 X_2 = 8.22 X_2$$

on Simplification we get

$$5.3 X_1 = 3.67 X_2 \longrightarrow (1)$$

$$3.67 X_1 = 2.55 X_2 \longrightarrow (2)$$

From (1) & (2)

$x_1 = 0.69 x_2$

From (2) the eigen vector is

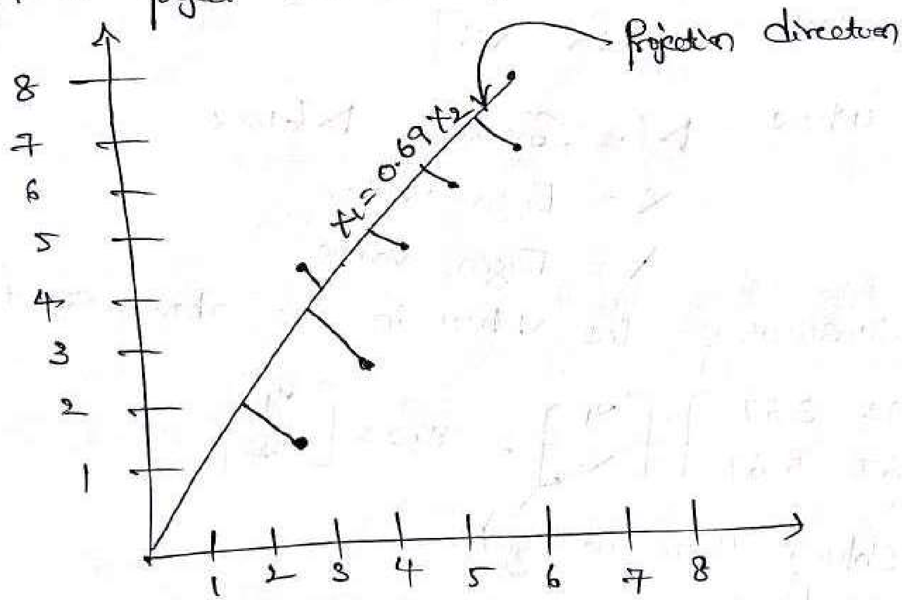
Eigen vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.55 \\ 3.67 \end{bmatrix}$

Thus, Principal Component for the given data set is

Principal Component

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.55 \\ 3.67 \end{bmatrix}$$

Lastly, we project the data points onto the new subspace as



(2) Use PCA Algorithm to transform the pattern (2,1) onto the eigen vector in the Previous question (25)

Sol: The given feature vector is (2,1)

Given Feature vector :- $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

The Feature vector gets transformed to =

Transpose of eigen vector X (Feature vector − Mean vector)

$$= \begin{bmatrix} 2.55 \\ 3.67 \end{bmatrix}^T \times \left(\begin{bmatrix} 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 4.5 \\ 5 \end{bmatrix} \right)$$

$$= \begin{bmatrix} 2.55 & 3.67 \end{bmatrix} \times \begin{bmatrix} -2.5 \\ -4 \end{bmatrix}$$

$$= \underline{\underline{-21.055}}$$

Problem on PCA

1) Given the following data, Use PCA to reduce the dimension from 2 to 1.

Feature	Example 1	Example 2	Example 3	Example 4
x	4	8	13	7
y	11	4	5	14

Sol:- Step-1 :- Data set :-

Feature	Example 1	Example 2	Example 3	Example 4
x	4	8	13	7
y	11	4	5	14

No of features, $n = 2$ (x, y)

No of Sample $N = 4$

Step-2 :- Computation of Mean of Variables

$$\bar{x} = \frac{4+8+13+7}{4} = 8$$

$$\bar{y} = \frac{11+4+5+14}{4} = 8.5$$

Step-3 :- Computation of Covariance Matrix

Ordered Pairs are (x, y) $n^2 = 2^2 = 4$

n variables

(x, x) (x, y) (y, x) (y, y)

26

1) Co-variance of all Ordered Pairs

$$Cov(x,x) = \frac{1}{N-1} \sum_{K=1}^{N} (x_{iK} - \bar{x}_i)(x_{jK} - \bar{x}_j)$$

$$= \frac{1}{4-1} \left[(4-8)^2 + (8-8)^2 + (13-8)^2 + (7-8)^2 \right]$$

$$= \underline{\underline{14}}.$$

$Cov(x,x) \longrightarrow$ If two variables are Same then

$$\boxed{Cov(x,x) = \frac{1}{N-1} \sum_{K=1}^{N} (x_i - \bar{x})^2}$$

$$Cov(x,y) = \frac{1}{4-1} \left[(4-8)(11-8.5) + (8-8)(4-8.5) + (13-8)(5-8.5) + (7-8)(14-8.5) \right]$$

$$= \underline{\underline{-11}}.$$

$$Cov(y,x) = Cov(x,y) = -11$$

$$Cov(y,y) = \frac{1}{4-1} \left[(11-8.5)^2 + (4-8.5)^2 + (5-8.5)^2 + (14-8.5)^2 \right]$$

$$= \underline{\underline{23}}$$

$\therefore$ we have

$$\boxed{\begin{aligned} Cov(x,x) &= 14 \\ Cov(x,y) &= -11 \\ Cov(y,x) &= -11 \\ Cov(y,y) &= 23. \end{aligned}}$$

Using all these 4 Co-varience values we are going to Construct Co-varience Method of size $n \times n$, 2×2.

$$S = \begin{bmatrix} Cov(x,x) & Cov(x,y) \\ Cov(y,x) & Cov(y,y) \end{bmatrix}$$ Co-varience matrix.

Covarience matrix values

Covarience Matrix $S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$ This is Co-varience Matrix S.

~~Fasch~~

Step 4 :- Eigen value, Eigen vector, Normalized Eigen vector.

1) Eigen value.

$$|S - \lambda I| = 0 \qquad I = 2\times 2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0 \qquad \lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = 0$$

$$\Rightarrow \left| \begin{bmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{bmatrix} \right| = 0.$$

27

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

$$\begin{vmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{vmatrix} = 0$$

$\Rightarrow (14-\lambda)(23-\lambda) - (-11 \times -11) = 0$

$\Rightarrow \lambda^2 - 37\lambda + 201 = 0.$

Now we find roots ∴ roots are.

$\lambda = \underline{30.3849,}\ 6.6151$

$\boxed{\lambda_1 > \lambda_2}$

$\lambda_1 = 30.3849$

$\lambda_2 = 6.6151$

Roots of Quadratic eqn

$\boxed{\frac{1}{2a} \sqrt{b^2 - 4ac}}$

$\frac{1}{2(1)} \sqrt{(-37)^2 - 4(1)(201)}$

$\Rightarrow \lambda$

Here

$a = 1$

$b = -37$

$c = 201$

The Largest eigen value by Principle.

∴ $\boxed{\lambda_1 = 30.3849}$ is Largest eigen value.

(ii) Eigen vector of λ_1.

$\boxed{(S - \lambda_1 I) U_1 = 0}$

~~Here~~ Where.

S = Co-Var matrix

I = Identity matrix

λ_1 = eigen value Largest

U_1 = eigen vector of λ_1

$$\begin{bmatrix} 14-\lambda_1 & -11 \\ -11 & 23-\lambda_1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0$$

$\boxed{U_1 = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_{2\times 1} \text{ matrix.}}$

$$\Rightarrow \begin{bmatrix} (14-\lambda_1)u_1 - 11u_2 \\ -11u_1 + (23-\lambda_1)u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\Rightarrow$

$$\begin{bmatrix} (14-\lambda_1)\mu_1 - 11\mu_2 \\ -11\mu_1 + (23-\lambda_1)\mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\Rightarrow (14-\lambda_1)\mu_1 - 11\mu_2 = 0$

~~$-11\mu_1 + (23-\lambda_1$~~

$-11\mu_1 + (23-\lambda_1)\mu_2 = 0$.

now we have to find value of μ_1, μ_2 from these Linear eqns.

$$\frac{\mu_1}{11} = \frac{\mu_2}{14-\lambda_1} = t \text{ (say)}$$

when $\boxed{t=1} \Rightarrow \mu_1 = 11$
$\mu_2 = 14-\lambda_1$

$\therefore$ Eigen vector U_1 of $\lambda_1 = \begin{bmatrix} 11 \\ 14-\lambda_1 \end{bmatrix}$

$$= \begin{bmatrix} 11 \\ 14-30.3849 \end{bmatrix} = \begin{bmatrix} 11 \\ -16.3849 \end{bmatrix}$$ eigenvector

(iii) Normalize the eigen vector U_1

$$e_1 = \begin{bmatrix} 11 \\ -16.3849 \end{bmatrix} = \begin{bmatrix} 11 / \sqrt{11^2 + (-16.3849)^2} \\ \dfrac{-16.3849}{\sqrt{11^2 + (-16.3849)^2}} \end{bmatrix}$$ (Length)

$$e_1 = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$$ Normalized eigen vector

28

For λ_2

$$e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$$

Step-5 :- Derive New data set.

1st Principal Component.

	Ex-1	Ex-2	Ex-3	Ex-4
First Principal Component PC1	P_{11} = ?	P_{12} = ?	P_{13} = ?	P_{14} = ?

$$P_{11} = e_1^T \begin{bmatrix} 4-8 \\ 11-8.5 \end{bmatrix} = \begin{bmatrix} 0.5574 & -0.8303 \end{bmatrix} \begin{bmatrix} -4 \\ 2.5 \end{bmatrix}$$

$$= \underline{-4.3052}. \quad 1\times1 \text{ matrix}$$

$$P_{12} = \begin{bmatrix} 0.5574 & -0.8303 \end{bmatrix} \begin{bmatrix} 8-8 \\ 4-8.5 \end{bmatrix} = \underline{3.7361}$$

Ily.

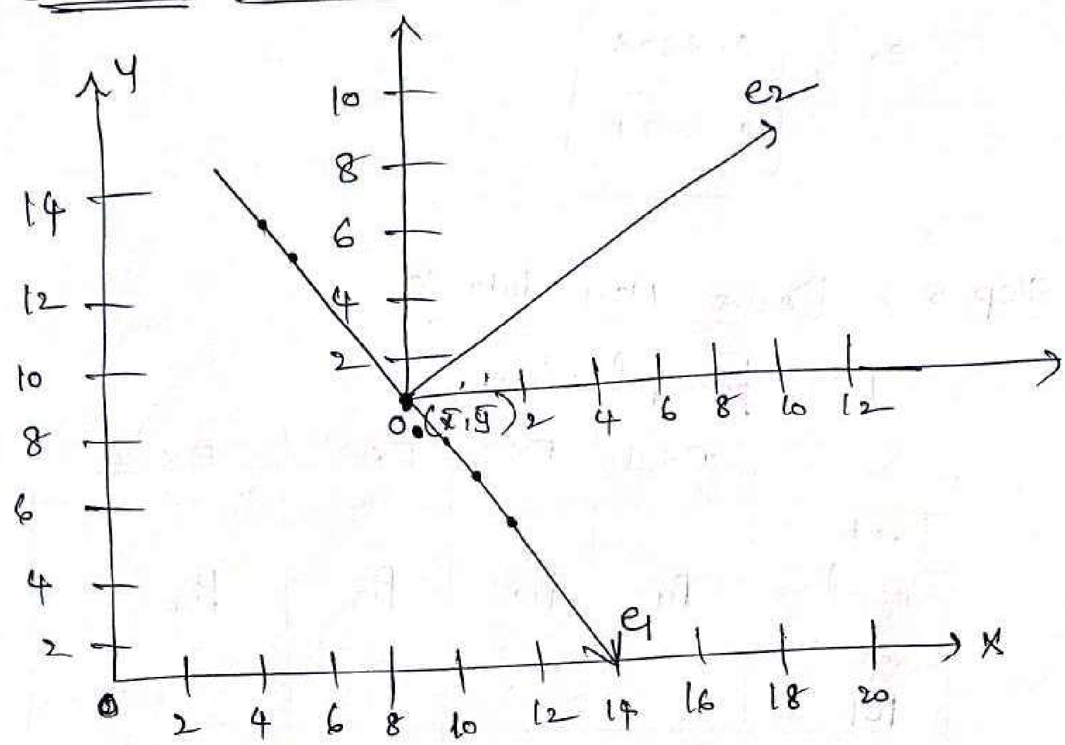
$P_{13} = 5.6928$

$P_{14} = -5.1238$

	Ex-1	Ex-2	Ex-3	Ex-4
PC1	P_{11} -4.3052	P_{12} 3.7361	P_{13} 5.6928	P_{14} -5.1238

New data set with reduced dimension 1

Co-ordinate System for Principal Components.



Mean Value of $x = 8$

" " " $y = 8.5$

$$e_1 = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$$

$$e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$$

Dimensionality Reduction

(29)

Dimension of an Instance ? (or) Length of Instance

⟶ Number of variables of instance.

Dimensionality reduction :-

Dimensionality reduction is the Process of reducing the number of Variables under Consideration by Obtaining a Smaller set of Principal Variables.

Advantage of reducing dimension :-

- ⟶ Decreases the Complexity of the algorithm
- ⟶ Saves the Cost of extracting an unnecessary input
- ⟶ Simple models can be choosed
- ⟶ Simplifying the Knowledge extraction
- ⟶ Easy to plot and analyse

Dimensionality Reduction

Dimensionality Reduction ⟶ Feature Selection, Feature extraction

Feature Selection

Find 'K' of d dimensions and discard (d-k) dimensions

⟶ Subset Selection.

Feature extraction

1) Find a new K dimensions that are Combinations of 'd' dimensions

2) PCA (Principal Component Analysis)

3) LDA (Linear Discriminant Analysis)

Subset Selection :-

→ Also known as Variable Selection, attribute selection, feature selection

→ {A, B, C}

{A} {B} {C} {AB} {BC} {AC} {ABC} {ϕ}

→ Advantages :-

→ Simplification of Models
→ Shorter training times
→ Enhanced generalization
→ To avoid the Curse of dimensionality.

Subset Selection

Forward Selection	Backward Selection.
→ Start with no Variables and add them one by one. at each step adding the One that decrease the error the most until any further addition does not decrease the error	Start with all Variables and remove them one by one at each step till the error become minimum

PCA (Principal Component Analysis) 30

(used in Machine Learning)

1) Find the PCA

x	2.5	0.5	2.2	1.9	3.1	2.3	2.0	1.0	1.5	1.1
y	2.4	0.7	2.9	2.2	3.0	2.7	1.6	1.1	1.6	0.9

$\boxed{n = 10}$ x & y are 2 Variables

$\bar{x} = \frac{\Sigma x}{n_1} = \underline{1.81}$ $\bar{y} = \frac{\Sigma y}{n_2} = \underline{1.91}$

$\therefore$ Mean $\Rightarrow \boxed{\begin{matrix} \bar{x} = 1.81 \\ \bar{y} = 1.91 \end{matrix}}$

Co-variance Matrix $= \begin{bmatrix} Cov(x,x) & Cov(x,y) \\ Cov(x,y) & Cov(y,y) \end{bmatrix}$

$$Cov(x,x) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$$

(A) $x_i - \bar{x}$	(B) $y_i - \bar{y}$	AB	A^2	B^2
0.69	0.49	0.3381		
−1.31	−1.21	1.5851		
0.39	0.99	0.3861		
0.09	0.29	0.0261		
1.29	1.09	1.4061		
0.49	0.79	0.3871		
0.19	−0.31	−0.0589		
−0.81	−0.81	0.6561		
−0.31	−0.31	0.0961		
−0.71	−1.01	0.7171		
		$\boxed{5.539}$		

A $x_i - \bar{x}$	B $y_i - \bar{y}$	AB	A^2	B^2
0.69	0.49	0.3381	0.4761	0.2401
−1.31	−1.21	1.5851	1.7161	1.4641
0.39	0.99	0.3861	0.1521	0.9801
0.09	0.29	0.0261	0.0081	0.0841
1.29	1.09	1.4061	1.6641	1.1881
0.49	0.79	0.3871	0.2401	0.6241
0.19	−0.31	−0.0589	0.0361	0.0961
−0.81	−0.81	0.6561	0.6561	0.6561
−0.31	−0.31	0.0961	0.0961	0.0961
−0.71	−1.01	0.7171	0.5041	1.0201
		5.539	5.549	6.449

Co-variance Matrix (A) $= \dfrac{\begin{bmatrix} 5.549 & 5.539 \\ 5.539 & 6.449 \end{bmatrix}}{n-1}$

$10 - 1 = 9$

$= \begin{bmatrix} 0.6166 & 0.6154 \\ 0.6154 & 0.7166 \end{bmatrix}$

31

Eigen values $|A - \lambda I| = 0$.

$$\Rightarrow (0.6166-\lambda)(0.7166-\lambda) - (0.6154)^2 = 0$$

$$\Rightarrow 0.4418 - 0.6166\lambda - 0.7166\lambda + \lambda^2 - 0.3787 = 0$$

$$\lambda^2 - 1.3332\lambda + 0.0631 = 0$$

$$ax^2 + bx + c.$$

$$\boxed{\begin{aligned}\lambda_1 &= 1.2841 \\ \lambda_2 &= 0.0491\end{aligned}}$$ are eigen value

$$\boxed{\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}}$$

$$\begin{bmatrix} 0.6166-\lambda & 0.6154 \\ 0.6154 & 0.7166-\lambda \end{bmatrix}$$

Now we substitute $\lambda_1\ \lambda_2$ to find eigen vectors.

(In place of $\lambda \rightarrow \lambda_1$ value
$\lambda \rightarrow \lambda_2$ value.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} -0.6634 & 0.6154 \\ 0.6154 & -0.5631 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

(λ_1 substituted)

$$\left.\begin{aligned} -0.6634\,a_{11} + 0.6154\,a_{12} &= 0 \\ 0.6154\,a_{11} - 0.5631\,a_{12} &= 0 \end{aligned}\right\} \text{Linear eqns}$$

$Z_1 = a_{11}x_1 + a_{12}x_2$
$Z_2 = a_{21}x_1 + a_{22}x_2$

$$\left.\begin{aligned} 0.6634\,a_{11} &= 0.6154\,a_{12} \\ 0.6154\,a_{11} &= 0.5631\,a_{12} \end{aligned}\right\} \text{add eqns both}$$

$$a_{11} = 1.1785\,a_{12}$$

$$a_{11} = \frac{1.5785}{1.2788}\,a_{12}$$

$$a_{11} = \underline{0.9215\,a_{12}}$$

$$\begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} \Rightarrow \begin{bmatrix} 0.9215 \\ 1 \end{bmatrix} \text{(eigen vector)} \Rightarrow \sqrt{(0.9215)^2 + 1^2}$$

$$= \sqrt{0.8491 + 1}$$

$$\Rightarrow \begin{bmatrix} \frac{0.9215}{1.3598} \\ \frac{1}{1.3598} \end{bmatrix} \qquad = \sqrt{1.8491}$$

$$= \underline{\underline{1.3598}}$$

$$\text{Eigen vectors} = \begin{bmatrix} 0.6677 \\ 0.7354 \end{bmatrix} \qquad \sqrt{(a_{11})^2 + (a_{12})^2} = \underline{\underline{1.3598}}$$

$$\begin{bmatrix} 0.5675 & 0.6154 \\ 0.6154 & 0.6675 \end{bmatrix} \begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

(λ_2 substitute)

$$0.5675\, a_{21} = -0.6154\, a_{22}$$

$$0.6154\, a_{21} = -0.6675\, a_{22} \quad \text{(add)}$$

$$1.1829\, a_{21} = -1.2829\, a_{22}$$

$$a_{21} = \frac{-1.2829}{1.1829} a_{22} = \underline{\underline{-1.0845\, a_{22}}}$$

$$\begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix} = \begin{bmatrix} -1.0845 \\ 1 \end{bmatrix} \Rightarrow \sqrt{(1.0845)^2 + 1^2}$$

$$= \sqrt{1.1761 + 1}$$

$$= \sqrt{2.1761} = \underline{\underline{1.4751}}$$

32

$$\begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{-1.0845}{1.4751} \\ \frac{1}{1.4751} \end{bmatrix}$$

$$\text{Eigen vectors} = \begin{bmatrix} -0.7352 \\ 0.6779 \end{bmatrix}$$

Total variance % = Eigen vector 1 Eigen vector 2

$$\left(\frac{\lambda_1}{\lambda_1+\lambda_2} \quad \frac{\lambda_2}{\lambda_1+\lambda_2} \right) \Rightarrow \frac{1.2841}{1.2841+0.0491} \qquad \frac{0.0491}{1.2841+0.0491}$$

$$= \frac{1.2841}{1.3332} \qquad = \frac{0.0491}{1.3332}$$

$$= 0.96 \qquad = 0.036$$

(96%) ✓ (= 3.6%) ✗

↓ High value ↓ Low value.

↓ Eigen vector 1 ↓ Eigen vect 2

(reduce the data value available)

33

Principal Component Analysis (PCA)

1) Given two attributes X & Y with values given in the table below

X	2.5	0.5	2.2	1.9	3.1	2.3	2	1	1.5	1.1
Y	2.4	0.7	2.9	2.2	3.0	2.7	1.6	1.1	1.6	0.9

Find the eigen vector & Principal Component from the given data.

Sol:- PCA Algorithm

(1) Get data
(2) Subtract the mean (subtract mean from data)
(3) Calculate the Covariance matrix
(4) Calculate the eigen vector & eigen values of the Covariance matrix
(5) Choosing Components & forming a feature vector
(6) Deriving the new data set, this is final step in PCA.

Sol :- (1) Get data (given data)

(2) Mean :-

$$\bar{x} = \frac{2.5 + 0.5 + 2.2 + 1.9 + 3.1 + 2.3 + 2 + 1 + 1.5 + 1.1}{10}$$

$$\therefore \boxed{\bar{x} = 1.81}$$

$$\bar{y} = \frac{2.4 + 0.7 + 2.9 + 2.2 + 3.0 + 2.7 + 1.6 + 1.1 + 1.6 + 0.9}{10}$$

$$\boxed{\bar{y} = 1.91}$$

(3) Covariance Matrix :- $\boxed{n = 10}$ no of terms.

$$C = \begin{bmatrix} Cov(x,x) & Cov(x,y) \\ Cov(y,x) & Cov(y,y) \end{bmatrix}$$

$$\boxed{Cov(x,y) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})(y_i - \bar{y})}{n-1}}$$

$$\boxed{Cov(x,x) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})(x_i - \bar{x})}{n-1}}$$

34

$$\boxed{Cov(x,x) = \sum_{i=1}^{n} \frac{(x_i-\bar{x})(x_i-\bar{x})}{n-1}}$$

x	$(x_i-\bar{x})$	$(x_i-\bar{x})(x_i-\bar{x})$
2.5	0.69	0.4761
0.5	−1.31	1.7161
2.2	0.39	0.1521
1.9	0.09	0.0081
3.1	1.29	1.6641
2.3	0.49	0.2401
2	0.19	0.0361
1	−0.81	0.6561
1.5	−0.31	0.0961
1.1	−0.71	0.5041
		$\sum(x_i-\bar{x})(x_i-\bar{x}) = 5.549 = 8$um

$\boxed{\bar{x} = 1.81}$

$$Cov(x\,x) = \frac{5.549}{n-1} \qquad \boxed{n=10}$$

$$= \frac{5.549}{9}$$

$$= 0.6165$$

$$\boxed{Cov(x,x) = 0.6165}$$

||ly $Cov(x,y)$

$Cov(y,x)$

$Cov(y,y)$

$$C = \begin{bmatrix} Cov(x,x) & Cov(x,y) \\ Cov(y,x) & Cov(y,y) \end{bmatrix}$$

$$C = \begin{bmatrix} 0.6165 & 0.6154 \\ 0.6154 & 0.7165 \end{bmatrix}$$ //

(4) Eigen value & Eigen vector :- (Co-variance matrix)

$$\boxed{C - \lambda I = 0}$$

$$\begin{bmatrix} 0.6165 & 0.6154 \\ 0.6154 & 0.7165 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} 0.6165 - \lambda & 0.6154 \\ 0.6154 & 0.7165 - \lambda \end{bmatrix} = 0 \qquad |C - \lambda I| = 0$$

$$\Rightarrow (0.6165)\left[0.7165 - \lambda\right] - \lambda\left[0.7165 - \lambda\right] - 0.6154 \times 0.6154 = 0$$

$$\Rightarrow 0.6165 \times 0.7165 - 0.6165\lambda - 0.7165\lambda + \lambda^2 - (0.6154)^2 = 0$$

$$\Rightarrow \lambda^2 - 1.333\lambda + 0.4417 - 0.3737 = 0$$

$$\Rightarrow \lambda^2 - 1.333\lambda + 0.063 = 0.$$

$$\therefore \left.\begin{array}{l} \boxed{\lambda_1 = 0.0490} \\ \boxed{\lambda_2 = 1.2840} \end{array}\right\} \text{are Eigen values}$$

Eigen vector :- For $\boxed{\lambda_1 = 0.0490}$

$$\begin{bmatrix} 0.6165 - 0.0490 & 0.6154 \\ 0.6154 & 0.7165 - 0.0490 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = 0$$

$$\Rightarrow \begin{bmatrix} 0.5674 & 0.6154 \\ 0.6154 & 0.6674 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

35

$$\begin{bmatrix} 0.5674 & 0.6154 \\ 0.6154 & 0.6674 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\implies 0.5674x_1 + 0.6154y_1 = 0 \longrightarrow (1)$$

$$0.6154x_1 + 0.6674y_1 = 0 \longrightarrow (2)$$

$$\implies \boxed{x^2 + y^2 = 1}$$ on solving (1) & (2)

$$y_1 = -\frac{0.5674}{0.6154} x_1$$

$$x_1^2 + \left(\frac{-0.5674}{0.6154} x_1\right)^2 = 1$$

$$\boxed{x_1 = 0.7351}$$

$$\boxed{y_1 = -0.6778}$$

For $\boxed{\lambda_2 = 1.2840}$

$$\begin{bmatrix} -0.6675 & 0.6154 \\ 0.6154 & -0.5675 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$-0.6675x_2 + 0.6154y_2 = 0$$

$$0.6154x_2 - 0.5675y_2 = 0.$$

$$y_2 = \frac{0.6675}{0.6154} x_2$$

now using $\boxed{x^2 + y^2 = 1}$ $\boxed{x_2 = 0.6773}$

~~$x_2 = 0.66$~~ $\boxed{y_2 = 0.7351}$

$$\begin{array}{cc} x_1 & y_1 \\ \end{array}$$
$$\begin{bmatrix} 0.7351 & -0.6778 \\ 0.6778 & 0.7351 \end{bmatrix}$$
$$\begin{array}{cc} x_2 & y_2 \end{array}$$

(5)

$\lambda_1 = 0.0490$

$\lambda_2 = 1.2840$ $\quad \lambda_2 > \lambda_1$ Clearly.

$$\begin{array}{c} \lambda_1 \\ \lambda_2 \end{array} \begin{bmatrix} 0.7357 & -0.6778 \\ \boxed{0.6778 \quad 0.7357} \end{bmatrix}$$

eigen vector

PCA

PCA :-

⟶ It is a way of identifying patterns in data and expressing the data in such a way to highlight their similarities & differences

⟹ Dimensionality Reduction.

- **Principal Component Analysis (PCA)**

is a statistical procedure that is used to reduce the dimensionality.

It uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. It is often used as a dimensionality reduction technique.

Steps involved in the PCA :-

Step-1 :- Standardise the data set.

Step-2 :- Calculate the covariance matrix for the features in the dataset.

Step-3 :- Calculate the eigen values and eigen vectors for the covariance matrix.

Step-4 :- Sort eigen values and their corresponding eigen vectors.

Step-5 :- pick k eigen values and form a matrix of eigen vectors

Step-6 :- Transform the original matrix.

Advantages of Dimensionality Reduction :-

(1) Dimensionality reduction helps in data Compression, and hence reduced the storage space.

(2). It reduces Computation time.

(3) It also helps remove redundant features If any.

(4) Dimensionality Reduction helps in data Compressing and reducing the Storage space required.

(5) It fastens the time required for Performing same Computations.

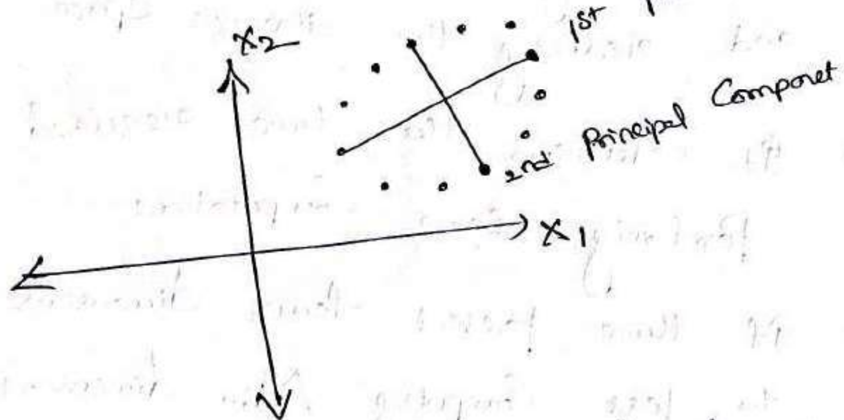
(6) If there present fewer dimensions then it leads to less Computing. Also dimensions can allow usage of algorithms unfit for a large number of dimensions.

(7) It takes Care of multiCollinearity that improves the model Performance. It removes redundant features.

for example, there is no point in storing a value in two different units (meters & inches)

(8) Reducing the dimensions of data to 2D or 3D may allow us to plot and visualize it Precisely. You can then Observe Patterns more clearly.

Below you can see that, how a 3D data is Converted into 2D, First it has identified the 2D plane then represented the points on these two new axes Z_1 and Z_2.



It is helpful in noise removal also and as a result of that, we can improve the Performance of models.

38

Dis-advantages of Dimensionality Reduction :-

(1) Basically, it may lead to some amount of data loss.

(2) Although, PCA tends to find linear correlations between variables, which is sometimes undesirable

(3) Also, PCA fails in cases where mean and covariance are not enough to define datasets.

(4) Further, we may not know how many principal components to keep - in practice some thumb rules are applied.

Importance of Dimensionality Reduction :-

1) why is Dimension Reduction is important in machine learning Predictive modeling ?

A) The Problem of unwanted increase in dimension is closely related to other. That was to fixation of measuring / recording data at a far granular level then it was done in past.

This is no way suggesting that this is a ~~recen~~ recent Problem..

It has started gaining more importance lately due to a surge in data.

Dis-advantages of Dimensionality Reduction :-

1) It may lead to some amount of data loss
2) PCA tends to find linear Correlations between Variables, which is sometimes undesirable.
3) PCA fails in Cases where mean and Covariance are not enough to define datasets.

Advantages of Dimensionality Reduction :-

1) It helps in data Compression and hence reduced Storage Space.
2) It reduces Computation time.
3) It also helps remove redundant features, If any

Machine Learning: Machine Learning is nothing but a field of study which allows Computers to "Learn" like humans without any need of explicit Programming.

What is Predictive Modeling :-

Predictive modeling is a Probabilistic Process that allows us to forecast outcomes, on the basis of some Predictors.

These Predictors are basically features that come into play when deciding the final result of the Model.

ie: the outcome of the Model.

What is Dimensionality Reduction ?

In machine learning classification Problems, there are often too many factors on the basis of which the final classification is done.

These factors are basically Variables called features.

The higher the number of features, the harder it gets to visualize the training set and then work on it.

Sometimes most of these features are Correlated and hence redundant. That is where dimensionality reduction algorithms come into play.

(40)

Dimensionality reduction is the Process of reducing the number of random variables under consideration, by obtaining a set of Principal Variables.

It can be divided into feature selection and feature extraction.

Components of Dimensionality Reduction :-

There are two Components of dimensionality reduction

1) Feature Selection :- In this, we try to find a subset of the Original set of Variables, or features, to get a smaller subset which can be used to model the Problem

It usually involves Three ways.

1) Filter
2) wrapper
3) Embedded.

2) Feature Extraction :- This reduces the data in a high dimensional space to a lower dimension space

ie:- a space with lesser no of dimensions.

Methods of Dimensionality Reduction :-

The various methods used for dimensionality reduction include

(1) Principal Component Analysis (PCA)

2) Linear Discriminant Analysis (LDA)

3) Generalized Discriminant Analysis (GDA)

Dimensionality reduction may be both linear or non-Linear, depending upon the method used.

(41)

- → **Principal Component Analysis (or) PCA** is a dimensionality reduction method that is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set.

→ Based on the dataset find a new set of orthogonal feature vectors in such a way that the data spread is maximum in the direction of the feature vector (or) dimension.

Covariance Formula :-

Covariance formula is a statistical formula which is used to assess the relationship between two variables.

In simple words, Covariance is one of the statistical measurement to know the relationship of the variance between the two variables.

The Covariance indicates how two variables are related and also helps to know whether the two variables vary together or change together.

The Covariance is denoted by Cov(X,Y) and the formula of Covariance are given below

Population Covariance formula:

$$\boxed{Cov(x, y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N}}$$

Sample Covariance formula :-

$$\boxed{Cov(x, y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N-1}}$$

These are the formulas to find sample and population Covariance.

Notations in Covariance formulas:

X_i = data value of X

Y_i = data value of Y

$\bar{X}$ = mean of X

$\bar{Y}$ = mean of Y

N = Number of data values.

43

Co-variance

$$\boxed{Cov(x,y) = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{N}}$$

1)

x_i	10	12	14	8
y_i	40	48	56	32

$\bar{x} = \frac{10+12+14+8}{4} = \frac{44}{4} = 11$ Mean of $x_i = 11$

$\therefore \bar{x} = 11$

$\bar{y} = \frac{40+48+56+32}{4} = \frac{176}{4} = 44$ Mean of $y_i = 44$

x_i	$x_i - \bar{x}$ ($\bar{x}=11$)	y_i	$y_i - \bar{y}$ ($\bar{y}=44$)
10	−1	40	−4
12	1	48	4
14	3	56	12
8	−3	32	−12

$Cov(xy) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{N}$

$= \frac{(-1)(-4) + (1)(4) + (3)(12) + (-3)(-12)}{4}$

$= \frac{(4) + (4) + (36) + (36)}{4}$

$= \frac{80}{4} = 20$

$\therefore Cov(xy) = 20$

Co-Variance :-

If X & Y are two random variables then Covariance between them is defined as

$$Cov(XY) = \sigma_{xy} = E(XY) - E(X)E(Y)$$

Proof :- $Cov(XY) = E[(X - E(X))(Y - E(Y))]$

$$= E[XY - XE(Y) - YE(X) + E(X)E(Y)]$$

$$= E(XY) - E(X)E(Y) - \cancel{E(Y)E(X)} + \cancel{E(X)E(Y)}$$

$$= \underline{E(XY) - E(X)E(Y)}$$

$$\left(\begin{array}{l} E(K) = K \\ E[E(X)] = E(X) \end{array} \right]$$

Properties of Covariance :-

(1) If X & Y are Independent, then $E(XY) = E(X)E(Y)$ and hence

$$Cov(XY) = E(X)E(Y) - E(X)E(Y) = 0$$

2) $Cov(X,X) = Var(X)$

3) $Cov(aX, bY) = ab\, Cov(XY)$

4) $Cov(X+a, Y+b) = Cov(X, Y)$

(44)

- (5) $\mathrm{Cov}\left(\frac{x-\bar{x}}{\sigma_x}, \frac{y-\bar{y}}{\sigma_y}\right)$

$$= \frac{1}{\sigma_x \sigma_y} \mathrm{Cov}(x, y)$$

6) $\mathrm{Cov}(x+y, z) = \mathrm{Cov}(x, z) + \mathrm{Cov}(y, z)$

Covariance of (x, y)

In Bi-variate distribution of (x_i, y_i) take values

$(x_1, y_1), (x_2, y_2) \ldots (x_n, y_n)$

$$\mathrm{Cov}(x, y) = \frac{1}{n} \sum (x-\bar{x})(y-\bar{y})$$

$$= \frac{1}{n} \sum \left[xy - x\bar{y} - \bar{x}y + \bar{x}\bar{y} \right]$$

$$= \frac{1}{n} \sum xy - \bar{y} \frac{\sum x}{n} - \bar{x} \frac{\sum y}{n} + \bar{x}\bar{y}$$

$$= \frac{1}{n} \sum xy - 2\bar{x}\bar{y} + \bar{x}\bar{y}$$

$$= \frac{1}{n} \sum xy - \bar{x}\bar{y}$$

$$\therefore \boxed{\mathrm{Cov}(x, y) = \frac{1}{n} \sum xy - \bar{x}\bar{y}}$$

45

(2)

x	1	2	3	4	5
y	2	3	4	6	10

$$\boxed{Cov(xy) = \frac{1}{n}\sum xy - \bar{x}\bar{y}}$$

$$\bar{x} = \frac{\sum x}{n} = \frac{1+2+3+4+5}{5} = \frac{15}{5} = 3$$

$$\therefore \boxed{\bar{x} = 3}$$

$$\bar{y} = \frac{\sum y}{n} = \frac{2+3+4+6+10}{5} = \frac{25}{5} = 5$$

$$\therefore \boxed{\bar{y} = 5}$$

x	y	xy
1	2	2
2	3	6
3	4	12
4	6	24
5	10	50
		$\sum xy = 94$

$$Cov(xy) = \frac{1}{n}\sum xy - \bar{x}\bar{y}$$

$$= \frac{1}{5} \times (94) - (3)(5)$$

$$= 18.8 - 15$$

$$= 3.8$$

Problem of Covariance formula :-

1) The table below describes the rate of economic growth (x_i) and the rate of return on the S&P 500 (y_i)

Using the Covariance formula, determine whether economic growth & S&P 500 returns have a positive or inverse relationship

Before you compute the Covariance, Calculate the Mean of X & Y.

Economic growth % (x_i)	2.1	2.5	4.0	3.6
S&P 500 Returns % (y_i)	8	12	14	10

Sol :- given x_i = 2.1, 2.5, 4.0 & 3.6 (economic growth)

y_i = 8, 12, 14, 10 (S&P 500 returns)

Find $\overline{x}$ & $\overline{y}$

$$\overline{x} = \frac{\Sigma x_i}{n} = \frac{2.1+2.5+4.0+3.6}{4} = \frac{12.2}{4} = 3.1$$

$$\therefore \boxed{\overline{x} = 3.1}$$

$$\overline{y} = \frac{\Sigma y_i}{n} = \frac{8+12+14+10}{4} = \frac{44}{4} = 11$$

$$\therefore \boxed{\overline{y} = 11}$$

(46)

Now substitute these values into the Covariance formula to determine the relationship between economic growth & S&P 500 returns.

$\overline{X} = 3.1$ $\overline{Y} = 11$

x_i	y_i	$x_i - \overline{x}$	$y_i - \overline{y}$
2.1	8	−1	−3
2.5	12	−0.6	1
4.0	14	0.9	3
3.6	10	0.5	−1

$$\therefore \; Cov(x,y) = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{N}$$

$$= \frac{(-1)(-3) + (-0.6)(1) + (0.9)(3) + (0.5)(-1)}{4}$$

$$= \frac{4.6}{4} = \underline{1.15}.$$

47

1) Co-Variance

What is Covariance in relation to Variance & Correlation

Two Data sets   5 elements data set

$X = (2, 4, 6, 8, 10)$

$Y = (1, 3, 8, 11, 12)$

Variance $= s^2 =$ A measure of how spread out the numbers of a data set are

X Average $(\bar{x}) = \frac{\sum x_i}{n} = \frac{2+4+6+8+10}{5} = \frac{30}{5} = 6$

Y Average $(\bar{y}) = \frac{\sum y_i}{n} = \frac{1+3+8+11+12}{5} = \frac{35}{5} = 7$

(X) Variance $(s_x^2) = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{(2-6)^2 + (4-6)^2 + \cdots + (10-6)^2}{5}$

$= \frac{16+4+0+4+16}{5} = \frac{40}{5} = 8$.

(Y) Variance $(s_y^2) = \frac{\sum (y_i - \bar{y})^2}{n} = \frac{(1-7)^2 + (3-7)^2 + \cdots + (12-7)^2}{5}$

$= \frac{36+16+1+16+25}{5} = \frac{94}{5} = 18.8$

Co-Variance : Cov(x y) = A measure of how the trends of 2 data sets are related

$Cov(x\,y) = \frac{\sum (x-\bar{x})(y-\bar{y})}{n} = \frac{(-4)(-6) + (-2)(-4) + (0)(0) + (2)(4) + (4)(5)}{5}$

$= \frac{24+8+0+8+20}{5} = \frac{60}{5} = 12$.

Correlation :- (r) = A measure of how the trends of 2 data sets are related $-1 \leq r \leq 1$

$$r = \frac{Cov(x,y)}{S_x S_y} = \frac{12}{\sqrt{8} \cdot \sqrt{18.8}} = \underline{\underline{0.98}}.$$

$$r = \frac{Cov(x,y)}{\sqrt{s_x^2} \cdot \sqrt{s_y^2}} = \underline{\underline{0.98}}.$$ (strong relationship between 2 sets)

2) How to Calculate the Variance :-

Data set $Z = (-4, -1, 5, 12, 18)$ (The whole population)

Step-1 : Find the Mean (Average)

$$\bar{Z} = \frac{\sum_i Z_i}{n} = \frac{(-4)+(-1)+(5)+(12)+(18)}{5}$$

$$\bar{Z} = \frac{30}{5} = \underline{\underline{6}}$$

Step-2 :- Find the Variance :

$$S_Z^2 = \frac{\sum_i (z_i - \bar{z})^2}{n} = \frac{(-4-6)^2 + (-1-6)^2 + (5-6)^2 + (12-6)^2 + (18-6)^2}{5}$$

$$\therefore \underline{\underline{S_Z^2 = 66}}.$$

$$= \frac{(-10)^2 + (-7)^2 + (-1)^2 + (6)^2 + (12)^2}{5}$$

$$= \frac{100+49+1+36+144}{5} = \frac{330}{5} = \underline{\underline{66}}.$$

(48)

3) Population vs Sample Variance :-

$$X = (2, 4, 6, 8, 10)$$

$$\bar{X} = \frac{\sum_i x_i}{n} = \frac{2+4+6+8+10}{5} = \frac{30}{5} = \underline{\underline{6}}.$$

$$\therefore \underline{\underline{\bar{X} = 6}}.$$

Sample Variance :-

$$S^2 = \frac{\sum_i (x_i - \bar{x})^2}{n-1} = \frac{(2-6)^2 + (4-6)^2 + (6-6)^2 + (8-6)^2 + (10-6)^2}{5-1}$$

$$= \frac{16+4+0+4+16}{4} = \frac{40}{4} = \underline{\underline{10}}.$$

$$\therefore \underline{\underline{S^2 = 10}}.$$

$$\sigma^2 = \frac{\sum_i (x_i - \bar{x})^2}{n} = \frac{(2-6)^2 + (4-6)^2 + (6-6)^2 + (8-6)^2 + (10-6)^2}{5}$$

$$= \frac{16+4+0+4+16}{5} = \frac{40}{5} = \underline{\underline{8}}.$$

4) **How to Calculate the Co-Variance :-**

we have 2 data sets.

$X = (2, 4, 6, 8, 10)$
$Y = (12, 11, 8, 3, 1)$

Step-1 :- Find the Mean (average) of both sets.

$$\bar{x} = \frac{\sum_i x_i}{n} = \frac{2+4+6+8+10}{5} = \frac{30}{5} = \underline{\underline{6}}$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{12+11+8+3+1}{5} = \frac{35}{5} = \underline{\underline{7}}$$

Step-2 :- Find the Variance of both sets.

$$\sigma_x^2 = \frac{\sum_i (x_i - \bar{x})^2}{n} = \frac{(2-6)^2 + (4-6)^2 + (6-6)^2 + (8-6)^2 + (10-6)^2}{5}$$

$$= \frac{4^2 + 2^2 + 0^2 + 2^2 + 4^2}{5}$$

$$= \frac{16+4+0+4+16}{5} = \frac{40}{5} = \underline{\underline{8}}.$$

$$\sigma_y^2 = \frac{\sum_i (y_i - \bar{y})^2}{n} = \frac{(12-7)^2 + (11-7)^2 + (8-7)^2 + (3-7)^2 + (1-7)^2}{5}$$

$$= \frac{5^2 + 4^2 + 1^2 + 4^2 + 6^2}{5}$$

$$= \frac{25+16+1+16+36}{5} = \frac{94}{5} = \underline{\underline{18.8}}.$$

Step-3 :- Find the Covariance :-

$$Cov(xy) = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{n}$$

$$= \frac{(-4)(5) + (-2)(4) + (0)(1) + (2)(-4) + (4)(-6)}{5}$$

$$= \frac{-20 - 8 + 0 - 8 - 24}{5} = \frac{-60}{5} = \underline{\underline{-12}}.$$

(5) **Covariance :- What is the Covariance Matrix**

The Covariance Matrix is an $n \times n$ matrix (where n = no of data sets) such that the diagonal elements represent the variances of each data set and the off-diagonal elements represent the Covariance between the data sets.

1) $X = 2, 4, 6, 8, 10$ $\quad \bar{X} = 6 \quad Var(x) = \sigma_x^2 = 8$

$Y = 1, 3, 8, 11, 12$ $\quad \bar{Y} = 7 \quad Var(y) = \sigma_y^2 = 18.8$

$Cov(xy) = 12 = Cov(y, x)$

$$Var(x) = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N}$$

$$Cov(xy) = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{N}$$

Covariance Matrix =

$$= \begin{array}{cc} & \begin{array}{cc} x & y \end{array} \\ \begin{array}{c} x \\ y \end{array} & \begin{bmatrix} Var(x) & Cov(xy) \\ Cov(yx) & Var(y) \end{bmatrix} \end{array}$$

$$= \begin{bmatrix} 8 & 12 \\ 12 & 18.8 \end{bmatrix} //.$$

2) Example of Co-Variance Matrix :

$X = 2, 4, 6, 8, 10.$

$Y = 7, 3, 5, 1, 9.$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{2+4+6+8+10}{5} = \frac{30}{5} = \underline{\underline{6}}. \quad \therefore \boxed{\bar{x} = 6}$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{7+3+5+1+9}{5} = \frac{25}{5} = \underline{\underline{5}}. \quad \therefore \boxed{\bar{y} = 5}$$

$$\sigma_x^2 = Var(x) = \frac{\sum_i (x_i - \bar{x})^2}{n} = \frac{(2-6)^2 + (4-6)^2 + (6-6)^2 + (8-6)^2 + (10-6)^2}{5}$$

$$\sigma_x^2 = \underline{\underline{8}}.$$

$$\sigma_y^2 = Var(y) = \frac{\sum_i (y_i - \bar{y})^2}{n} = \frac{(7-5)^2 + (3-5)^2 + (5-5)^2 + (1-5)^2 + (9-5)^2}{5}$$

$$\sigma_y^2 = \underline{\underline{8}}.$$

$$Cov(x\,y) = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{n}$$

$$= \frac{(2-6)(7-5) + (4-6)(3-5) + (6-6)(5-5) + (8-6)(1-5) + (10-6)(9-5)}{5}$$

$$= \frac{(-4)(2) + (-2)(-2) + (0)(0) + (2)(-4) + (4)(4)}{5}$$

$$= \frac{4}{5} = \underline{\underline{0.8}}.$$

$$\therefore Cov(x\,y) = \underline{\underline{0.8}}.$$

50

$$\text{Co-variance Matrix} = \begin{bmatrix} Var(x) & Cov(x\,y) \\ Cov(y\,x) & Var(y) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_x & Cov(x\,y) \\ Cov(y\,x) & \sigma_y \end{bmatrix}$$

$$= \begin{bmatrix} 8 & 0.8 \\ 0.8 & 8 \end{bmatrix}$$

3) Example of Covariance Matrix :-

$x = 2, 4, 6, 8, 10$

$y = 10, 8, 6, 4, 2$

$$\bar{x} = \frac{\sum_i x_i}{n} = \frac{2+4+6+8+10}{5} = \frac{30}{5} = 6 \Rightarrow \boxed{\bar{x} = 6}$$

$$\bar{y} = \frac{\sum_i y_i}{n} = \frac{10+8+6+4+2}{5} = \frac{30}{5} = 6 \Rightarrow \boxed{\bar{y} = 6}$$

$$\sigma_x^2 \; (Var(x)) = \frac{\sum_i (x_i - \bar{x})^2}{n} = \frac{(2-6)^2 + (4-6)^2 + (6-6)^2 + (8-6)^2 + (10-6)^2}{5}$$

$$= 8$$

$$\sigma_y^2 \; Var(y) = \frac{\sum_i (y_i - \bar{y})^2}{n} = \frac{(10-6)^2 + (8-6)^2 + (6-6)^2 + (4-6)^2 + (2-6)^2}{5}$$

$$= 8$$

$$Cov(xy) = \frac{\sum_{i} (x_i - \bar{x})(y_i - \bar{y})}{n}$$

$$= \frac{(2-6)(10-6) + (4-6)(8-6) + (6-6)(6-6) + (8-6)(4-6) + (10-6)(2-6)}{5}$$

$$= \frac{(-4)(4) + (-2)(2) + (0)(0) + (2)(-2) + (4)(-4)}{5}$$

$$= \frac{-40}{5} = \underline{\underline{-8}}.$$

$$\text{Covariance Matrix} = \begin{bmatrix} Var(x) & Cov(xy) \\ Cov(yx) & Var(y) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_x & Cov(xy) \\ Cov(yx) & \sigma_y \end{bmatrix}$$

$$= \begin{bmatrix} 8 & -8 \\ -8 & 8 \end{bmatrix}$$

51

8) Covariance :-

What is Correlation Coefficient :-

$x = 2, 4, 6, 8, 10 \qquad \bar{x} = 6$

$y = 1, 3, 8, 11, 12 \qquad \bar{y} = 7.$

$$\text{Var}(x) = \sigma_x^2 = \frac{\sum_i (x_i - \bar{x})^2}{n} = \underline{8}.$$

$$\text{Var}(y) = \sigma_y^2 = \frac{\sum_i (y_i - \bar{y})^2}{n} = \frac{94}{5} = \underline{\underline{18.8}}$$

$$\text{Cov}(xy) = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{n}$$

$$= \frac{(2-6)(1-7) + (4-6)(3-7) + (6-6)(8-7) + (8-6)(11-7) + (10-6)(12-7)}{5}$$

$$= \frac{60}{5} = \underline{\underline{12}}.$$

$$\text{Cov Matrix} = \begin{bmatrix} 8 & 12 \\ 12 & 18.8 \end{bmatrix}$$

$$\text{Correlation Co-eff } r = \frac{\text{Cov}(xy)}{\sqrt{\text{Var}(x)}\sqrt{\text{Var}(y)}} = \frac{\text{Cov}(xy)}{\sqrt{\sigma_x^2} \cdot \sqrt{\sigma_y^2}}$$

$$= \frac{12}{\sqrt{8} \cdot \sqrt{18.8}} = \underline{\underline{0.98}} \qquad \underline{\underline{(-1 \leq r \leq 1)}}$$

9) Calculate the Correlation Co-efficient.

$x = 2, 4, 6, 8, 10$

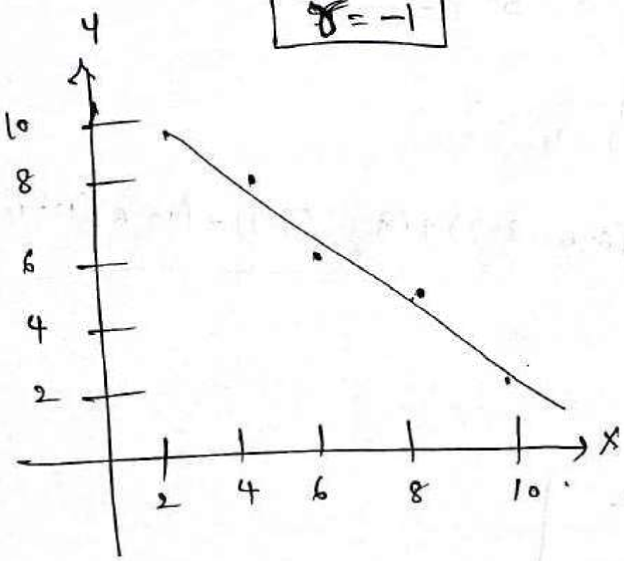
$y = 10, 8, 6, 4, 2$

$\bar{x} = 6 \quad \bar{y} = 6 \quad \sigma_x = 8 \quad \sigma_y = 8$

$Cov(x y) = -8$

$$\gamma = \frac{Cov(x, y)}{\sqrt{\sigma_x^2} \cdot \sqrt{\sigma_y^2}} = \frac{-8}{\sqrt{8} \cdot \sqrt{8}}$$

$$\boxed{\gamma = -1}$$



There is strong Correlation (line)

$x = 2, 4, 6, 8, 10$

$y = 7, 3, 5, 1, 9$

$\bar{x} = 6, \ \bar{y} = 5, \ \sigma_x = 8 \ \ \sigma_y = 8$

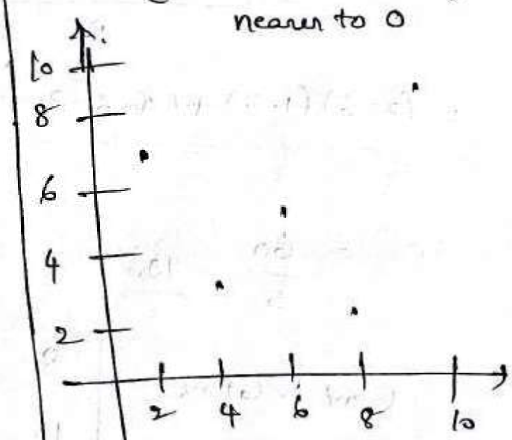
$Cov(x y) = 0.8$

$$\gamma = \frac{Cov(x y)}{\sqrt{\sigma_x^2} \cdot \sqrt{\sigma_y^2}} = \frac{0.8}{\sqrt{8} \cdot \sqrt{8}}$$

$$= \frac{0.8}{8}$$

$$\boxed{\gamma = 0.1}$$

(little Correlation)
nearer to 0



There is no strong Correlation Co-eff (no line)

52

(10) **Covariance :- Covariance Matrix with 3 data sets.**

$X = 2, 4, 6, 8, 10 \qquad \sum X = 30 \quad \bar{X} = 6$

$Y = 3, 6, 9, 12, 15 \qquad \sum y = 45 \quad \bar{y} = 9$

$Z = 9, 7, 5, 3, 1 \qquad \sum y = 25 \quad \bar{Z} = 5$

$$Var(X) = \frac{\sum_1 (x_i - \bar{x})^2}{n} = \frac{(2-6)^2 + (4-6)^2 + (6-6)^2 + (8-6)^2 + (10-6)^2}{5}$$

$$= \frac{40}{5} = \underline{\underline{8}}$$

$$Var(y) = \frac{\sum_1 (y_i - \bar{y})^2}{n} = \frac{(3-9)^2 + (6-9)^2 + (9-9)^2 + (12-9)^2 + (15-9)^2}{5}$$

$$= \frac{90}{5} = \underline{\underline{18}}$$

$$Var(Z) = \frac{\sum_1 (z_i - \bar{Z})^2}{n} = \frac{(9-5)^2 + (7-5)^2 + (5-5)^2 + (3-5)^2 + (1-5)^2}{5}$$

$$= \frac{40}{5} = \underline{\underline{8}}.$$

$$\text{Cov Matrix} = \begin{array}{c c} & \begin{array}{ccc} x & y & z \end{array} \\ \begin{array}{c} x \\ y \\ z \end{array} & \begin{bmatrix} Var(x) & Cov(xy) & Cov(xz) \\ Cov(yx) & Var(y) & Cov(y,z) \\ Cov(zx) & Cov(zy) & Var(Z) \end{bmatrix} \end{array}$$

$$= \begin{bmatrix} 8 & & \\ & 18 & \\ & & 8 \end{bmatrix}$$

$$Cov(xy) = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{n}$$

$$= \frac{(2-6)(3-9) + (4-6)(6-9) + (6-6)(9-9) + (8-6)(12-9) + (10-6)(15-9)}{5}$$

$$= \frac{(-4)(-6) + (-2)(-3) + (0)(0) + (2)(3) + (4)(6)}{5}$$

$$= \frac{24 + 6 + 0 + 6 + 24}{5} = \frac{60}{5} = \underline{\underline{12}}.$$

$$\therefore \boxed{Cov(xy) = 12.}$$

$$Cov(xz) = \frac{\sum_i (x_i - \bar{x})(z_i - \bar{z})}{n} =$$

$$= \frac{(2-6)(9-5) + (4-6)(7-5) + (6-6)(5-5) + (8-6)(3-5) + (10-6)(1-5)}{5}$$

$$= \frac{-40}{5} = \underline{\underline{-8}}. \quad \therefore \boxed{Cov(xz) = -8}$$

$$Cov(y,z) = \frac{\sum_i (y_i - \bar{y})(z_i - \bar{z})}{n} = \frac{-60}{5} = \underline{\underline{-12}}.$$

$$\therefore \boxed{Cov(y,z) = -12}$$

53

$$\therefore \text{Cov Matrix} = \begin{bmatrix} 8 & 12 & -8 \\ 12 & 18 & -12 \\ -8 & -12 & 8 \end{bmatrix}$$

$$r_{xy} = \frac{Cov(x,y)}{\sqrt{var(x)}\sqrt{var(y)}} = \frac{12}{\sqrt{8}\cdot\sqrt{18}} = \underline{\underline{1}}$$

$$r_{xz} = \frac{Cov(x,z)}{\sqrt{var(x)}\cdot\sqrt{var(z)}} = \frac{-8}{\sqrt{8}\ \sqrt{8}} = \underline{\underline{-1}}$$

$$r_{yz} = \frac{Cov(y,z)}{\sqrt{var(y)}\cdot\sqrt{var(z)}} = \frac{-12}{\sqrt{18}\cdot\sqrt{8}} = \underline{\underline{-1}}$$

13) **Covariance Matrix Using Sample data :-**

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \\ 3 & 1 & 1 \\ 4 & 1 & 2 \end{bmatrix} \text{ Sample data}$$

$$\bar{x} = 2.5 \qquad \bar{y} = 2 \qquad \bar{z} = \frac{7}{4} = 1.75$$

$$Var(x) = S_x^2 = \frac{\sum_i (x_i - \bar{x})^2}{n-1} = \frac{(1-2.5)^2 + (2-2.5)^2 + (3-2.5)^2 + (4-2.5)^2}{4-1}$$

$$Var(x) = \underline{\underline{1.67}}.$$

$$Var(y) = S_y = \frac{\sum_i (y_i-\bar{y})^2}{n-1} = \frac{(2-2)^2+(4-2)^2+(1-2)^2+(1-2)^2}{4-1}$$

$$= \frac{6}{3} = \underline{\underline{2}}.$$

$$\therefore \boxed{Var(y) = 2}$$

$$Var(Z) = S_Z = \frac{\sum_i (z_i-\bar{z})^2}{n-1} = \frac{(3-1.75)^2+(1-1.75)^2+(1-1.75)^2+(2-1.75)^2}{4-1}$$

$$= \frac{2.75}{3} = \underline{\underline{0.9167}}$$

$$\boxed{Var(Z) = 0.9167}$$

$$Cov(xy) = \frac{\sum_i (x_i-\bar{x})(y_i-\bar{y})}{n-1} = \frac{(-1.5)(0)+(-0.5)(2)+(0.5)(-1)+(1.5)(-1)}{4-1}$$

$$= \frac{-3}{3} = \underline{\underline{-1}}.$$

$$\therefore \boxed{Cov(xy) = -1}$$

$$Cov(x, Z) = \frac{\sum (x_i-\bar{x})(z_i-\bar{z})}{n-1} = \frac{(-1.5)(1.25)+(-0.5)(-0.75)+(0.5)(-0.75)+(1.5)(0.25)}{4-1}$$

$$= \frac{-1.5}{3} = \underline{\underline{-0.5}}.$$

$$\therefore \boxed{Cov(xZ) = -0.5}$$

54

$$Cov(y,z) = \frac{\sum_i (y_i - \bar{y})(z_i - \bar{z})}{n-1}$$

$$= \frac{(0)(1.25) + (2)(-0.75) + (-1)(-0.75) + (-1)(0.25)}{4-1}$$

$$= -\frac{1}{3} = -0.333$$

$$\therefore \boxed{Cov(y,z) = -0.333}$$

$$\text{Covariance Matrix} = \begin{bmatrix} 1.67 & -1 & -0.5 \\ -1 & 2 & -0.333 \\ -0.5 & -0.333 & 0.9167 \end{bmatrix}$$

(55)

Singular Value Decomposition (SVD)

The Singular Value Decomposition (SVD) of a matrix is a factorization of that matrix into three matrices.

It has some interesting algebraic Properties and ~~Conveys~~ Conveys important geometrical and theoretical insights about linear transformations

It also has some important applications in data science.

SVD :-

The SVD of $m \times n$ matrix A is given by the formula

$$\boxed{A = UWV^T}$$

where U = $m \times n$ matrix of the Orthonormal eigen vectors of AA^T

V^T = Transpose of a $n \times n$ matrix Containing the Orthonormal eigen vectors of A^TA

W = a $n \times n$ diagonal matrix of the singular values which are the square roots of eigen values of A^TA

SVD :-

$$\boxed{C_{m\times n}} = \boxed{U_{m\times r}} \times \boxed{\Sigma_{r\times r}} \times \boxed{V'_{r\times n}}$$

Problem :-

1) Find the SVD for the matrix

$$A = \begin{bmatrix} 3 & 3 & 2 \\ 2 & 3 & -2 \end{bmatrix}_{2\times 3}$$

Sol :- To Calculate the SVD First we need to Compute the Singular Values by finding eigen values of AA^T

$$AA^T = \begin{bmatrix} 3 & 3 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 3 & 3 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

The Characteristic equation for the above matrix is

$$W - \lambda I = 0$$

$$AA^T - \lambda I = 0$$

$$\lambda^2 - 34\lambda + 225 = 0$$

$$\Longrightarrow (\lambda - 25)(\lambda - 9)$$

So our Singular values are $\sigma_1 = 5$ $\sigma_2 = 3$

(56)

Now we find the right singular vectors

ie :- Orthonormal set of eigen vectors of A^TA

The eigen values of A^TA are 25, 9 & 0

and since A^TA is symmetric we know that the eigen vectors will be Orthogonal.

For $\boxed{\lambda = 25}$

$$AA^T - 25I = \begin{bmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{bmatrix}$$

which can be row reduces to

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

A unit vector in the direction of it is

$$v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

|||ly for $\boxed{\lambda = 9}$ the eigen vector is

$$v_2 = \begin{bmatrix} 1/\sqrt{18} \\ -1/\sqrt{18} \\ 4/\sqrt{18} \end{bmatrix}$$

For the 3rd eigen vector, we could use the property that it is $\perp r$ to V_1 & V_2 such that

$$V_1^T V_3 = 0$$

$$V_2^T V_3 = 0$$

Solving the above equation to generate the 3rd eigen vector

$$V_3 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ -a \\ -a/2 \end{bmatrix} = \begin{bmatrix} 2/3 \\ -2/3 \\ -1/3 \end{bmatrix}$$

Hence our Final SVD equation becomes

$$A = A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ \frac{2}{3} & \frac{-2}{3} & \frac{1}{3} \end{bmatrix}$$

57

(2) Find the SVD of a 2×3 matrix A having values

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Sol :-

Step-1 :- Find A^T & then Compute A^TA

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \text{ then } A^T = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Step-2 :- Find eigen values associated with matrix A^TA

Eigen values associated with A^TA

$$\lambda = 0, 1 \text{ \& } 3$$

Step-3 :- Find the singular values corresponding to the obtained eigen values using formula $\sigma_i = \sqrt{\lambda_i}$

$$\boxed{\sigma_i = \sqrt{\lambda_i}}$$

Singular values associated with A^TA

$\lambda = 3, 1$ & 0

$\lambda_1 = 3 \implies \sigma_1 = \sqrt{3}$

$\lambda_2 = 1 \implies \sigma_2 = 1$

$\lambda_3 = 0 \implies \sigma_3 = 0$

Step-4 :- Compute diagonal matrix Σ using the values of σ keeping the above discussed cases in mind.

As $(m = 2 < n = 3)$

Case-1 is applied and matrix Σ is

$$\Sigma = \begin{bmatrix} \sqrt{3}, & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Step-5 :- Find the eigen vectors & corresponding normalized eigen vectors associated with the matrix A^TA

Eigen vectors associated with A^TA

For $\lambda_1 = 3 \implies X_1 = [1, 2, 1]$

For $\lambda_2 = 1 \implies X_2 = [-1, 0, 1]$

For $\lambda_3 = 0 \implies X_3 = [1, -1, 1]$

where X_1 X_2 X_3 are eigen vectors of matrix A^TA.

(58)

Normalized eigen vectors associated with A^TA

For $X_1 = [1, 2, 1]$

$$\Rightarrow V_1 = [(1/\sqrt{6}), (2/\sqrt{6}), (1/\sqrt{6})]$$

For $X_2 = [-1, 0, 1]$

$$\Rightarrow V_2 = [(-1/\sqrt{2}), 0, (1/\sqrt{2})]$$

For $X_3 = [1, -1, 1]$

$$\Rightarrow V_3 = [(1/\sqrt{3}), (-1/\sqrt{3}), (1/\sqrt{3})]$$

where V_1, V_2, & V_3 are eigen vectors of matrix A^TA

Step-6:- Use eigen vectors Obtained to Compute matrix V.

$$V = \begin{bmatrix} (1/6) & (-1/2) & (1/3) \\ (2/6) & 0 & (-1/3) \\ (1/6) & (1/2) & (1/3) \end{bmatrix}$$

Step-7:- Use the above given equation to Compute the Orthogonal matrix U.

$$u_1 = \frac{Av_1}{\sigma_1}$$

$$= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} (1/6) \\ (2/6) \\ (1/6) \end{bmatrix} = \begin{bmatrix} (1/2) \\ (1/2) \end{bmatrix}$$

$$u_2 = \frac{Av_2}{\sigma_2}$$

$$= \frac{1}{1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} (-1/2) \\ 0 \\ (1/2) \end{bmatrix} = \begin{bmatrix} (-1/2) \\ (1/2) \end{bmatrix}$$

$\therefore$ Orthogonal matrix U is

$$U = \begin{bmatrix} (1/2) & (-1/2) \\ (1/2) & (1/2) \end{bmatrix}$$

<u>Step-8</u>:- Compute the SVD of A using the equation given below

$$A = U \Sigma V^T$$

$\therefore$ Using SVD, A can be expressed as

$$A = U \Sigma V^T$$

$$= \begin{bmatrix} (1/2) & (-1/2) \\ (1/2) & (1/2) \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} (1/6) & (-1/2) & (1/3) \\ (2/6) & 0 & (-1/3) \\ (1/6) & (1/2) & 1/3 \end{bmatrix}$$

59

Singular Value Decomposition:

(SVD)

For Rectangular matrix

$$\boxed{A = U \Sigma V^T}$$

$A \rightarrow$ Given input matrix : $m \times n$
$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad r \quad c$

(m – rows
  n – Columns)

For V : $n \times n$ Columns, are EV of $A^T A$

For U : $m \times m$ Columns are EV of $A A^T$

Σ : diagonal matrix : $m \times n$

Problem

1) Find Singular value decomposition of matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}$$

(1) Find Singular value decomposition of matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}_{\substack{3\times 2 \\ m\times n}}$$

Sol: To find $\boxed{A = U \Sigma V^T}$

Step-1 :- Compute $A^T A$ for V : EV : $n\times n = 2\times 2$

AA^T for U : EV : $m\times m = 3\times 3$

$$A^T A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1+0+1 & 1+0-1 \\ 1+0-1 & 1+1+1 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}_{\substack{2\times 2 \\ n\times n}} \quad (\text{for } V)$$

$$AA^T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1+1 & 0+1 & -1+1 \\ 0+1 & 0+1 & 0+1 \\ -1+1 & 0+1 & 1+1 \end{bmatrix}$$

$$AA^T = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}_{\substack{3\times 3 \\ m\times m}} \quad (\text{for } U)$$

(60)

Step-2 :- Find Eigen values for $A^TA \Rightarrow V$

so we can calculate V^T

Characteristic equation $|A - \lambda I| = 0$

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$$

$$\begin{vmatrix} 2-\lambda & 0 \\ 0 & 3-\lambda \end{vmatrix} = 0$$ Here $A^TA = A$

$$(2-\lambda)(3-\lambda) = 0$$

$$6 - 2\lambda - 3\lambda + \lambda^2 = 0$$

$$\boxed{\lambda^2 - 5\lambda + 6 = 0}$$

$\therefore$ $\boxed{\lambda = 3, 2}$ are eigen values

write in decending Order.

Singular Values are Obtained from eigen values.

Eigen vectors for $\boxed{\lambda = 3}$:-

$$(A - \lambda I)X = 0$$

$$\begin{bmatrix} 2-\lambda & 0 \\ 0 & 3-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Put $\boxed{\lambda = 3}$

$$\begin{bmatrix} 2-\lambda & 0 \\ 0 & 3-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 2-3 & 0 \\ 0 & 3-3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Let $x_2 = x_2 \quad x_1 = 0$

$\therefore$ Eigen vector $X_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$V_1 = \frac{X_1}{\|X_1\|} = \frac{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}{1} \qquad \|x_1\| = \sqrt{0^2+1^2} = \sqrt{1^2} = \underline{\underline{1}}$$

$$\|x_1\| = 1$$

$$V_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Eigen vector for $\boxed{\lambda = 2}$ (61)

$$(A - \lambda I)\, X = 0$$

$$\begin{bmatrix} 2-\lambda & 0 \\ 0 & 3-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Put $\boxed{\lambda = 2}$

$$\Rightarrow \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

F P

Let $x_1 = x_1$ $x_2 = 0$

$x_1/x_1 = 1$

Eigen vector $X_2 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$$V_2 = \frac{X_2}{\|X_2\|} = \frac{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}{1} \qquad \|X_2\| = \sqrt{1^2 + 0^2} = \sqrt{1^2} = 1.$$

$$V_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Note:- $\|x_1\| = \sqrt{0^2 + 1^2} = 1$

$\|x_2\| = \sqrt{1^2 + 0^2} = 1$

Eigen vectors are

$$v_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Step-3 : Find matrix U : $m \times m = 3 \times 3 = AA^T$

we have

$$U = \begin{bmatrix} \frac{Av_1}{\sigma_1}, & \frac{Av_2}{\sigma_2} \end{bmatrix}$$

$$\quad\quad (u_1) \quad\quad (u_2)$$

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{3}} \begin{bmatrix} 0+1 \\ 0+1 \\ 0+1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

$$\therefore U_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

62

$$U_2 = \frac{AV_2}{\sigma_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \frac{1}{\sqrt{2}}\begin{bmatrix} 1+0 \\ 0+0 \\ -1+0 \end{bmatrix}$$

$$U_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$$

$U_1^T U_3 = 0$ $\qquad$ $U_2^T U_3 = 0$

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} \qquad \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}$$

$$= \frac{\alpha}{\sqrt{3}} + \frac{\beta}{\sqrt{3}} + \frac{\gamma}{\sqrt{3}} = 0 \qquad \frac{\alpha}{\sqrt{2}} - \frac{\gamma}{\sqrt{2}} = 0$$

$$\frac{\alpha+\beta+\gamma}{\sqrt{3}} = 0 \qquad \frac{\alpha-\gamma}{\sqrt{2}} = 0$$

$$\beta + 2\alpha = 0 \qquad \alpha - \gamma = 0$$

$$\boxed{\beta = -2\alpha} \qquad \boxed{\alpha = \gamma}$$

$$\therefore \text{Vector :-} \begin{bmatrix} \alpha \\ -2\alpha \\ \alpha \end{bmatrix} = \alpha\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

$$U_3 = \frac{U_3}{\|U_3\|} = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

$$\|U_3\| = \sqrt{1^2+4+1} = \sqrt{6}$$

$$\|U_3\| = \sqrt{6}$$

$$\therefore U = \{U_1, U_2, U_3\}$$

$$U = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}_{3\times 3}$$

Step-4 :-

$$\therefore \Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}_{\substack{3\times 2 \\ m\ \ n}}$$

$\therefore$ SVD is method of decomposing a rectangular matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}_{3\times 2}$ into 3 matrices

63

First matrix $U = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}_{\substack{m \times m \\ 3 \times 3}}$

Second matrix $\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}_{\substack{3 \times 2 \\ m \times n}}$

Third matrix $V^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}_{2 \times 2}$.

A is decomposed into 3 matrix.

First matrix
Second matrix
Third matrix

Verification:
$$\begin{bmatrix} 1+0+0 & 0+1+0 \\ 1+0+0 & 0+0+0 \\ 1+0+0 & 0-1+0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\boxed{U \Sigma V^T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} = A.}$$

Singular value Decomposition (SVD)

Let A be a $m \times n$ matrix

Then the SVD divides this matrix into 2 Unitary matrices that are Orthogonal in nature and a rectangular diagonal matrix Containing Singular values till 'r'

Mathematically it is expressed as

$$\boxed{A = U \Sigma V^T}$$

where $\Sigma \longrightarrow (m \times n)$ Orthogonal matrix

$U \longrightarrow (m \times m)$ Orthogonal matrix

$V \longrightarrow (n \times n)$ diagonal matrix with first r rows having Only Singular.

64

Gram Schmidt Process

(1) Apply the Gram Schmidt Process to the vectors

$\beta_1 = (1,0,1)$, $\beta_2 = (1,0,-1)$ $\beta_3 = (0,3,4)$

Sol:- Given three vectors are

$\beta_1 = (1,0,1)$, $\beta_2 = (1,0,-1)$ $\beta_3 = (0,3,4)$

let $\{\alpha_1, \alpha_2, \alpha_3\}$ be Orthogonal basis

step (1) Let $\alpha_1 = \beta_1 = (1,0,1)$

$\|\alpha_1\| = \sqrt{1^2+0^2+1^2} = \sqrt{2}$ $\therefore \|\alpha_1\| = \sqrt{2}$.

step (2)

$$\boxed{\alpha_2 = \beta_2 - \frac{\langle \beta_2 \cdot \alpha_1 \rangle}{\|\alpha_1\|^2} \cdot \alpha_1}$$

$\alpha_2 = (1,0,-1) - \frac{0}{\sqrt{2}} \cdot (1,0,1)$

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

$\boxed{\alpha_2 = (1,0,-1)}$

$\langle \beta_2 \cdot \alpha_1 \rangle =$

$\beta_2 = (1,0,-1)$

$\alpha_1 = (1,0,1)$

$\langle \beta_2 \cdot \alpha_1 \rangle = (1+0,-1)$

$= \underline{0}$.

Step-3

$$\boxed{\alpha_3 = \beta_3 - \frac{\langle \beta_3 \cdot \alpha_1 \rangle}{\|\alpha_1\|^2} \cdot \alpha_1 - \frac{\langle \beta_3 \cdot \alpha_2 \rangle}{\|\alpha_2\|^2} \cdot \alpha_2}$$

$$\alpha_3 = (0,3,4) - \frac{\cancel{4}^{2}}{\cancel{2}_{1}}(1,0,1) - \frac{-\cancel{4}^{2}}{\cancel{2}}(1,0,-1)$$

$$= (0,3,4) - 2(1,0,1) + 2(1,0,-1)$$

$$= (0,3,4) + (-2,0,-2) + (2,0,-2)$$

$$= (0,3,4) + (-2,0,-2) + (2,0,-2)$$

$$= \underline{(0,3,0)}$$

$\langle \beta_3 \cdot \alpha_1 \rangle$	$\langle \beta_3 \cdot \alpha_2 \rangle$	$\|\alpha_2\|^2$
$\beta_3 = (0,3,4)$ $\alpha_1 = (1,0,1)$ $\langle \beta_3 \cdot \alpha_1 \rangle =$ $(0+0+4)$ $= \underline{4}$.	$\beta_3 = (0,3,4)$ $\alpha_2 = (1,0,-1)$ $\langle \beta_3 \cdot \alpha_2 \rangle =$ $(0+0-4)$ $= -4$	$= \sqrt{1^2+0^2+(-1)^2}$ $= 1+0+1 = \sqrt{2}$. $\|\alpha_2\| = \sqrt{2}$

$\|\alpha_3\|^2$

$= \sqrt{0^2+3^2+0^2}$

$= \sqrt{3^2}$

$= \sqrt{9} = \underline{\underline{3}}$

$\underline{\|\alpha_3\| = 3}$

65

Step-4

$\therefore$ Orthogonal Basis is :-

$$= \{(1,0,1)\ \ (1,0,-1)\ \ (0,3,0)\}$$
$$\quad \alpha_1 \qquad\quad \alpha_2 \qquad\quad \alpha_3$$

Step-5. Orthonormal basis is :-

$$= \left\{ \frac{\alpha_1}{\|\alpha_1\|}, \frac{\alpha_2}{\|\alpha_2\|}, \frac{\alpha_3}{\|\alpha_3\|} \right\}$$

$$= \left\{ \frac{(1,0,1)}{\sqrt{2}}, \frac{(1,0,-1)}{\sqrt{2}}, \frac{(0,3,0)}{3} \right\}.$$

$$= \left\{ \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right) (0,1,0) \right\}$$

This is Orthonormal Basis

(2) Apply the Gram Schmidtte Process to the vectors

$u_1 = (1,1,1)$ $u_2 = (0,1,1)$ $u_3 = (0,0,1)$

Sol:- Given vectors are

$u_1 = (1,1,1)$

$u_2 = (0,1,1)$

$u_3 = (0,0,1)$

Step-1:- $v_1 = \hat{u}_1 = (1,1,1)$

$\boxed{v_1 = (1,1,1)}$

Step-2:-

$$\boxed{v_2 = u_2 - \frac{\langle u_2 \cdot v_1 \rangle}{\|v_1\|^2} \cdot v_1}$$

$= (0,1,1) - \frac{2}{3} \cdot (1,1,1)$

$= (0,1,1) - \left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right)$

$\frac{1}{1} - \frac{1}{3} - \frac{1}{6}$

$\frac{6-2-1}{6}$

$= \frac{3}{6} = \frac{1}{2}$

$\langle u_2 \cdot v_1 \rangle$

$u_2 = (0,1,1)$

$v_1 = (1,1,1)$

$\langle u_2 \cdot v_1 \rangle =$

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

$\|v_1\|^2 = \sqrt{1^2+1^2+1^2}$

$= \sqrt{3}$.

$\|v_1\|^2 = (\sqrt{3})^2 = 3$.

$\frac{0}{1} - \frac{2}{3}$

$\frac{0-2}{3} = \frac{-2}{3}$

66

$$V_2 = (0,1,1) - \left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right)$$

$$V_2 = \left(0-\frac{2}{3},\ 1-\frac{2}{3},\ 1-\frac{2}{3}\right)$$

$$\boxed{V_2 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)}$$

Step-3 :-

$$\boxed{V_3 = \mu_3 - \frac{\langle \mu_3 . V_1 \rangle}{\|V_1\|^2} . V_1 - \frac{\langle \mu_3 . V_2 \rangle}{\|V_2\|^2} . V_2}$$

$$V_3 = (0,0,1) - \frac{1}{3}(1,1,1) - \frac{\frac{1}{\cancel{3}}}{\frac{2}{\cancel{3}}}\left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

$$= (0,0,1) - (1/3, 1/3, 1/3) - \left(-\frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right)$$

$$\boxed{V_3 = \left(0, -\frac{1}{2}, \frac{1}{2}\right)}$$

$\langle \mu_3 . V_1 \rangle$

$\mu_3 = (0,0,1)$

$V_1 = (1,1,1)$

$\langle \mu_3 . V_1 \rangle$

$= (0+0+1) = \underline{\underline{1}}$.

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

$\|V_1\|^2 = \sqrt{1^2+1^2+1^2}$

$= \sqrt{3}$

$(\sqrt{3})^2 = \underline{3}$.

$\|V_2\|^2 = \sqrt{\left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2}$

$= \sqrt{\frac{4}{9}+\frac{1}{9}+\frac{1}{9}} = \sqrt{\frac{6}{9}} = \frac{2}{3}$ //

$\langle \mu_3 . V_2 \rangle$

$\mu_3 = (0,0,1)$

$V_2 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$

$= \left(0+0+\frac{1}{3}\right)$

$= \frac{1}{3}$

Thus the vectors are

$$v_1 = (1,1,1)$$

$$v_2 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

$v_3 = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$ are the Orthogonal vectors

$\therefore$ The ~~Orog~~ Orthogonal Basis is.

$$= \left\{ \underset{v_1}{(1,1,1)}\ \underset{v_2}{\left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)}\ \underset{v_3}{\left(0, -\frac{1}{2}, \frac{1}{2}\right)} \right\}.$$

The Orthonormal vectors are:

$$= \left\{ \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \frac{v_3}{\|v_3\|} \right\}$$

$$q_1 = \frac{v_1}{\|v_1\|} = \frac{(1,1,1)}{\sqrt{3}} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

$$q_2 = \frac{v_2}{\|v_2\|} = \frac{\left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)}{\frac{\sqrt{6}}{3}} = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$$

$$q_3 = \frac{v_3}{\|v_3\|} = \frac{\left(0, -\frac{1}{2}, \frac{1}{2}\right)}{1/\sqrt{2}} = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

67

$$\|v_2\| = \sqrt{\left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2}$$

$$= \sqrt{\frac{4}{9} + \frac{1}{9} + \frac{1}{9}}$$

$$= \sqrt{\frac{6}{9}} = \frac{2}{3}$$

$$= \frac{\sqrt{6}}{3}$$

(3) Apply gram schmidt Process to the Vectors

$$\underset{u_1}{(1,0,1)} \quad \underset{u_2}{(1,1,1)} \quad \underset{u_3}{(-1,1,0)}$$

Sol: $\omega_1 = (1,0,1)$

$$\omega_2 = u_2 - \frac{\langle u_2 \cdot \omega_1 \rangle}{\langle \omega_1\ \omega_1 \rangle} \cdot \omega_1$$

$$= (1,1,1) - \frac{2}{2}(1,0,1) = (0,1,0)$$

$\langle u_2 \cdot \omega_1 \rangle$

$u_2 = (1,1,1)$

$\omega_1 = (1,0,1)$

$= (1+0+1)$

$= 2$.

$\omega_2 = (0,1,0)$

$$w_3 = u_3 - \frac{\langle u_3\, w_1\rangle}{\langle w_1, w_1\rangle}\cdot w_1 - \frac{\langle u_3\, w_2\rangle}{\langle w_2\, w_2\rangle}\cdot w_2$$

$$= (-1,1,0) - \frac{(-1)}{2}(1,0,1) - \frac{1}{1}(0,1,0)$$

$$= (-1,1,0) + \left(\frac{1}{2}, 0, \frac{1}{2}\right) + (0,-1,0)$$

$$= \left(-\frac{1}{2}, 0, \frac{1}{2}\right)$$

$\therefore \left\{ (1,0,1)\ (0,1,0),\ \left(-\frac{1}{2}, 0, \frac{1}{2}\right)\right\}$ is an

Orthogonal basis of V

(4) $(0,0,1)\ (0,1,1)\ (1,1,1)$

(5) $(2,2,1)\ (1,3,1)\ (1,2,2)$

(68)

• The Gram Schmidt Process

To Convert a basis vectors $u_1\ u_2 \ldots u_n$ into an Orthogonal vectors perform the following Computations

Step-1 :- $v_1 = u_1$

Step-2 :- $v_2 = u_2 - \dfrac{\langle u_2 \cdot v_1 \rangle}{\|v_1\|^2} \cdot v_1$

Step-3 :- $v_3 = u_3 - \dfrac{\langle u_3 \cdot v_1 \rangle}{\|v_1\|^2} \cdot v_1 - \dfrac{\langle u_3 \cdot v_2 \rangle}{\|v_2\|^2} \cdot v_2$

Step-4 :- $v_4 = u_4 - \dfrac{\langle u_4 \cdot v_1 \rangle}{\|v_1\|^2} \cdot v_1 - \dfrac{\langle u_4 \cdot v_2 \rangle}{\|v_2\|^2} \cdot v_2 - \dfrac{\langle u_4 \cdot v_3 \rangle}{\|v_3\|^2} \cdot v_3 \ldots$

The vectors Obtained $v_1, v_2, v_3 \ldots$ are Orthogonal vectors.

This Process is called as Gram Schmidt Process.

Gram Schmidt

→ This Process is used to Convert set of all Ordinary vectors into Orthonormal vectors

→ Orthogonal vectors :-

The vectors $v_1 . v_2 . . . . v_n$ are Said to be Orthogonal If the inner Product of any two different vectors equals to Zero.

ie:- $< v_i . v_j > = 0 \quad \forall \; i \neq j$

(Inner Product (or) dot Product)

→ The Gram- Schmidt Process (or Procedure) is a sequence of Operations that allow us to transform a set of Linearly independent vectors into a set of Orthonormal vectors.

→ The Gram schmidt Process is used to transform a set of linearly independent vectors into a set of Orthonormal vectors forming an Orthonormal basis.

69

- → The Gram schmidt algorithm makes it Possible to Construct, for each list of linearly independent vectors (basis), a Corresponding Orthonormal list (Orthonormal basis)

→ we say that 2 vectors are Orthogonal If they are $\perp r$ to each Other

ie :- The dot Product of the two vectors is Zero.

Single Variable Distribution

Variable:-

A variable is a quality which changes or varies. the change may occur due to time factor or any factor.

Eg:- Height of the person with age, Height and weight of the person.

Variable is of 2 types:

1) Discrete 2) Continuous

Random Variable:-

A real variable x whose value is determined by the outcome of a random experiment is called a random variable. Random variables are two types

1) Discrete Random variable

2) Continuos Random Variable

1) **Discrete Random Variable:-**

A random variable x which can takes only a finite number of discrete values in an interval of domain is called a discrete random variable.

Eg:- Tossing of coin, Tossing of die etc.

2) **Continuous Random Variable:**

A random variable x which can take values continuosly then the variable is called continuos random variable.

Eg:- Temperate, Time, Height, Age etc.

Probability function of a discrete random variable:-

If for a discrete random variable "x" the real value function $p(x)$

i.e. $p(X=x) = p(x)$

Properties :-

(i) $p(x) \geq 0$

(ii) $\sum_{i=1}^{n} p(x_i) = 1$

(iii) $p(x)$ lies between 0 and 1

(iv) $p(x)$ cannot be negative for any value of x.

Probability distribution function:-

$$F(x) = P(X \leq x)$$

Properties:-

(i) $0 \leq F(x) \leq 1$

(ii) $F(-\infty) = 0.$

(iii) $F(\infty) = ~~\infty~~ 1$

Cummulative distributive function of a discrete random variable :-

$$F(x) = P(X \leq x)$$

Probability density function:

$$f(x) = \frac{d}{dx}\left[F(x)\right]$$

Expectation, Mean, Variance and standard deviation of a discrete random variable :-

Expectation: $E(x) = \sum_{i=1}^{n} p_i x_i$

Properties:-

(i) $E(x+k) = E(x)+k$

(ii) $E(x \pm y) = E(x) \pm E(y)$

(iii) $E(kx) = k \cdot E(x)$

(iv) $E(ax \pm b) = aE(x) \pm b.$

(v) $E(xy) = E(x)E(y)$

(vi) $E\left(\frac{1}{x}\right) = \frac{1}{E(x)}$

Mean:-

$\mu = E(X)$

$\mu = \sum_{i=1}^{n} P_i x_i$

Variance:-

$Var(X)\ (\sigma^2) = E(X^2) - [E(X)]^2$

$= E(X^2) - \mu^2$

Standard deviation:-

$S.D = \sqrt{Variance}$

$= \sqrt{E(x^2) - \mu^2}$

properties:-

(i) $V(K) = 0$

(ii) $V(KX) = K^2 V(X)$

(iii) $V(X+K) = V(X)$.

(iv) $V(aX+K) = a^2 V(X)$

(v) $V(X+Y) = V(X) + V(Y)$

Problems:-

1) A random variable X has the following probability function.

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

(i) Determine k

(ii) Evaluate probability $P(X<6)$, $P(X \geq 6)$, $P(0<X<5)$, $P(0 \leq X \leq 4)$

(iii) Mean (iv) Variance (v) Distribution function of x

(vi) If $P(X \leq k) > \frac{1}{2}$ then find k.

Sol:- (i) $\sum_{i=1}^{n} P(x_i) = 1$

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

$9K + 10K^2 = 1$

$K = 1/10.$

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

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

$= 8k+k^2$

$= \frac{8}{10}+\frac{1}{100}$

$= \frac{81}{100} = 0.81$

$P(x\geq 6) = P(x=6)+P(x=7)$

$= 2k^2+7k^2+k$

$= 9k^2+k$

$= \frac{9}{100}+\frac{1}{10}$

$= \frac{19}{100} = 0.19$

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

$= k+2k+2k+3k$

$= 8k$

$= \frac{8}{10} = 0.8$

$P(0\leq x\leq 4) = P(x=1)+P(x=2)+P(x=3)+P(x=4)$

$= k+2k+2k+3k$

$= 8k = \frac{8}{10} = 0.8$

(iii) Mean:-

$\mu = \sum_{i=0}^{7} p_i x_i$

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

$= k+4k+6k+12k+5k^2+12k^2+49k^2+7k$

$= 30k+66k^2$

$= \frac{30}{10}+\frac{66}{100}$

$= 3.66$

(iv) Variance:-

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

$$= \sum_{i=0}^{7} p_i x_i^2 - \mu^2$$

$$= [0(0)^2 + k(1)^2 + 2k(2^2) + 2k(3^2) + 3k(4)^2 + k^2(5^2) + (2k^2)(6)^2 + (7k^2+k)(7^2)] - (3.66)^2$$

$$= k + 4k + 18k + 48k + 25k^2 + 72k^2 + 343k^2 + 49k - (3.66)^2$$

$$= 124k + 440k^2 - (3.66)^2$$

$$= \left[\frac{124}{10} + \frac{440}{100}\right] - (3.66)^2$$

$$= 3.4044$$

(v) Distribution function of x:-

X	$F(x) = P(X \leq x)$
0	0
1	$k = 1/10 = 0.1$
2	$3k = 3/10 = 0.3$
3	$5k = 0.5$
4	$8k = 0.8$
5	$8k + k^2 = 0.81$
6	$8k + 3k^2 = 0.83$
7	$9k + 10k^2 = 1$

(vi) If $P(X \leq k) > \frac{1}{2}$

We know that $P(X \leq 4) > 0.5$

$\therefore k = 4$

2) The probability density function of a variable x is as follows

x	0	1	2	3	4	5	6
p(x)	k	3k	5k	7k	9k	11k	13k

(i) Determine k

(ii) Evaluate $p(x<4)$, $p(x\geq5)$, $p(3<x\leq6)$

(iii) Mean

(iv) Variance

Sol:- (i) $\sum_{i=1}^{n} p(x_i)=1$

$$k+3k+5k+7k+9k+11k+13k=1$$

$$49k=1$$

$$k=\frac{1}{49}=0.02$$

(ii) $p(x<4)=p(x=0)+p(x=1)+p(x=2)+p(x=3)$

$$=k+3k+5k+7k$$

$$=16k$$

$$=\frac{16}{49}=16(0.02)$$

$$=0.32$$

$p(x\geq5)=p(x=5)+p(x=6)$

$$=11k+13k$$

$$=24k$$

$$=24(0.02)$$

$$=(0.48)$$

$p(3<x\leq6)=p(x=4)+p(x=5)+p(x=6)$

$$=9k+11k+13k$$

$$=33k=33(0.02)$$

$$=0.66.$$

(iii) Mean:

$$\mu = \sum_{i=0}^{6} p_i x_i$$

$= 0(k) + 1(3k) + 2(5k) + 3(7k) + 4(9k) + 5(11k) + 6(13k)$

$= 0 + 3k + 10k + 21k + 36k + 55k + 78k$

$= 203k$

$= 203(0.02)$

$= 4.06.$

(iv) Variance :-

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

$= k(0^2) + 3k(1^2) + 5k(2^2) + 7k(3^2) + 9k(4^2) + 11k(5^2) + 13k(6^2) - \mu^2$

$= [k + 3k + 20k + 63k + 144k + 275k + 468k] - \mu^2$

$= 973k - \mu^2 = 973(0.02) - \mu^2$

$= 2.9764.$

3) Let X denote the number of heads in a single toss of 4 coins. Determine

(i) $P(X<2)$ (ii) $P(1<X\leq 3)$ (iii) Mean (iv) Variance

(v) Standard deviation.

Sol:- Given no. of coins tossing at a time = 4

No. of possibilities $= 2^4 = 16$.

Let x denote the number of heads.

The required probability of getting heads as follows

x	0	1	2	3	4
$P(x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

$\frac{4C_0}{2^4}, \frac{4C_1}{2^4}, \frac{4C_2}{2^4}, \frac{4C_3}{2^4}, \frac{4C_4}{2^4}$

(i) $P(x<2) = P(x=0) + P(x=1)$

$$= \frac{1}{16} + \frac{4}{16} = \frac{5}{16}$$

(ii) $P(1<x\leq 3) = P(x=2) + P(x=3)$

$$= \frac{6}{16} + \frac{4}{16}$$

$$= \frac{10}{16} = \frac{5}{8}$$

(iii) Mean $\mu = \sum_{i=0}^{4} p_i x_i$

$$= 0 + \frac{4}{16} + \frac{12}{16} + \frac{12}{16} + \frac{4}{16}$$

$$= \frac{32}{16}$$

$$= 2$$

(iv) Variance $\sigma^2 = \sum_{i=0}^{4} p_i x_i^2 - \mu^2$

$$= \left(0 + \frac{4}{16} + \frac{24}{16} + \frac{36}{16} + \frac{16}{16}\right) - (2)^2$$

$$= 5 - 4$$

$$= 1$$

(v) standard deviation $= \sqrt{\text{variance}}$

$$= \sqrt{1} = 1$$

4) A sample of 4 items is selected at random from a box containing 12 items of which 5 are defective. Find the expected number of defective items.

Sol:- Total number of items $= 12$

Number of Sample items selected $= 4$

Number of ways selected from 12 of 4 $= {}^{12}C_4$

$${}^{12}C_4 = 495$$

Let x denote the number of defective items.

Number of defective items = 5

Number of good items = 7

The required probability of getting defective items as follows:

$$P(x=0) = \frac{5C_0 \times 7C_4}{12C_4} = \frac{7}{99}$$

$$P(x=1) = \frac{5C_1 \times 7C_3}{12C_4} = \frac{35}{99}$$

$$P(x=2) = \frac{5C_2 \times 7C_2}{12C_4} = \frac{42}{99}$$

$$P(x=3) = \frac{5C_3 \times 7C_1}{12C_4} = \frac{14}{99}$$

$$P(x=4) = \frac{5C_4 \times 7C_0}{12C_4} = \frac{1}{99}.$$

X	0	1	2	3	4
P(x)	$\frac{7}{99}$	$\frac{35}{99}$	$\frac{42}{99}$	$\frac{14}{99}$	$\frac{1}{99}$

Expected Value = $E(x) = \sum_{i=0}^{4} p_i x_i$

$$= 0 + \frac{35}{99} + \frac{84}{99} + \frac{42}{99} + \frac{4}{99}$$

$$= \frac{165}{99} = \frac{5}{3} = 1.667$$

5) A Sample of 3 items is selected at random from a box containing 10 items of 4 are defective. Find the expected value.

Sol:-

Total number of items = 10.

No. of Sample items selected = 3

Number of ways selected from 10 of 3 = $10C_3$

$$10C_3 = 120$$

Number of good items = 6

Number of defective = 4

Let x be denote number of defective.

The required probabilities are as follows:

$$P(x=0) = \frac{4C_0 \times 6C_3}{10C_3} = \frac{1}{6}$$

$$P(x=1) = \frac{4C_1 \times 6C_2}{10C_3} = \frac{1}{2}$$

$$P(x=2) = \frac{4C_2 \times 6C_1}{10C_3} = \frac{3}{10}$$

$$P(x=3) = \frac{4C_3 \times 6C_0}{10C_3} = \frac{1}{30}$$

x	0	1	2	3
P(x)	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{30}$

$\therefore$ Expected Value $= E(x) = \sum_{i=0}^{3} p_i x_i$

$$= 0 + \frac{1}{2} + \frac{6}{10} + \frac{3}{30}$$

$$= \frac{1}{2} + \frac{3}{5} + \frac{1}{10}$$

$$= 0.5 + 0.6 + 0.1$$

$$= 1.2$$

6) Find the mean of the probability distribution of the number of heads obtained in tossing 3 coins.

Sol:- Let X denotes the number of heads

Number of coins tossed at a time = 3

Total possibilities $= 2^3 = 8$

The required probability distribution is as follows:

x	0	1	2	3
$p(x)$	1/8	3/8	3/8	1/8

Expected Value $= E(x) = \frac{3}{8} + \frac{6}{8} + \frac{3}{8}$

$$= \cancel{\frac{9}{8}} = \frac{12}{8} = \frac{3}{2} = 1.5$$

Continuos probability distribution:-

Let $f(x)$ be a continuous function in the interva (a,b) is called continuous probability distribution. It is denoted by $\int_a^b f(x)\,dx$.

properties:-

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$
3. $P(a \leq x \leq b) = \int_a^b f(x)\,dx$

Cummulative distribution function of a continuous random variable:-

$$F(x) = P(X \leq x)$$
$$= \int_{-\infty}^{\infty} f(x)\,dx.$$

Properties:-

1. $0 \leq F(x) \leq 1$
2. $F'(x) = f(x)$
3. $F(-\infty) = 0$
4. $F(\infty) = 1$

Measures of central tendency for continuos probability distribution:-

Mean:-

$$\mu = E(x) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$$

Median:- Median is the point which divides the entire distibution in to two equal parts. Suppose the median point is taken as M then

$$\int_a^M F(x)dx = \int_M^b f(x)dx = \frac{1}{2} \qquad \text{Here } [a,b]$$

<u>Mode</u>:- Mode is the value of x for which $f(x)$ is maximum. Mode is calculated by $F'(x)=0$, $f''(x)<0$ for $a<x<b$.

<u>Variance</u>:- Variance $\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$

<u>Standard deviation</u>:-

$$S.D \quad \sigma = \sqrt{var(x)}$$

$$= \sqrt{\int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2}$$

<u>Mean deviation</u>:-

Mean deviation about the mean μ is given by $\int_{-\infty}^{\infty} [x-\mu] f(x)dx$.

7) If a continuous random variable has the probability density function $f(x)$ has $f(x) = \begin{cases} 2e^{-2x} & \text{for } x>0. \\ 0 & \text{for } x\leq 0 \end{cases}$ Find the probabilities

(i) between 1 and 3 (ii) greater than 0.5

<u>Sol</u>:- Given,

$$f(x) = \begin{cases} 2e^{-2x} & \text{for } x>0 \\ 0 & \text{for } x\leq 0 \end{cases}$$

(i) $P(1<x<3) = \int_{-\infty}^{\infty} f(x)\,dx$

$$= \int_{-\infty}^{0} (0)dx + \int_{0}^{\infty} 2e^{-2x}dx$$

$$= 0 + \int_1^3 2e^{-2x}dx$$

$$= 2\int_1^3 e^{-2x}dx$$

$$= 2\left[\frac{e^{-2x}}{-2}\right]_1^3$$

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

$$= \frac{2}{-2}\left[e^{-6} - e^{-2}\right]$$

$$= -\left[e^{-6} - e^{-2}\right]$$

$$= e^{-2} - e^{-6} = 0.132$$

(ii) $$P(x>0.5) = \int_{0.5}^{\infty} f(x)\,dx$$

$$= \int_{0.5}^{\infty} 2e^{-2x}\,dx$$

$$= 2\left(\frac{e^{-2x}}{-2}\right)_{0.5}^{\infty}$$

$$= -\left(e^{-\infty} - e^{-1}\right)$$

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

$$= \frac{1}{e} = 0.367$$

8) The probability density function $f(x)$ of a continuous random variable is given by $f(x) = ce^{-|x|}$, $-\infty < x < \infty$. To find (i) c (ii) Mean (iii) Variance (iv) $P(0<x<4)$

Sol:- Given,

$$f(x) = ce^{-|x|}$$

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

$$\int_{-\infty}^{\infty} ce^{-|x|} = 1$$

$$2c\int_{0}^{\infty} e^{-x}\,dx = 1$$

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

$$-2c\left[e^{-\infty} - e^{-0}\right] = 1$$

$$-2c\,[0-1] = 1$$

$$2c = 1 \Rightarrow c = \frac{1}{2}$$

$$\therefore f(x) = \frac{1}{2} e^{-|x|}$$

(ii) Mean :-

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

$$= \int_{-\infty}^{\infty} x \frac{1}{2} e^{-|x|} dx$$

$$= \frac{1}{2}\left[\int_{-\infty}^{\infty} x e^{-|x|} dx\right]$$

$\mu = 0$ since it is an odd function.

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

$$= \int_{-\infty}^{\infty} x^2 \frac{1}{2} e^{-|x|} dx - 0^2$$

$$= \cancel{2} \times \frac{1}{\cancel{2}} \int_{0}^{\infty} x^2 e^{-|x|} dx$$

$$= \int_{0}^{\infty} x^2 e^{-x} dx$$

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

$$= x^2 e^{-x} - \int_{0}^{\infty} e^{-x}(2x)\, dx$$

$$= (x^2 e^{-x})_0^{\infty} - 2\left[\left[x e^{-x}\right]_0^{\infty} + \left[e^{-x}\right]_0^{\infty}\right]$$

$$= -2\left[e^{-\infty} - e^{-0}\right]$$

$$= -2[0 - 1]$$

$$= 2$$

(iv) $P(0 < x < 4) = \int_{0}^{4} f(x)\, dx$

$$= \int_{0}^{4} \frac{1}{2} e^{-|x|} dx$$

$$= \frac{1}{2}\left(\frac{e^{-x}}{-1}\right)_0^{1}$$

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

$= \frac{-1}{2}\left[e^{-4}-1\right]$

$= \frac{1}{2}\left[1-e^{-4}\right]$

$= 0.4908.$

→ The probability density function of a random variable 'x' is

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

find (i) $E(x)$ (ii) $E(x^2)$ (iii) $Var(x)$ (iv) standard deviation.

Sol:-

(i) $E(x) = \int_{-\infty}^{\infty} x e^{-x} dx$ $\quad \because E(x) = \int_{-\infty}^{\infty} x f(x) dx$

$= \int_{-\infty}^{0} x(0)dx + \int_{0}^{\infty} x e^{-x} dx$

$= \int_{0}^{\infty} x e^{-x} dx$

$= \left[x\left[\frac{e^{-x}}{-1}\right] - (1)\left(\frac{e^{-x}}{(-1)(-1)}\right)\right]_0^{\infty} = (0-0)-(0-1) = 1$

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

$= \int_{-\infty}^{0} x^2(0)dx + \int_{0}^{\infty} x^2 e^{-x} dx$ $\quad \int_0^{\infty} uv = u\int v - \int\left[\frac{d}{dx}u\int v\right]$

$= \int_{0}^{\infty} x^2 e^{-x} dx$

$= x^2\left(\frac{e^{-x}}{-1}\right) - 2x\left(\frac{e^{-x}}{(-1)(-1)}\right) + 2\left(\frac{e^{-x}}{(-1)(-1)(-1)}\right)\Big]_0^{\infty}$

$= (0-0+0)-(0-0+2) = 2$

(iii) $Var(x) = E(x^2) - (E(x))^2$

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

(iv) Standard deviation, $\sigma = \sqrt{Var(x)}$

$= \sqrt{1}$

$= 1$

→ Suppose a continuous Random Variable x has the probability density function.

$$f(x) = \begin{cases} K(1-x^2) & \text{for } 0<x<1 \\ 0 & \text{otherwise} \end{cases}$$

Find (i) K (ii) mean (iii) Variance

Sol:-

(i) We know that $\int_{-\infty}^{\infty} f(x)dx = 1$ ∵ Total probability.

$$\int_{-\infty}^{0} f(0)dx + \int_{0}^{\infty} K(1-x^2)dx + \int_{1}^{\infty} (0)dx = 1$$

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

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

$$K\left[(1-1/3) - (0-0)\right] = 1$$

$$K[2/3] = 1$$

$$K = 3/2$$

(ii) Mean, $\mu = E(x)$

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

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

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

$$= 3/2 \int_{0}^{1} x - x^3 \, dx$$

$$= 3/2 \left(\frac{x^2}{2} - \frac{x^4}{4}\right)_0^1$$

$$= 3/2\left[\left(\frac{1}{2} - \frac{1}{4}\right) - (0-0)\right]$$

$$= 3/2\left[\frac{2-1}{4}\right]$$

$$= 3/2[1/4]$$

$$= 3/8$$

(iii) Variance $= E(x^2) - \mu^2$

$$= \int_{-\infty}^{\infty} x^2 f(x)\,dx - (3/8)^2$$

$$= \int_0^1 x^2\, 3/2(1-x^2)\,dx - (3/8)^2$$

$$= 3/2 \int_0^1 x^2(1-x^2)\,dx - (3/8)^2$$

$$= 3/2 \left[x^3/3 - \frac{x^5}{5} \right]_0^1 - (3/8)^2$$

$$= 3/2 \left[(1/3 - 1/5) - (0-0) \right] - (3/8)^2$$

$$= 3/2 \left[\frac{5-3}{15} \right] - (3/8)^2$$

$$= 3/2 \left[2/15 \right] - (9/64).$$

$$= 6/30 - 9/64$$

$$= 1/5 - 9/64$$

$$= \frac{19}{320}$$

→ Suppose for a continuous Random Variable 'x' is $f(x) = \begin{cases} kx^2e^{-x} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$ find (i) k (ii) Mean (iii) Variance (iv) S.D

Sol:- Given that, $f(x) = \begin{cases} kx^2e^{-x} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$

(i) We know that $\int_{-\infty}^{\infty} f(x)\,dx = 1$

$$\int_{-\infty}^{0} (0)\,dx + \int_0^{\infty} kx^2e^{-x}\,dx = 1$$

$$k\int_0^{\infty} x^2e^{-x}\,dx = 1 \Rightarrow k\left[x^2 \int_0^{\infty} -e^{-x}\,dx - \frac{d}{dx}(x^2) \int_0^{\infty} e^{-x}\,dx\right] = 1$$

$$k\left[x^2\left(\frac{e^{-x}}{-1}\right) - 2x\left(-e^{-x}\right)\right]_0^{\infty} = 1$$

$$k(2) = 1$$

$$k = 1/2$$

(ii) Mean $= E(x)$

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

$$= \int_{-\infty}^{0} x(0) dx + \int_{0}^{\infty} x^2 k e^{-x} dx$$

$$= \int_{0}^{\infty} (1/2) x^3 e^{-x} dx$$

$$= 1/2 \int_{0}^{\infty} x^3 e^{-x} dx = 1/2 \left[x^2\left(\frac{e^{-x}}{-1}\right) - \int \frac{d}{dx} x^2 \int e^{-x}\right] dx$$

$$= 1/2 \left[x^2\left(\frac{e^{-x}}{-1}\right) - \int 3x \left(\frac{e^{-x}}{-1}\right)\right] dx$$

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

$$= 1/2 [0-0-0-0] - [0-0-0-6] = 1/2 (6)$$

$$= 1/2 (6)$$

$$= 3$$

(iii) Variance $E(x)^2 = \mu^2$

$$= \int_{-\infty}^{\infty} x^2 f(x) dx - (3)^2$$

$$= \int_{0}^{\infty} x^2 (1/2) [x^2 e^{-x}] dx - 9$$

$$= 1/2 [x^4 e^{-x} - x^2 e^{-x}] dx - 9$$

$$= 1/2 \left[x^4 \frac{e^{-x}}{-1} - 4x^3 \frac{e^{-x}}{(-1)(-1)} + 2x^2 \frac{e^{-x}}{(-1)(-1)(-1)} - 24x \frac{e^{-x}}{(-1)(-1)(-1)(-1)} + 24 \frac{e^{-x}}{(-1)(-1)(-1)(-1)(-1)}\right)_0^{\infty} - (3)^2$$

$$= 1/2 [0-0-0-0-0] - [0-0-0-0-24] - 9$$

$$= 1/2 [24] - 9$$

$$= 12 - 9 = 3$$

(iv) Standard deviation ; $\sigma = \sqrt{Var(x)}$

$$= \sqrt{3}$$

$$= 1.732$$

→ For a continuous random variable x is

$$f(x) = \begin{cases} cx(2-x), & \text{for } 0<x<2 \\ 0 & \text{, otherwise} \end{cases}$$

find (i) c (ii) mean (iii) variance

Sol:- Given $f(x) = \begin{cases} cx(2-x), & \text{for } 0<x<2 \\ 0 & \text{, otherwise} \end{cases}$

we know that

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

$$\int_{-\infty}^{0} (0)dx + \int_{0}^{2} cx(2-x)dx + \int_{2}^{\infty} (0)dx = 1$$

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

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

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

$$= c\left[\left[4 - 8/3\right] - (0-0)\right] = 1$$

$$\Rightarrow c\left[4/3\right] \Rightarrow c = 3/4$$

Mean $\mu = E(x)$

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

$$= \int_{0}^{\infty} x\left(x(2-x)\right)dx$$

$$= 3/4 \int_{0}^{\infty} x^2(2-x)dx$$

$$= 3/4 \int_{0}^{\infty} (2x^2 - x^3)dx$$

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

$$= 3/4 \left[2/3(8) - \frac{16}{4}) - (0-0)\right]$$

$$= 3/4 \left[16/3 - 16/4\right]$$

$$= 3/4 \left[4/3\right] = 1$$

(iii) Variance

$$V(x) = E(x^2) - \mu^2$$

$$= \int_{-\infty}^{\infty} x^2 f(x)\,dx - (1)^2$$

$$= \int_0^2 x^2 (x(2-x))\,dx - 1$$

$$= 3/4 \int_0^2 (2x^3 - x^4)\,dx - 1$$

$$= 3/4 \left[2\cdot\frac{x^4}{4} - \frac{x^5}{5}\right] - 1$$

$$= 3/4 \left[16/2 - 32/5\right] - 1$$

$$= 3/4 \left[8 - 32/5\right] - 1$$

$$= 3/4 \left[8/5 - 1\right]$$

$$= 6/5 - 1$$

$$= 1/5$$

→ A continuas random variable X has the distribution function

$$F(x) = \begin{cases} 0 & \text{if } x \le 1 \\ k(x-1)^4 & \text{if } 1 < x < 3 \\ 1 & \text{if } x > 3 \end{cases}$$

(i) find $f(x)$ (ii) k (iii) Mean

Sol:- Given,

$$F(x) = \begin{cases} 0 & \text{if } x \le 1 \\ k(x-1)^4 & \text{if } 1 < x < 3 \\ 1 & \text{if } x > 3 \end{cases}$$

(i) $$f(x) = \frac{d}{dx}\left[F(x)\right]$$

$$f(x) = \begin{cases} 0 & \text{if } x \le 1 \\ 4k(x-1)^3 & \text{if } 1 < x < 3 \\ 0 & \text{if } x > 3 \end{cases}$$

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

$$\int_{-\infty}^{1}(0)dx + \int_{1}^{3} 4k(x-1)^3 dx + \int_{3}^{\infty}(0)dx = 1$$

$$0 + 4k\left[\int_{1}^{3}(x-1)^3dx\right] + 0 = 1$$

$$4k\left[\int_{1}^{3}(x-1)^3dx\right] = 1$$

$$4k\left(\frac{(x-1)^4}{4}\right)_1^3 = 1$$

$$k\left((3-1)^4 - (1-1)^4\right) = 1.$$

$$k\left((2)^4\right) = 1$$

$$16k = 1$$

$$k = 1/16.$$

(iii) Mean $\mu = E(x)$

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

$$= \int_{-\infty}^{1}(0)dx + \int_{1}^{3} x \cdot \frac{4}{16}(x-1)^3 dx + 0\,dx.$$

$$= \int_{1}^{3} x \cdot \frac{4}{16}(x-1)^3 dx$$

$$= \frac{1}{4}\int_{1}^{3} x(x-1)^3 dx$$

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

$$= \frac{1}{4}\left[\int_{1}^{3}(x^4 - x - 3x^3 + 3x^2)dx\right]$$

$$= \frac{1}{4}\left[\frac{x^5}{5} - \frac{x^2}{2} - \frac{3x^4}{4} + \frac{3x^3}{3}\right]_1^3$$

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

$$= \frac{1}{4}\left[\left[\frac{243}{5} - \frac{9}{2} - \frac{3(81)}{4} + \frac{3(27)}{3}\right] - \left[\frac{1}{5} - \frac{1}{2} - \frac{3}{4} + \frac{3}{3}\right]\right]$$

$$= \frac{13}{5}$$

→ If X is a continuos random variable and $Y = aX+b$

prove that $E(Y) = aE(X)+b$ and $V(Y) = a^2 V(X)$

Sol:- Given $Y = aX+b$

We know that $E(X) = \int_{-\infty}^{\infty} x f(x) dx$

$$E(Y) = E(aX+b)$$

$$= \int_{-\infty}^{\infty} (ax+b) f(x) dx$$

$$= \int_{-\infty}^{\infty} ax f(x) dx + \int_{-\infty}^{\infty} b f(x) dx$$

$$= a \int_{-\infty}^{\infty} x f(x) dx + b \int_{-\infty}^{\infty} f(x) dx$$

$$= aE(X)+b$$

$$\therefore E(Y) = aE(X)+b \longrightarrow ①$$

$$Y = aX+b \longrightarrow ②$$

$$② - ① \Rightarrow Y - E(Y) = a[X - E(X)]$$

Squaring on both sides

$$[Y-E(Y)]^2 = a^2[X-E(X)]^2$$

Taking Expectation

$$E(Y-E(Y))^2 = a^2 E[X-E(X)]^2$$

$$V(Y) = a^2 V(X) \;//$$

→ If X is a continuos random variable and k is constant then prove that $V(X+k) = V(X)$.

(ii) $V(kX) = k^2 V(X)$

Sol:- We know that

$$V(X) = E(X^2) - (E(X))^2$$

$$= \int_{-\infty} x^2 f(x)\,dx - \left[\int_{-\infty} x f(x)\,dx\right]$$

$$V(x+k) = \int_{-\infty}^{\infty} (x+k)^2 f(x)\,dx - \left[\int_{-\infty}^{\infty} (x+k) f(x)\,dx\right]^2$$

$$= \int_{-\infty}^{\infty} (x^2+2k+k^2) f(x)\,dx - \left[\int_{-\infty}^{\infty} x f(x)\,dx + k\int_{-\infty}^{\infty} f(x)\,dx\right]^2$$

$$= \left[\int_{-\infty}^{\infty} x^2 f(x)\,dx + 2k\int_{-\infty}^{\infty} x f(x)\,dx + k^2\int_{-\infty}^{\infty} f(x)\,dx\right] - \left[\int_{-\infty}^{\infty} x f(x)\,dx + k\int_{-\infty}^{\infty} f(x)\,dx\right]^2$$

$$= \left[\int_{-\infty}^{\infty} x^2 f(x)\,dx + 2k\int_{-\infty}^{\infty} x f(x)\,dx + k^2\int_{-\infty}^{\infty} f(x)\,dx\right] - \left[(E(x)) + k\right]^2$$

$$= \left[E(x^2) + 2kE(x) + k^2(1)\right] - \left[(E(x))^2 + k^2 + 2kE(x)\right]$$

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

$$V(x+k) = V(x)_{//}$$

(ii) $$V(kx) = \left[E(k^2x^2) - [E(kx)]^2\right]$$

$$= \int_{-\infty}^{\infty} x^2k^2 f(x)\,dx - \left[\int_{-\infty}^{\infty} kx f(x)\,dx\right]^2$$

$$= k^2\left[\int_{-\infty}^{\infty} x^2 f(x)\,dx\right] - k^2\left[\int_{-\infty}^{\infty} x f(x)\,dx\right]^2$$

$$= k^2\left[E(x^2) - [E(x)]^2\right]$$

$$= k^2 V(x)$$

$$V(kx) = k^2 V(x)_{//}$$

Distributions:-

There are two types of theoritical distributions

(i) Discrete theoretical distibutions

(2) Continuos theoretical distributions

1) Discrete theoretical distribution:

In this we study Binomial, poisson distributions

2) Continuous theoretical distributions:-

In this we study Normal, Uniform, Exponential distributions.

Bernoulli Distribution:-

A random Variable x which takes two values 0 and 1 with probability p and q respectively i.e.;

$P(x) = 1 = P$

$P(x) = 0 = q$

$q = 1-p$ is called Bernoulli Distribution.

This is shown as

$$P(x) = p^x q^{1-x} \quad (\text{where } x = 0, 1)$$

$$= p^x (1-p)^{1-x}$$

Bernoulli's Theorem:-

$$P(X=r) = n_{C_r} p^r q^{n-r}$$

where $p+q=1$

Binomial Distribution:-

It was discovered by James Bernoulli in the year 1700 and it is a discrete probability distribution

Eg:- Tossing of a coin, Birth of baby.

Definition:-

A random Variable X has a binomial distribution if it assumes only non negative values and its probability density function is

$$P(X=r) = n_{C_r} p^r q^{n-r} \quad (\text{where } r = 0, 1, 2 \ldots n)$$

Conditions of Binomial Distribution:-

There are n independent trails. Each trail has two possible outcomes. The probabilities of two outcomes are constant.

Mean of the Binomial Distribution:

Mean $\mu = np$

we know that the Binomial distribution

$$p(r) = {}^nC_r\, p^r q^{n-r} \quad (r = 0, 1, 2 \dots n)$$

$$\text{Mean } \mu = \sum_{r=0}^{n} r\, p(r)$$

$$= \sum_{r=0}^{n} r\, {}^nC_r\, p^r q^{n-r}$$

$$= 0 + {}^nC_1\, p q^{n-1} + 2\, {}^nC_2\, p^2 q^{n-2} + \dots + n \cdot {}^nC_n\, p^n q^0$$

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

$$= np\left[q^{n-1} + (n-1) p q^{n-2} + \dots + p^{n-1}\right]$$

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

$$= np(1) = np$$

Variance of Binomial Distribution:-

$\sigma^2 = npq$

we know that $p(r) = {}^nC_r\, p^r q^{n-r}, \ (r = 0, 1, 2 \dots n)$

$$\text{Variance } \sigma^2 = \sum_{r=0}^{n} r^2 p(r) - (\mu)^2$$

$$= \sum_{r=0}^{n} \left[r(r-1) + r\right] p(r) - n^2 p^2$$

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

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

$$= \left((2)(1)\, {}^nC_2\, p^2 q^{n-2} + (3)(2)\, {}^nC_3\, p^3 q^{n-3} + \dots + n(n-1)\, {}^nC_n\, p^n q^0\right) + np - n^2 p^2$$

$$= \left[(2)(1)\frac{n(n-1)}{2}p^2q^{n-2} + (3)(2)\frac{n(n-1)(n-2)}{(3)(2)}p^3q^{n-3} + \cdots + n(n-1)p^n\right] + np - n^2p^2$$

$$= n(n-1)p^2\left[q^{n-2} + (n-2)pq^{n-3} + \cdots + p^{n-2}\right] + np - n^2p^2$$

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

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

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

$$= np(1-p) = np(q) = npq$$

$$\sigma^2 = npq$$

Mode of the Binomial Distribution:-

Mode of the binomial distribution is the value of 'x' at which $p(x)$ is maximum value.

$$\text{Mode} = \begin{cases} [(n+1)p] & \text{if non-integer} \\ (n+1)p \text{ and } (n+1)p-1 & \text{if integer} \end{cases}$$

Recurrance Relation for Binomial Distribution:-

$$p(r+1) = \frac{(n-r)p}{(r+1)q}\,p(r)$$

proof:

By Binomial distribution we have

$$p(r) = n_{c_r}p^rq^{n-r} \longrightarrow ①$$

$$p(r+1) = n_{c_{r+1}}p^{r+1}q^{n-(r+1)}$$

$$= n_{c_{r+1}}p^{r+1}q^{n-r-1} \longrightarrow ②$$

$$\frac{②}{①} = \frac{p(r+1)}{p(r)} = \frac{n_{c_{r+1}}p^{r+1}q^{n-r-1}}{n_{c_r}p^rq^{n-r}}$$

$$= \frac{\frac{n-r}{r+1}n_{c_r}p^{r+1}q^{n-r-1}}{n_{c_r}p^rq^{n-r}}$$

$$= \frac{n-r}{r+1}pq^{-1} = \left(\frac{n-r}{r+1}\right)\frac{p}{q}$$

1. A fair coin is tossed 6 times. Find the probability of getting 4 heads.

Given $n=6$.

p = probability of getting a heads = $1/2$

$q = 1/2$

$r = 4$ heads

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

$$p(4) = 6c_4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^{6-4}$$

$$= 6c_4 \cdot \frac{1}{2^4} \frac{1}{2^2}$$

$$= \frac{15}{2^6} = \frac{15}{64} //.$$

2. 10 coins are thrown simultaneously. Find the probability of getting (i) atleast 7 heads (ii) 6 heads

Given $n = 10$ coins

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

$$p(r \geq 7) = p(r=7) + p(r=8) + p(r=9) + p(r=10)$$

$$= 10c_7 \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^{10-7} + 10c_8 \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^{10-8} + 10c_9 \left(\frac{1}{2}\right)^9 \left(\frac{1}{2}\right)^{10-9} + 10c_{10} \left(\frac{1}{2}\right)^{10} \left(\frac{1}{2}\right)^{10-10}$$

$$= \frac{176}{1024}$$

$$= 0.171$$

(ii) $p(r \geq 6) \Rightarrow p(r=6) + p(r \geq 7)$

$$\Rightarrow nc_6 \left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^{10-6} + 0.171$$

$$\Rightarrow 0.376 //.$$

3. The mean and Variance of binomial distribution of 4 and 4/3 respectively. Find $p(x \geq 1)$.

Given mean $\mu = np = 4 \rightarrow ①$

Variance $\sigma^2 = npq = 4/3 \longrightarrow ②$

So $\frac{②}{①} = \frac{npq}{np} = \frac{4/3}{4} = \frac{1}{3}$

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

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

Also $np = 4$

$n\left(\frac{2}{3}\right) = 4$

$n = 6.$

i.e $n=6, p = 2/3, q = 1/3$

$p(x \geq 1) = 1 - p(x=0)$

$= 1 - 6C_0 \left(\frac{2}{3}\right)^2 \left(\frac{1}{3}\right)^6$

$= 1 - (1)\left(\frac{1}{36}\right)$

$= 1 - \frac{1}{129}$

$= 0.998$

4. In 8 throws of a die 5 or 6 is considered a success. Find the mean and standard deviation.

Sol:- Given $n=8$

Let P is the probability of success when 5 or 6 fallen

$P = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$

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

Mean $\mu = np = 8\left(\frac{1}{3}\right)$

$= 2.66$

Variance $\sigma^2 = npq$

$= \left(\frac{8}{3}\right)\left(\frac{2}{3}\right)$

$= \frac{16}{9}$

$= \sqrt{\frac{16}{9}} = \frac{4}{3}$

$= 1.33$

5. In a family of 5 children find the probability that there are 2 boys, atleast 1 boy, All are boys, No boys.

Sol:- Given $n=5, p=1/2, q=1/2$

$$P(X=x) = n_{C_x} p^x q^{n-x}$$
$$= 5_{C_x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{5-x}$$

(i) $P(X=2) = 5_{C_2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{5-2} = 10\left(\frac{1}{2}\right)^5 = 0.3125$

(ii) $P(X\geq 1) = 1 - P(X=0) = 1 - 5_{C_0} \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^5 = 1 - \frac{1}{2^5} = 0.968$

(iii) $P(X=5) = 5_{C_5} \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^0 = \frac{1}{32} = 0.031$

(iv) $P(X=0) = 5_{C_0} \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^5 = \frac{1}{32} = 0.031$

6. Determine the probability of getting a sum of 9 exactly twice in 3 throws with a pair of fair dice

Given $n=3$

Let P be the probability of getting sum of 9 in pair of dice

No.of possibility cases $= 4 \Rightarrow (5,4)(4,5),(6,3),(3,6)$

Total No.of cases $\Rightarrow 6^2 = 36$.

$\therefore P = \frac{4}{36} = \frac{1}{9} \Rightarrow q = \frac{8}{9}$

Probability of getting sum of 9 exactly 2 times in 3 throws $= P(X=2) = 3_{C_2} \left(\frac{1}{9}\right)^2 \left(\frac{8}{9}\right)^1 = 0.033$.

7. 20% of the items produced from a factory are defective Find the probability that in a sample of 5 chosen at random (i) None is defective (ii) 1 is defective (iii) $P(1<x<4)$

$n=5 \quad p = \frac{20}{100} = 0.2 \qquad q = 0.8$

(i) $P(X=0) = 5_{C_0} (0.2)^0 (0.8)^5 = 0.327$

(ii) $P(X=1) = 5_{C_1} (0.2)^1 (0.8)^4 = 5(0.2)(0.8)^4 = 0.409$

(iii) $P(1<x<4) = P(X=2) + P(X=3) = 5_{C_2} (0.2)^2 (0.8)^3 + 5_{C_3} (0.2)^3 (0.8)^2$

$= 10(0.2)^2(0.8)^3 + 10(0.2)^3(0.8)^2 = 0.2048 + 0.0312$

$= 0.256$

8. Fit a Binomial distribution to the following data.

x	0	1	2	3	4	5
f	2	14	20	34	22	8

Sol:- Binomial distribution $\rightarrow N(p+q)^n$

where $N = \Sigma f$

Given; x values, f values (table)

$N = \Sigma fi$

$= 2+14+20+34+22+8$

$= 100$

Mean $= u = np = \Sigma xifi / \Sigma fi$

$5p = 0+14+40+102+88+40 / 100$

$5p = 284/100$

$p = \frac{284}{500}$

$= \frac{2.84}{5}$

$= 0.568$

$q = 1-p = 1-0.568$

$= 0.432$

Binomial distribution $= N(p+q)^n$

$= 100(0.568+0.432)^5$

$= 100[5C_0(0.568)^0(0.432)^5 + 5C_1(0.568)^1(0.432)^4 + 5C_2(0.568)^2(0.432)^3 + 5C_3(0.568)^3(0.432)^2 + 5C_4(0.568)^4(0.432)^1 + 5C_5(0.568)^5(0.432)^0]$

$= 100[0.015+0.099+0.260+0.341+0.225+0.059]$

$= [1.4+9.9+26+34.1+22.5+5.9]$

$= 1+10+26+34+23+6 = 100.$

x	0	1	2	3	4	5
f	2	14	20	34	22	8
B.D	1	10	26	34	23	6

9. 4 Coins are tossed 160 times. The no. of times x heads occur are given below. Find the Binomial distribution.

x	0	1	2	3	4
f	8	34	69	43	6

Given $N = \Sigma f_i$

$= 8 + 34 + 69 + 43 + 6$

$= 160.$

Mean $\mu = np = \dfrac{\Sigma x_i f_i}{\Sigma f_i}$

$4P = \dfrac{0 + 34 + 138 + 129 + 24}{160}$

$4p = \dfrac{325}{160}$

$4p = 2.03125$

$p = 0.5078125$

$= 0.508$

$q = 1 - p = 1 - 0.5078125$

$= 0.4921815 \simeq 0.492$

$B.D = N(p+q)^n$

$= 160[0.508 + 0.492]^4$

$= 160\left[4c_0(0.508)^0(0.492)^4 + 4c_1(0.508)^1(0.492)^3 + 4c_2(0.508)^2 (0.492)^2 + 4c_3(0.508)^3(0.492)^1 + 4c_4(0.508)^4(0.492)^0\right]$

$= 160[0.058 + 0.242 + 0.374 + 0.258 + 0.066]$

$= 9.28 + 38.72 + 59.84 + 41.28 + 10.56.$

$= 9 + 39 + 60 + 41 + 11 = 160$

x	0	1	2	3	4
f	8	34	69	43	6
B.D	9	39	60	41	11

10. A die is thrown 8 times if getting a 2 (or) 4 is a success, find the probability of

(i) 4 is success (ii) $P(x \leq 3)$ (iii) $P(x \geq 2)$ (iv) $P(x \geq 1)$

Sol:- Given $N = 8$

$$p = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

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

(i) $P(x=4) = {}^8C_4 \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^4$

$= 0.1707$

(ii) $P(x \leq 3) = P(x=0) + P(x=1) + P(x=2)$

$= \left[{}^8C_0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^8 + {}^8C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^7 + {}^8C_2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^6\right] + {}^8C_3 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^5$

$= [0.039 + 0.156 + 0.273] + 0.273$

$= 0.468 + 0.273 = 0.741$

(iii) $P(x \geq 2) = 1 - P(x=0) + P(x=1)$

$= 1 - \left[{}^8C_0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^8 + {}^8C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^7\right]$

$= 1 - [0.039 + 0.156]$

$= 1 - 0.195$

$= 0.805$

(iv) $P(x \geq 1) = 1 - P(x=0)$

$= 1 - {}^8C_0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^8$

$= 1 - 0.039$

$= 0.961$

Poisson Distribution:-

This is interoduced by SD poisson in 1837. The poisson distribution can be derived as a limiting case of binomial distribution under following conditions:

(1) The probability of occurance of event is very small

(2) n is very large

Definition of poisson distribution:

$$P(X=x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} & \text{where } x = 0,1,2,3\ldots\ldots \\ 0 & \text{otherwise} \end{cases}$$

Mean of the poisson distribution:

$$\text{Mean } \mu = \Sigma x p(x)$$

$$= \Sigma x \left(\frac{e^{-\lambda}\lambda^x}{x!}\right)$$

$$= e^{-\lambda}\left[\frac{\Sigma x \lambda^x}{x(x-1)!}\right]$$

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

$$= e^{-\lambda}\left[\frac{\lambda}{1} + \frac{\lambda^2}{1!} + \frac{\lambda^3}{2!} + \ldots\right]$$

$$= e^{-\lambda}\lambda\left[1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \ldots\right]$$

$$= e^{-\lambda}\lambda e^{\lambda}\left[e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!}\ldots\right]$$

$$= \lambda$$

Variance of poisson Distribution:-

$$\text{Variance } \sigma^2 = \Sigma x^2 p(x) - \mu^2$$

$$= \Sigma x^2 \frac{e^{-\lambda}\lambda^x}{x!} - \lambda^2$$

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

$$= e^{-1}\left[\Sigma \frac{((x-1)+1)}{(x-1)!}\lambda^x\right] - \lambda^2$$

$$= e^{-\lambda}\left[\sum \frac{(x-1)\lambda^x}{(x-1)!} + \frac{\sum \lambda^x}{(x-1)!}\right] - \lambda^2$$

$$= e^{-\lambda}\left[\sum_{i=2}^{n} \frac{\lambda^x}{(x-2)!} + \sum_{i=1}^{n} \frac{\lambda^x}{(x-1)!}\right] - \lambda^2$$

$$= e^{-\lambda}\left[\left[\frac{\lambda^2}{1} + \frac{\lambda^3}{1!} + \frac{\lambda^4}{2!} + \ldots\right] + \left[\frac{\lambda}{1} + \frac{\lambda^2}{1!} + \frac{\lambda^3}{2!} + \ldots\right]\right] - \lambda^2$$

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

$$= e^{-\lambda}\left[\lambda^2 e^{\lambda} + \lambda e^{\lambda}\right] - \lambda^2$$

$$= \lambda^2 + \lambda - \lambda^2$$

$$= \lambda$$

Variance $= \lambda$

Mode of the poisson Distribution:

Mode of the poisson distribution lies between $(\lambda - 1)$ & λ

Note: If λ is integer then we have 2 modes i.e; $\lambda-1$ & λ. If λ is not integer then mode is integer part of λ.

Recurrance Relation for poisson distribution:

$$P(x+1) = \frac{\lambda}{(x+1)} P(x)$$

→ Average number of accidents on any day on a national highway is 1.8. Determine the probability that the number of accidents are

(i) atleast one (ii) atmost one

Sol:- Given Mean = Average = $\lambda = 1.8$

By poisson distribution;

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

$$= \frac{e^{-1.8}(1.8)^x}{x!}$$

(i) p(atleast one) = $p(x \geq 1)$

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

$$= 1 - \frac{e^{-1.8}(1.8)^0}{0!}$$

$$= 0.8347$$

(ii) p(atmost one) = $p(x \leq 1)$

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

$$= \frac{e^{-1.8}(1.8)^0}{0!} + \frac{e^{-1.8}(1.8)^1}{1!}$$

$$= e^{-1.8}[1+1.8]$$

$$= e^{-1.8}[2.8]$$

$$= 0.4628$$

→ A sample of 3 items is selected at random from a box containing 10 items of which 4 are defective. Find the expected number of defective items.

Sol: The probability of defective items $p = \frac{4}{10} = \frac{2}{5} = 0.4$

Sample of items $n = 3$

$$\text{Mean } \mu = np$$

$$= 3(0.4)$$

$$= 1.2$$

→ If 2 cards are drawn from a pack of 52 cards which are diamonds using poisson distribution, find the poisson probability of getting 2 diamonds atleast 3 times in 51 consecutive trials of 2 cards.

Sol:- The probability of getting 2 diamonds from a pack of 52 cards $\Rightarrow p = \frac{13C_2}{52C_2} = \frac{1}{17}$

$n = 51$

Mean $\lambda = np$

$= 51\left(\frac{1}{17}\right)$

$= 3$

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

$= \frac{e^{-3}\,3^{x}}{x!}$

$p(\text{atleast } 3 \text{ times}) = p(X \geq 3)$

$= 1 - [p(X=0) + p(X=1) + p(X=2)]$

$= 1 - \left[\frac{e^{-3}(3)^{0}}{0} + \frac{e^{-3}3^{(1)}}{1!} + \frac{e^{-3}(3)^{2}}{2!}\right]$

$= 1 - e^{-3}\left[1 + 3 + \frac{9}{2}\right]$

$= 1 - e^{-3}[\ ] = 0.5678$

→ A hospital switch board receives an average of 4 emergency calls in a 10min interval. What is the probability that

(i) there are atmost 2 emergency calls in 10min interval

(ii) there are exactly 3 emergency calls in 10min interval

Sol:- Given Average $\lambda = 4$

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

$p(X=x) = \frac{e^{-4}.4^{x}}{x!}$

(i) $p(\text{atmost } 2) = p(X \leq 2)$

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

$= \frac{e^{-4}4^{0}}{0!} + \frac{e^{-4}4^{1}}{1!} + \frac{e^{-4}4^{2}}{2!} = e^{-4}[1+4+8]$

$= e^{-4}[13] = 0.238$

(ii) p(exactly 3) = $p(x=3)$

$$= \frac{e^{-4}4^3}{3!} = \frac{e^{-4}4^3}{6}$$

$$= 0.1954$$

→ If $p(1) = p(2)$ then find ① mean ② $p(4)$ ③ $p(x \geq 1)$ ④ $p(1 < x < 4)$

Sol:- Given $p(1) = p(2)$

$$\frac{e^{-\lambda}\lambda^1}{1!} = \frac{e^{-\lambda}\lambda^2}{2!} \Rightarrow 2\lambda = \lambda^2 \Rightarrow \lambda^2 - 2\lambda = 0$$

$$\lambda = 0, \lambda = 2$$

② $p(4) = \frac{e^{-\lambda}\lambda^4}{4!} = \frac{e^{-2}(2)^4}{4!} = 0.09$

③ $p(x \geq 1) = 1 - p(x=0) = 1 - \frac{e^{-2}(2)^0}{0!} = 0.864$

④ $p(1 < x < 4) = p(x=2) + p(x=3) = \frac{e^{-2}(2)^2}{2!} + \frac{e^{-2}(2)^3}{3!}$

$$= e^{-2}\left[\frac{4}{2} + \frac{8}{6}\right] = e^{-2}\left[3.33\right] = 0.45$$

→ If a poisson distribution is such that $p(x=1)\frac{3}{2} = p(x=3)$ find $p(x \geq 1)$, $p(x \leq 3)$, $p(2 \leq x \leq 5)$

Sol:- Given $p(x=1)\frac{3}{2} = p(x=3)$

$$\frac{e^{-\lambda}(\lambda)^1}{1!}\left(\frac{3}{2}\right) = \frac{e^{-\lambda}(\lambda)^3}{3!}$$

$$\frac{3\lambda}{2} = \frac{\lambda^3}{6}$$

$$9\lambda = \lambda^3$$

$$\lambda^3 - 9\lambda = 0$$

$$\lambda(\lambda^2 - 9) = 0$$

$$\lambda = 0, \lambda = \pm 3$$

$$\lambda \neq 0 \text{ (or) } -3$$

$$\boxed{\lambda = 3}$$

(i) $p(x \geq 1) = 1 - p(x=0)$

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

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

$= 1 - e^{-3}$

$= 0.9502$

(ii) $p(x \leq 3) = p(x=0) + p(x=1) + p(x=2) + p(x=3)$

$= \frac{e^{-\lambda}(\lambda)^0}{0!} + \frac{e^{-\lambda}(\lambda)^1}{1!} + \frac{e^{-\lambda}(\lambda)^2}{2!} + \frac{e^{-\lambda}(\lambda)^3}{3!}$

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

$= e^{-3}\left[1 + 3 + \frac{9}{2} + \frac{\cancel{27}^{9}}{\cancel{6}_{2}}\right]$

$= e^{-3}\left[1 + 3 + \frac{18}{2}\right]$

$= e^{-3}\left[4 + 9\right]$

$= e^{-3}\left[13\right]$

$= 0.647$

(iii) $p(2 \leq x \leq 5)$

$= p(x=2) + p(x=3) + p(x=4) + p(x=5)$

$= \frac{e^{-\lambda}(\lambda^2)}{2!} + \frac{e^{-\lambda}(\lambda^3)}{3!} + \frac{e^{-\lambda}(\lambda)^4}{4!} + \frac{e^{-\lambda}(\lambda)^5}{5!}$

$= e^{-\lambda}\left[(\lambda)^2\right]\left[\frac{1}{2} + \frac{\lambda}{3} + \frac{\lambda^2}{24} + \frac{\lambda^3}{120}\right]$

$= e^{-3}(9)\left[\frac{1}{2} + 1 + \frac{9}{24} + \frac{27}{120}\right]$

$= e^{-3}(9)(2.1)$

$= 0.940$ //

→ If $2p(x=0) = p(x=2)$ find (i) $p(x \le 3)$ (ii) $p(2 < x \le 5)$ (iii) $p(x \ge 3)$

Given: $2p(x=0) = p(x=2)$

$$2\frac{e^{-\lambda}(\lambda)^0}{0!} = \frac{e^{-\lambda}(\lambda)^2}{2!}$$

$$2 = \frac{\lambda^2}{2}$$

$$4 = \lambda^2$$

$$\lambda = 2.$$

(i) $p(x \le 3) = p(x=0) + p(x=1) + p(x=2) + p(x=3)$

$$= \frac{e^{-\lambda}(\lambda)^0}{0!} + \frac{e^{-\lambda}(\lambda)^1}{1!} + \frac{e^{-\lambda}(\lambda)^2}{2!} + \frac{e^{-\lambda}(\lambda)^3}{3!}$$

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

$$= e^{-2}\left[1 + 2 + 2 + \frac{4}{3}\right] = \cancel{e^{-2}[6]}$$

$$= e^{-2}\left[\frac{19}{3}\right] = 0.857.$$

(ii) $p(2 < x \le 5) = p(x=3) + p(x=4) + p(x=5)$

$$= \frac{e^{-\lambda}(\lambda^3)}{3!} + \frac{e^{-\lambda}(\lambda)^4}{4!} + \frac{e^{-\lambda}(\lambda)^5}{5!}$$

$$= e^{-\lambda}\left[\frac{8}{6} + \frac{16}{24} + \frac{32}{120}\right] = e^{-2}(8)\left[\frac{1}{6} + \frac{2}{24} + \frac{4}{120}\right]$$

$$= \frac{e^{-2}(8)}{6}\left[1 + \frac{1}{2} + \frac{1}{5}\right]$$

$$= e^{-2}\left(\frac{4}{3}\right)\left[\frac{17}{10}\right] = e^{-2}(2.267) = 0.3068.$$

(iii) $p(x \ge 3) = 1 - p(x \le 3)$

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

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

$$= 1 - \left[e^{-2}\left[1 + \frac{\lambda}{1} + \frac{\lambda^2}{2}\right]\right]$$

$$= 1 - \left[e^{-2}[1+2+2]\right] = 1 - e^{-2}[5] = 1 - 5e^{-2}$$

$$= 0.3233$$

→ Fit a poisson distribution to the following table

x	0	1	2	3	4	5
f	142	156	69	27	5	1

Sol:- $N = \Sigma f = 400$

$$\text{Mean} = \lambda = \frac{\Sigma f_i x_i}{\Sigma f_i}$$

$$= \frac{0+156+138+81+20+5}{400}$$

$$= \frac{400}{400} = 1$$

poisson Distribution $= N p(x)$ · where $x = 0,1,2,3,4,$

$x=0$, $N p(0) = 400 \cdot \frac{e^{-1}(1)^0}{0!} = \frac{400}{e} = 147$

$x=1$, $N p(1) = 400 \cdot \frac{e^{-1}(1)^1}{1!} = \frac{400}{e} = 147$

$x=2$, $N p(2) = 400 \; \frac{e^{-1}(1)^2}{2!} = \frac{400}{2e} = 74$

$x=3$, $N p(3) = 400 \; \frac{e^{-1}(1)^3}{3!} = \frac{400}{3e} = 25$

$x=4$, $N p(4) = 400 \; \frac{e^{-1}(1)^4}{4!} = \frac{400}{24e} = 6$

$x=5$, $N p(5) = 400 \cdot \frac{e^{-1}(1)^5}{5!} = \frac{400}{120e} = 1.$

x	0	1	2	3	4	5
f	142	156	69	27	5	1
P.D	147	147	74	25	6	1

Normal Distribution:

Normal distribution was first discovered by De Moivre in 1733 and further developed by Laplace and Gauss. This is known as Gaussian distribution. It is limiting form of Binomial distribution of large values of n when p and q are not very small.

Definition of normal distribution:

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

$$-\infty < x < \infty, \; -\infty < \mu < \infty, \; \sigma > 0.$$

Note: In Normal distribution mean = median = mode.

Characteristics of normal distribution:

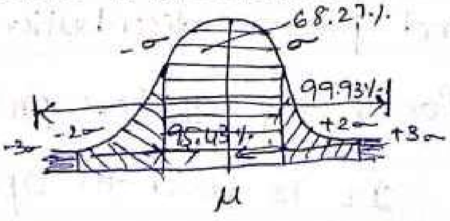
1) The shape of the graph of the normal distribution is bell shaped

2) Area under the normal curve represents the total population.

3) Mean, median and mode are coincide at middle of the curve.

4) The x-axis never touches the curve.

5) Area under the normal curve is distributed as follows:



6) Area of normal curve between $\mu - \sigma$ and $\mu + \sigma$ is 68.27%.

7) Area of normal curve between $\mu - 2\sigma$ and $\mu + 2\sigma$ is 95.43%.

8) Area of normal curve between $\mu - 3\sigma$ and $\mu + 3\sigma$ is 99.95%.

Standard Normal [illegible] distribution with mean $\mu=0$ and

The normal distribution with mean $\mu=0$ and standard deviation $\sigma=1$ then it is called standard normal distribution.

Uses of Normal distribution:

→ The normal distribution can be used to approximate Binomial & poisson distribution.

→ It helps us to estimate parameter from statistic and to find confidence limits to the parameter.

→ It is widely use in to test the hypothesis and test the significance of the population.

Importance & Applications of Normal distribution:

Normal distribution plays a very important role in statistical theory because of the following reasons

→ Data obtained from psychological, physical & Bilogical measurements approximates follows normal distributions.

Eg:- Height & weight of individuals, IPL score's

Normal distribution is limiting case of Binomial and poisson distribution. It is used to approximate for many applied problems in different branches.

It is used to approximate any statistic value. It is used in SQC → Statiscal Quality control in Industr for finding control limits.

Find the probability Density of Normal Curve:

The probability that the normal variant x with mean & signa lies between ...

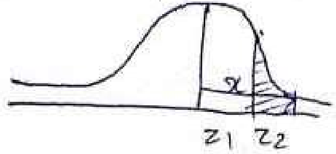
can be obtained using area under the standard normal curve as follows:

→ To find $z = \frac{x-\mu}{\sigma}$

→ $p(x_1 \le x \le x_2)$

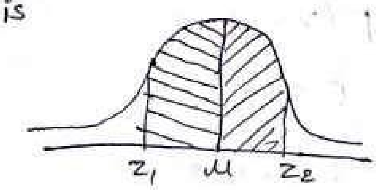
$$z_1 = \frac{x_1-\mu}{\sigma}, \quad z_2 = \frac{x_2-\mu}{\sigma}$$

→ z_1 and z_2 are positive then the normal curve is



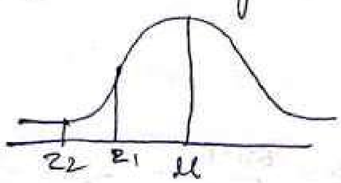
$$p(x_1 \le x \le x_2) = A(x_1 \le x \le x_2)$$
$$= A(z_2) - A(z_1)$$

→ If z_1 and z_2 are opposite signs then the normal curve is



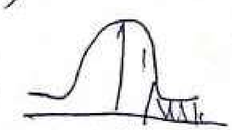
$$p(x_1 \le x \le x_2) = A(z_1 \le z \le z_2)$$
$$= A(z_2) + A(z_1)$$

→ If z_1 and z_2 are negative then normal curve

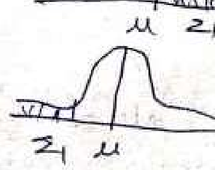


$$p(x_1 \le x \le x_2) = p(z_1 \le z \le z_2)$$
$$= A(z_2) - A(z_1)$$

→ If $z_1 > 0$; $p(z > z_1) = 0.5 - A(z_1)$

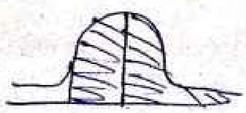


If $z_1 < 0$; $p(z < z_1) = 0.5 + A(z_1)$



If $z < 0$ & $p(z > z_1)$ ⇒

$$p(z > z_1) = 0.5 + A(z_1)$$



Problems:

iii

→ If the weights of 300 students are normally distributed with mean 68kgs and standard deviation 3 kgs. How many students have weights

(i) > 72kgs (ii) ≤ 64kgs (iii) Between 65 and 71kgs.

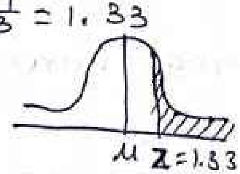
Sol:- Given $\mu = 68$kgs

S.D $\sigma = 3$kgs

(i) $p(x > 72) =$

When $x = 72 \Rightarrow z = \frac{x-\mu}{\sigma}$

$z = \frac{72-68}{3} = \frac{4}{3} = 1.33$



$\therefore P(x > 72) = p(z > 1.33)$

$= 0.5 - A(1.33)$

$= 0.5 - 0.4082$

$= 0.0918$

(ii) ~~p(x~~ No.of students more than 72 kgs $= 0.0918 \times 300$

$= 28$.

(ii) $p(x \leq 64) =$

$x = 64$, $z = \frac{64-68}{3} = \frac{-4}{3} = -1.33$

$p(x \leq 64) = p(z \leq -1.33)$

$= 0.5 - A(1.33)$

$= 0.5 - 0.4082$

$= 0.0918$

No.of students less on or equal to 64kgs $= 0.0918 \times 300$

$= 28$ students.

(iii) $p(65 \le x \le 71)$

$x_1 = 65 \Rightarrow z_1 = \frac{65-68}{3}$. $x_2 \Rightarrow 71 \Rightarrow z_2 = \frac{71-68}{3}$

$= -1$ $\qquad = 1$

$p(65 \le x \le 71) = p(-1 \le z \le 1)$

$= A(1) + A(1)$

$= 2A(1)$

$= 2 \times 0.3413$

$= 0.6826$

No.of students between 65 and 71kgs $= 0.6826 \times 300$

$= 204.78$

$= 205$ students

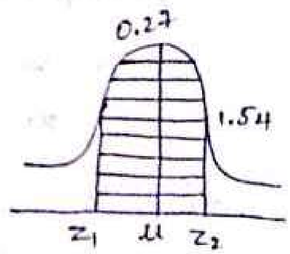
→ The mean deviation of marks obtained by 1000 students in an examination are respectively 34.5 & 16.5. Assuming the normality distribution find approximately no.of students expected to obtain marks between 30 & 60.

Sol:- Given $\mu = 34.5$, $\sigma = 16.5$, $n = 1000$

$P(30 \le x \le 60) = z = \frac{x-\mu}{\sigma}$

where $x_1 = 30$, $z_1 = \frac{30-34.5}{16.5} = -0.27$

$x_2 = 60$, $z_2 = \frac{60-34.5}{16.5} = 1.54$



$p(30 \le x \le 60) = p(-0.27 \le z \le 1.54)$

$= A(0.27) + A(1.54)$

$= 0.1064 + 0.4382 = 0.5446.$

∴ No.of students who get marks between 30 & 60

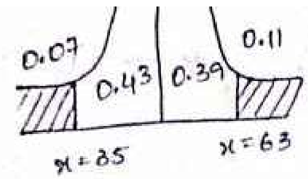
$= 0.5446 \times 1000$

$= 544.6 = 545$

→ In a Normal distribution 7% of items are under 35 & 89% are under 63 determine the mean & Variance of distribution

Sol:- Given, 7% of items are under 35 & 89% are under 63

These are shown in fig



$\therefore P(x<35)=0.09,\ P(x\le 63)=0.89,\ P(x>63)=0.11$

where $x_1=35,\quad z_1=\frac{x-\mu}{\sigma}=\frac{35-\mu}{\sigma}=-1.48 \longrightarrow ①$

$x_2=63,\quad z_2=\frac{x-\mu}{\sigma}=\frac{63-\mu}{\sigma}=1.23 \longrightarrow ②$

from ① ; $\mu-1.48\sigma=35$

from ② $\quad \mu+1.23\sigma=63$

$-2.71\sigma=-2.8 \Rightarrow \sigma=10.33$

Variance $\sigma^2=(10.33)^2=106.70$

$\mu-(1.48)(10.33)=35$

$\mu=50.29$

→ In a Normal distribution 31% of items under 45 and 8% are over 64 find the Mean & Variance

Sol:- $P(x<45)=\cancel{0.3\%}\ 0.31$, $P(x>64)=0.08$

where $x_1=45 \Rightarrow z_1=\frac{x-\mu}{\sigma}=\frac{45-\mu}{\sigma}=-0.50$

$x_2=64 \Rightarrow z_2=\frac{x_2-\mu}{\sigma}=\frac{64-\mu}{\sigma}=1.41$

from ① $\quad \mu-0.50\sigma=45 \rightarrow ①$ $\qquad \mu+1.41\sigma=64 \rightarrow ②$

$\mu+1.41\sigma=64$

$-1.9\sigma=-19 \Rightarrow \sigma=19.9$

$\sigma^2=99.02$

→ In a sample of 1000 cases the mean of certain test is 14 & $\sigma=2.5$. Assuming the distribution is normal
(i) find how many students score between 12 & 15
(ii) How many score above 18 (iii) How many score below 18.

Sol:- Given $\mu=14,\ \sigma=2.5,\ n=1000$

$P(12\le x\le 15)$, when $x_1=12,\ z_1=\frac{x_1-\mu}{\sigma}=\frac{12-14}{\sigma}=-0.8$

when $x_2=15,\ z_2=\frac{x_2-\mu}{\sigma}=\frac{15-14}{2.5}=0.4$

$=P(-0.8)\le x\le 0.4)$

$= A(0.8) + A(0.4) = 0.2881 + 0.1554$

No.of students between 12 & 15 $= 0.4435 \times 1000$

$= 444.$

(ii) $p(x>18) = p(z<1.6)$

$= 0.5 + A(1.6) = 0.5 + 0.4452$

$= 0.94452 \times 1000 = 945$

(iii) $p(x>18) = 0.5 - A(1.6) = 0.0548 \times 1000$

$= 55$

→ The marks obtained in Mathematics by 1000 students is normal distribution with $\mu = 78\%$, $\sigma = 11\%$.

(i) Determine how many students got marks above 90%

(ii) What was the highest mark obtained by the lowest 10% of students.

(iii) Within what limits did the middle of 90% of students lie.

→ If x is normally distributed with Mean 2 & Variance 0.1 then find $P(|x-2| \geq 0.01)$

Sol:- Given $\mu = 2$; $\sigma^2 = 0.1 \Rightarrow \sigma = 0.316$.

$P(|x-2| < 0.01) = P(1.99 < x < 2.01)$

when $x_1 = 1.99$ then $z_1 = \frac{1.99-2}{0.316} = -0.03$

when $x_2 = 2.01$ then $z_2 = \frac{2.01-2}{0.316} = 0.03$

$P(|x-2| < 0.01) = P(-0.03 < z < 0.03)$

$= A(0.03) + A(0.03)$

$= 2A(0.03) = 2(0.0120) = 0.024.$

$P(|x-2| \geq 0.01) = 1 - P(|x-2| < 0.01)$

$= 1 - 0.024 = 0.976.$

→ If x is a normal variate with Mean 30 & $\sigma = 5$ find probabilities that (i) $x \geq 45$ (ii) $26 \leq x \leq 40$ (iii) $x \leq 25$

Sol:- Given $\mu = 30$, $\sigma = 5$.

(i) $P(x \geq 45)$; when $x = 45$, $z = \frac{45-30}{5} = 3$

$p(x \geq 45) = p(z \geq 3) = 0.5 - A(3) = 0.5 - 0.498 + = 0.0013$

ii) $p(26 \leq x \leq 40)$; when $x_1 = 26$, $z_1 = \frac{26-30}{5} = -0.8$

$$x_2 = 40, \; z_2 = \frac{40-30}{5} = 2$$

$$p(26 \leq x \leq 40) = p(-0.8 \leq z \leq 2)$$

$$= A(-0.8) + A(2) = 0.2881 + 0.4772$$

$$= 0.7653.$$

(iii) $p(x \leq 25) \Rightarrow z = \frac{25-30}{5} = -1$

$$p(x \leq 25) + p(z \leq -1)$$

$$= 0.5 + A(1) = 0.5 + 0.3413$$

$$= 0.8413.$$

Approximation for the binomial distribution:-

→ 8 coins are tossed together. Find the probability of getting to 4 heads in single toss, using Normal approximation

Sol:- Given the coin $p = 1/2$, $q = 1/2$, $n = 8$

$$\mu = np = 8\left(\tfrac{1}{2}\right) = 4 \qquad \sigma = npq = 8\left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right) = \sqrt{2}$$

$P(1 \leq x \leq 4) =$

when $x_1 = 1 \Rightarrow \frac{1^{1/2} - 4}{\sqrt{2}} = \frac{-7}{2\sqrt{2}} = -2.47 \quad \because z_1 = \frac{(x_1 - \frac{1}{2}) - 4}{\sigma}$

when $x_2 = 4 \Rightarrow \frac{4^{1/2} - 4}{\sqrt{2}} = \frac{1}{2\sqrt{2}} = 0.35 \qquad z_2 = \frac{(x_2 + \frac{1}{2}) - 4}{\sigma}$

$$P(1 \leq x \leq 4) = \int_{-2.47}^{0.35} \phi(z)\,dz$$

$$= P(-2.47 \leq z \leq 0.35)$$

$$= A(-2.47) + A(0.35) = 0.4932 + 0.1368 = 0.63$$

→ Find the probability of getting even number on face 3 to 5 times in throwing 10 dice together.

Sol:- Given the dice;

Let p = Probability of getting even number indies $= \left(\frac{1}{6} + \frac{1}{6} + \frac{1}{6}\right) = 1/2$

$q = 1/2 \quad n = 10 \quad \mu = np = 10\left(\tfrac{1}{2}\right) = 5 \quad \sigma = \sqrt{npq} = \sqrt{10\left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right)}$

$p(3 \le x \le 5)$

when $x_1 = 3$; $z_1 = \frac{(3-1/2)-5}{1.58} = -1.58$

when $x_2 = 5$; $z_2 = \frac{(5+1/2)-5}{1.58} = 0.32$.

$$p(3 \le x \le 5) = \int_{-1.58}^{0.32} \phi(z)\,dz = (p(-1.58) \le z \le 0.32)$$

$$= A(1.58) + A(0.32) = 0.4429 + 0.1255$$

$$p(3 \le x \le 5) = 0.5684$$

→ Find the probability that by guess work a student can correctly answer 25 to 30 questions in a multiple choice quiz consisting of 80 questions. Assume that in each question with 4 choice, only one choice is correct and student has no knowledge on subject.

Sol:- Given $p = 1/4$; $q = \frac{3}{4}$, $n = 80$

Mean $\mu = np = 80\left(\frac{1}{4}\right) = 20$

$\sigma = \sqrt{npq} = \sqrt{80\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)} = \sqrt{15} = 3.872$

$p(25 \le x \le 30)$

when $x_1 = 25$, $z = \frac{(25-\frac{1}{2})-20}{3.872} = 1.16$

when $x_2 = 30$; $z = \frac{(30+\frac{1}{2})-20}{3.872} = 2.71$

$$p(25 \le x \le 30) = \int_{1.16}^{2.71} \phi(z)\,dz = p(1.16) \le z \le 2.71)$$

$$= A(2.71) - A(1.16)$$

$$= 0.4966 - 0.3770$$

$$p(25 \le x \le 30) = 0.1196$$

Uniform distribution:-

In the Uniform distribution every point has same probability.

Eg:- Taking a coin → probability $\frac{1}{2}$

Taking a die → probability $\frac{1}{6}$

Taking 100 numbers → probability of getting any no → $\frac{1}{100}$

Uniform distribution is of 2 types. They are:

1) Discrete Uniform distribution

2) Continuous Uniform distribution

1) Discrete Uniform distribution:-

In this case the discrete random variable each of its values with the same probability. Suppose the sample space S contains points with same probability $\frac{1}{m}$. This is shown below

$$f(x) = \frac{1}{m} \quad \text{for} \quad x = x_1, x_2 \ldots\ldots x_n$$

x	x_1	x_2	....	x_m
$f(x)$	$\frac{1}{m}$	$\frac{1}{m}$	....	$\frac{1}{m}$

The mean and Variance of discrete Uniform distribution are given by $\mu = E(x) = \sum_{i=1}^{n} x_i f(x_i)$

$$= \frac{1}{m}\sum_{i=1}^{n} x_i$$

$$\text{Variance } \sigma^2 = \frac{1}{m}\sum_{i=1}^{n}(x_i-\mu)^2$$

Problems:-

→ Find the mean and Variance of the following table

x	1	2	3	4	5	6
$f(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Sol:- Given; $x = 1,2,3,4,5,6$ & $f = \frac{1}{6}$

$$\text{Mean } \mu = \frac{1}{m}\sum_{i=1}^{n} x_i$$

$$= \frac{1}{6}[1+2+3+4+5+6]$$

$$= \frac{1}{6}[21] = \frac{1}{2}[7)$$

$$\mu = \frac{7}{2} = 3.5$$

$$\therefore \sigma^2 = \frac{1}{m}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$= \frac{1}{6}\left[(1-3.5)^2+(2-3.5)^2+(3-3.5)^2+(4-3.5)^2+(5-3.5)^2+(6-3.5)^2\right]$$

$\frac{1}{6}\left[6.25+2.25+0.25+0.25+2.25+6.25\right]$

$= \frac{1}{6}\left[17.5\right]$.

$= 2.916$ //

→ Find the mean and Variance of following Table

x	2	4	6	8
f	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

Sol:- Given; $x = 2, 4, 6, 8$ & $f = \frac{1}{4}$

$$\text{Mean} ; \mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$= \frac{1}{4}\left[2+4+6+8\right]$$

$$= \frac{20}{4} = 5$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$= \frac{1}{4}\left[(2-5)^2+(4-5)^2+(6-5)^2+(8-5)^2\right]$$

$$= \frac{1}{4}\left[9+1+1+9\right]$$

$$= \frac{20}{4} = 5.$$

Continuos Uniform distribution:-

If x is a continuos random variable then the uniform distribution is

$$f(t) = \begin{cases} \frac{1}{T} & \text{for } 0 < t < T \\ & \text{otherwise } 0 \end{cases}$$

$$\text{Mean } \mu = \int_0^T t\, f(t)\, dt$$

$$\text{Variance } \sigma^2 = \int_0^T t^2 f(t)\, dt - \mu^2$$

Problems:

→ Find the mean and variance of $f(t) = \begin{cases} \frac{1}{6} & 0<t<6 \\ \text{Otherwise } 0. \end{cases}$

Given $f(t) = \begin{cases} \frac{1}{6} & 0<t<6 \\ \text{otherwise } 0 \end{cases}$

$$\text{Mean } \mu = \int_0^6 t\, f(t)\,dt$$

$$= \int_0^6 t\left(\frac{1}{6}\right)dt = \frac{1}{6}\int_0^6 t\,dt$$

$$= \frac{1}{6}\left(\frac{t^2}{2}\right)_0^6 = \frac{1}{6}\left[\frac{36}{2}\right] = 3$$

$$\text{Variance } \sigma^2 = \int_0^6 t^2\left(\frac{1}{6}\right)dt - \mu^2$$

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

$$= \frac{1}{6}\left[\frac{36\times 6}{3}\right] - 3^2$$

$$= 12 - 9 = 3.$$

→ A random Variable x has a uniform PDF period with T=10, find the probability 1) $1 \le x \le 3$ 2) $1 \le x \le 9.3$ 3) $x > 2.9$ 4) $x < 7.2$ 5) $-1 < x < 2$ 6) $9.1 < x < 12.3$

Sol:- (i) $p(1 \le x \le 3) = \int_1^3 f(t)\,dt$

$$= \int_1^3 \frac{1}{10}\,dt = \frac{1}{10}\left[t\right]_1^3 = \frac{3}{10} - \frac{1}{10} = \frac{2}{10}$$

$$= \frac{1}{5} = 0.2$$

0.83
0.71
0.72
0.2
0.3

(ii) $p(1 \le x \le 9.3) = \int_1^{9.3} f(t)\,dt.$

$$= \int_1^{9.3} \frac{1}{10}\,dt = \frac{1}{10}\left[t\right]_1^{9.3} = \frac{9.3}{10} - \frac{1}{10}$$

$$= 0.83$$

i) $p(x > 2.9) = \int_{2.9}^{10} f(t)\,dt$

$= \frac{1}{10}[t]_{2.9}^{10} = \frac{1}{10}[10-2.9] = [\frac{1}{10}][7.1]$

$= 0.71$

ii) $p(x < 7.2) = \int_{0}^{7.2} f(t)\,dt$

$= \frac{1}{10}[t]_{0}^{7.2} = \frac{1}{10}[7.2-0] = \frac{7.2}{10}$

$= 0.72$

iii) $p(-1 < x < 2) = \int_{0}^{2} f(t)\,dt$

$= \frac{1}{10}[t]_{0}^{2} = \frac{1}{10}[2]$

$= 0.2$

iv) $p(9.1 < x < 12.3) = \int_{9.1}^{10} f(t)\,dt$

$= \frac{1}{10}[t]_{9.1}^{10} = \frac{1}{10}[10-9.1]$

$= 0.09$

Exponential Distribution:-

Let t be the time between events happening then the exponential PDF $F(t)$ is given by

$$F(t) = \begin{cases} \alpha e^{-\alpha t}, & \text{for } t \geq 0 \\ 0 & ; \text{ Otherwise} \end{cases}$$

→ Mean of expected value

$$\mu = \int_0^\infty \alpha t e^{-\alpha t}\,dt$$

$$= \alpha \int_0^\infty t e^{-\alpha t}\,dt = \alpha \left[t\left(\frac{e^{-\alpha t}}{-\alpha}\right) - (1)\left(\frac{e^{-\alpha t}}{(-\alpha)^2}\right) \right]_0^\infty$$

$$= \alpha \left[(0-0) - \left(0 - \frac{1}{\alpha^2}\right) \right]$$

$$= \frac{\alpha}{\alpha^2} = 1/\alpha$$

$$\therefore \mu = 1/\alpha$$

Problems:-

1) The time between breakdown of a machine follows an exponential distribution with a mean of 17 days. Find the probability that a machine breaksdown in a 15-day period.

SOl:- Given the mean time between breakdowns

$$=17(\alpha)$$

$$\therefore \mu = 1/\alpha = \frac{1}{17}$$

$$\alpha = 17$$

The probability density function

$$F(t) \text{ is } f(t) = \alpha e^{-t/\alpha} = \frac{1}{17} e^{-t/17}$$

$$P(0 \le t \le 15) = \int_0^{15} f(t)dt$$

$$= \int_0^{15} \frac{1}{17} e^{-t/17} dt$$

$$= \frac{1}{17}\int_0^{15} e^{-t/17} dt = \frac{1}{17}\left[\frac{e^{-t/17}}{(-1/17)}\right]_0^{15}$$

$$= \frac{1}{17} \times \frac{17}{-1}\left[e^{-t/17}\right]_0^{15} = -1\left[e^{-15/17} - 1\right]$$

$$= 1 - e^{-15/17} = 0.586$$

2) The mean time between breakdowns for a machine in 400hrs. Find the probability that the time between the breakdown for a machine is (i) >450hrs (ii) <350hrs.

Given the mean time between breakdowns = 400

$$\mu = 1/\alpha = \frac{1}{400}$$

$$\alpha = 400$$

$$\therefore F(t) = \alpha e^{-t/\alpha} = 400 e^{-t/400}$$

(i) $P(t > 450)$

$$= \int_{450}^{\infty} e^{-t/\alpha} \alpha \, dt = 400 \int_{450}^{\infty} e^{-t/400}$$

$$= 400\left[\frac{e^{-t/400}}{-1/400}\right]_{450}^{\infty}$$

$$= -1\left[0 - e^{-450/400}\right] = e^{-9/8} = 0.3247$$

(ii) $p(t < 350)$

$$= 400\int_0^{350} e^{-t/400}\,dt = 400\left[\frac{e^{-t/400}}{-1/400}\right]_0^{350}$$

$$= -1\left[e^{-350/400} - 1\right]$$

$$= 1 - e^{-350/400} = 0.5831$$

(1)

Stochastic Processes and Markov Chains

Introduction :-

Stochastic is a Greek word which means "Random" (or) "Chance"

Stochastic Analysis deals with models which involve uncertainties or randomness.

A random Variable is a rule (or function) that assigns a real number to every outcome of a random experiment, while a random Process is a rule (or a function) that assigns a time function to every outcome of a random experiment.

Stochastic (Random Process) :-

Def :- A stochastic (or random) Process is defined as a Collection of random Variables

$$\{ X(t_n) \; ; \; n = 1, 2, 3 \cdots \}$$

The random Variable $X(t)$ stands for Observation at time 't'

The number of states 'n' may be finite (or) Infinite depending upon the time range.

For example the Poisson distribution

$$P_n(t) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, \quad n = 1, 2, 3 \ldots$$

represents a stochastic (or random) Process with infinite number of States.

Here the random Variable 'n' denotes the number of occurrences between the time interval 0 and t (assuming that the System starts at 0 times).

Thus the states of the system at any time t are given by $n = 0, 1, 2 \ldots$

Markov Process Definition :-

Stochastic (or random) System is Called a Markov Process If the Occurrence of a future state depends on the immediately Preceding state and only on it.

Thus If $t_0 < t_1 < \ldots < t_n$ represents the Points in time scale then the family of random Variables $\{X(t_n)\}$ is Said

2

to be a Markov Process Provided it holds the Markovian Property.

$$P\left[X(t_n) = x_n / X(t_{n-1}) = x_{n-1}, \dots X(t_0) = x_0\right]$$

$$= P\left[X(t_n) = x_n / X(t_{n-1}) = x_{n-1}\right]$$

Markov Process is a Sequence of 'n' experiments in which each experiment has 'n' possible outcomes $x_1 x_2 --- x_n$.

Each individual outcome is called a state and the Probability (that a Particular outcome occurs) depends only on the Probability of the outcome of the Preceeding experiment.

Characteristics of Markov Process :-

Markov analysis is based on the following Characteristics :-

(1) The states are both collectively exhaustive and mutually exclusive.

(2) The Problem must have a finite number of states, none of them "absorbing" in nature.

(3) The transition Probabilities are Stationary

(4) The Probability of moving from One state to another depends only on the immediately Preceding state.

(5) The transition Probabilities of moving two alternatives states in the next time Period, given a state in the Current time Period must sum to Unity.

(6) The Process has a set of initial Probabilities that may be either given or determined.

Transition Probability :-

Def :- The Probability of moving from One state to another or remaining in the Same state during a single time Period is Called **Transition Probability**

Mathematically the Probability

$$P_{x_{n-1}, x_n} = P\left[X(t_n) = x_n \,/\, X(t_{n-1}) = x_{n-1}\right]$$

is Called the **Transition Probability**.

③

This Conditional Probability is known as One step transition Probability, because it describes the system during the time interval (t_{n-1}, t_n).

Since each time a new result occurs, the Process is said to have stepped or incremented One step. Here 'n' indicates the number of steps or increments.

If $n=0$; it represents the initial state.

Transition Probability Matrix :-

The transition Probabilities can be arranged in a matrix form and such a matrix is called a One step transition Probability matrix denoted by

$$P = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22} & \cdots & P_{2m} \\ & \vdots & & \\ P_{m1} & P_{m2} & \cdots & P_{mm} \end{bmatrix}$$

The matrix P is a Square matrix whose each element is non-negative and sum of elements of each row is Unity.

In general, any matrix P, whose elements are non-negative and sum of elements either in each row or column is Unity is called a transition matrix or a Probability matrix.

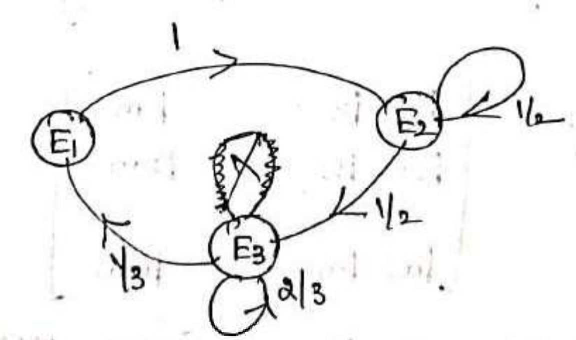
Thus a transition matrix is a square stochastic matrix and it gives the complete description of the Markov Process.

Diagrammatic representation of transition Probabilities

Transition diagram :-

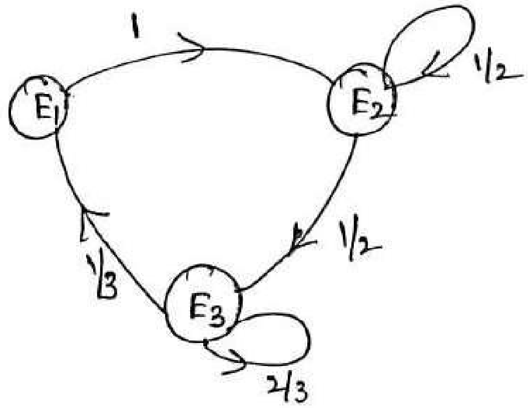
It shows the transition Probabilities (or) shifts that can occur in any Particular solution.

Ex :-



4

Ex :-



The arrows from each state indicate the possible states to which a process can move from the given state.

The matrix of transition probabilities corresponding to the above diagram is

$$P = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array} \begin{array}{c} \begin{array}{ccc} E_1 & E_2 & E_3 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \end{bmatrix} \end{array}$$

Transition Probability matrix

A zero element indicates that the transition is not possible.

5

First and Higher Order Markov Processes

The First Order markov Process is based on the following three assumptions

(1) The set of Possible outComes is finite

(2) The Probability of the next outCome depends only on the immediately Preceding outCome.

(3) The transition Probabilities are Constant over time.

depends on Preceding state.

x_{n-1} $\quad x_n$ $\quad x_{n+1}$

The Second Order markov Process assumes that the Probability of the next outCome state may depend on the two Previous outComes.

Likewise a Third Order markov Process assumes that the Probability of the next outCome state Can be Calculated by Obtaining and taking account of the outComes of the past three outComes.

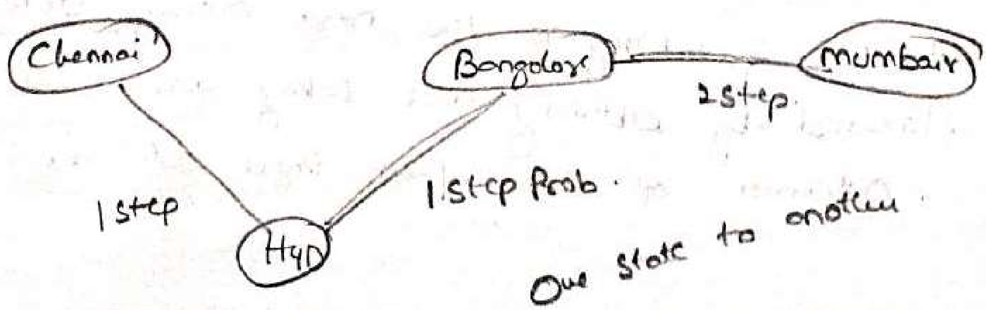
n - step transition Probabilities :-

Suppose the system which occupies the state E_1 at time $t=0$, then we may be interested in finding out the Probability that the system moves to the state E_j at time $t=n$ (these time Periods are referred to as number of steps)

If the n - step transition Probability is denoted by $P_{ij}^{(n)}$, then these Probabilities can be represented in matrix form as given below

$$
P^{(n)} = \begin{array}{c c}
 & \begin{array}{cccc} E_1 & E_2 & \cdots & E_m \end{array} \\
\begin{array}{c} E_1 \\ E_2 \\ \vdots \\ E_m \end{array} & \begin{pmatrix} P_{11}^{(n)} & P_{12}^{(n)} & \cdots & P_{1m}^{(n)} \\ P_{21}^{(n)} & P_{22}^{(n)} & \cdots & P_{2m}^{(n)} \\ \vdots & \vdots & & \vdots \\ P_{m1}^{(n)} & P_{m2}^{(n)} & \cdots & P_{mm}^{(n)} \end{pmatrix}
\end{array}
$$

Here, for example $P_{21}^{(n)}$ means the Probability that the system which occupies state E_2 will move to the state E_1 after n steps.



Markov Chain - Introduction :-

(6)

Let $P_j^{(0)}$ $(j = 0, 1, 2 \ldots)$ be the absolute Probability such that the system be in state E_j at time t_0 where E_j $(j = 0, 1, 2, 3 \ldots)$ denote the exhaustive and mutually exclusive outcomes of a system at any time.

we define

$$P_{ij} = P[X(t_n) = j \,/\, X(t_{n-1}) = i]$$

as the one-step transition Probability of going from state i at time t_{n-1} and to state j at time t_n.

It is also assumed here: that these Probabilities from state E_i to state E_j $\begin{pmatrix} i = 0, 1, 2 \ldots \\ j = 0, 1, 2 \ldots \end{pmatrix}$ are expressed in the matrix form as follows.

$$P = \begin{array}{cc} & E_0 \\ E_0 & \begin{pmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ P_{20} & P_{21} & P_{22} & \cdots \\ \vdots & \vdots & \vdots & \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \end{array}$$

This matrix, P is known as stochastic matrix (or) Homogeneous matrix.

The Probabilities P_{ij} must also satisfy the boundary Conditions

$$\sum P_{ij} = 1 \quad \forall \text{ and } P_{ij} \geq 0 \quad \forall \; i \,\&\, j$$

Definition :-

Markov Chain :-

The transition matrix P as defined above, together with the initial Probabilities $\{P_j^{(0)}\}$ associated with the state E_j ($j = 0, 1, 2 \ldots$) Completely define a markov Chain.

The markov chain are of two types

(i) ergodic markov chain

(ii) regular markov Chain.

(i) Ergodic markov chain :-

An ergodic markov chain has the Probability that it is possible to pass from one state to another in a finite number of steps regardless of Present state.

A special type of ergodic markov chain is the regular markov chain.

(ii) Regular markov chain :-

A regular markov chain is defined as a Chain having a transition matrix P such that for some power of P it has only non-zero Positive Probability values.

Thus all regular chains must be ergodic chains.

7

The easiest way to "Check if an ergodic Chain is regular" is to Continue squaring the transition matrix p until all the Zeros are removed.

The transition Probability may or may not be independent of n. if it is independent of n, then the markov Chain is Said to be Homogeneous or to have Stationary transition Probabilities.

If it is dependent on n, then the chain is said to be non- Homogeneous.

A markov chain, $\{X_n / n \geq 0)$ with K states when K is finite, is said to be a finite markov Chain. The transition matrix in this case is a Square matrix with K rows & Columns

If the possible values of X_n are ..... -2, -1, 0, 1, 2 - - . then the markov Chain is said to be denumerably infinite.

A stochastic matrix is a random matrix with non-negative elements and unit row sums

A stochastic matrix p is said to be regular if all the entries of some power p^m are Positive.

A stochastic matrix P is not regular If '1' occurs in the Principal diagonal.

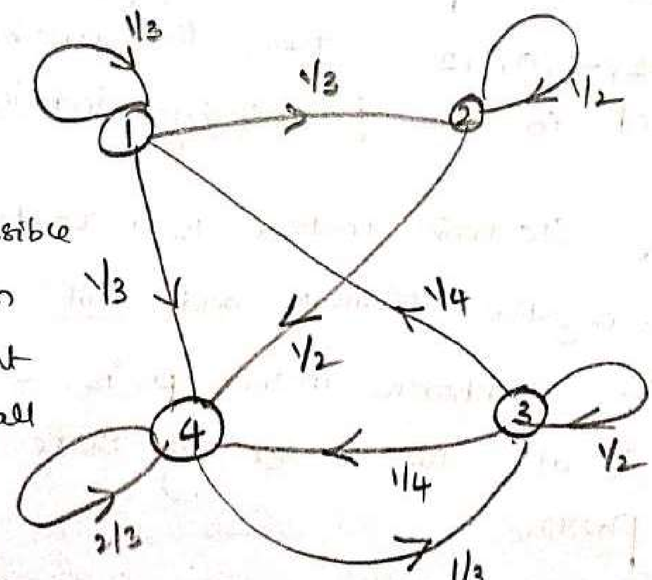
Problems

1) Determine if the following transition matrix is ergodic markov chain.

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

Sol:

$$\begin{array}{cc} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix} \end{array}$$



Here it is Possible to go from every Present state to all Other state

∴ It is ergodic markov chain

∴ The given transition matrix is ergodic markov chain

(8)

(2) Which of the following matrices are stochastic?

Sol:- (1) The row sums must equal to 1

(2) It must be a square matrix

(3) no negative values

(a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}_{2\times3}$ $\longrightarrow$ No it is not stochastic
Since it is not square matrix

(b) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{2\times2}$ $\longrightarrow$ Yes stochastic matrix
square
row sum = 1
no negative elements

(c) $\begin{pmatrix} 0 & 1 \\ 1/3 & 1/4 \end{pmatrix}_{2\times2}$ $\longrightarrow$ It is <u>not</u> <u>stochastic</u>.
$0+1=1$
$\frac{1}{3}+\frac{1}{4} \neq 1$

(d) $\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}_{2\times2}$ $\longrightarrow$ It is <u>stochastic</u> matrix
$1/2+1/2=1$
square.
no negative.

(e) $\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}_{2\times2}$ $\longrightarrow$ No it is not stochastic.
$\longrightarrow$ negative value is there (-1)

(f) $\begin{pmatrix} 0 & 2 \\ 1/4 & 1/4 \end{pmatrix}$ $\longrightarrow$ No it is not stochastic
row sum $\neq 1$

(3) which of the following stochastic matrices are regular.

Sol :- regular :-
(P is matrix $P^m \rightarrow$ only the non zero values)
no zero. m such that the non zero value
Power.

(a) $\begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{pmatrix} = P$

Principal diagonal if it '1' then it is not regular.
not regular since '1' is Present in the Principal diagonal.

(b) $\begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} = P$

$P^2 = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 3/8 & 3/8 & 1/4 \end{pmatrix}$ $\quad P^3 = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 7/16 & 7/16 & 1/8 \end{pmatrix}$

using Calculator

||ly P^4, P^5.. Here Zero Cannot be removed.
P_{13} P_{23} are Zero in all Powers $\therefore$ P is not regular.

9

(c) $P = \begin{pmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}$

$P^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix}$ (sign of changing zero's)

P^4 only one zero

Ily $P^5 = \begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 1/8 & 1/2 & 3/8 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}$ non-zero. all are +ve.

$\therefore$ P is regular (powers)

10

Classification of States, Examples of Markov Chain

In Markov Process, the states are Classified in Order to find the Communicating Classes.

The states of markov chain can be Partitioned into these Communicating Classes

Two states Communicate $\Longleftrightarrow$ it is Possible to go from each to other

ie:- states A & B Communicate $\Longleftrightarrow$ it is Possible to go from A to B & B to A

1) Transient
2) Periodic
3) Ergodic

1) Transient :- A state is Said to be transient if it is Possible to leave state and never return back.

2) Periodic :- A state is Said to be Periodic if it is not transient and that state is returned to only on multiples of some Positive integer greater than '1'. This integer is Known as Period of the state.

(3) Ergodic :- A state is said to be ergodic if it is neither transient nor Periodic.

1) The three state markov chain is given by the transition Probability matrix

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

Prove that the chain is Irreducible

Sol :- Given that the three state of markov chain is given by the transition matrix

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

The Condition for Irreducible markov chain is

$\boxed{P_{ij}^{(n)} > 0}$ for some n value and $\forall\, i \,\&\, j$

Then it can be said that a state can be reached from every other state.

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 0 & 2/3 & 1/3 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix} \end{array} \qquad P^{n} = \begin{bmatrix} 0.5 & 0.1666 & 0.3333 \\ 0.25 & 0.5833 & 0.1666 \\ 0.25 & 0.3333 & 0.41666 \end{bmatrix}$$

11

$$P^3 = \begin{bmatrix} 0.25 & 0.5 & 0.25 \\ 0.375 & 0.25 & 0.375 \\ 0.375 & 0.375 & 0.25 \end{bmatrix}$$

for P^2 & P^3, $P_{ij} > 0$

$\therefore$ It can be said that the given markov chain is Irreducible.

(a) The transition Probability matrix of a markov chain is given by $\begin{bmatrix} 0.3 & 0.7 & 0 \\ 0.1 & 0.4 & 0.5 \\ 0 & 0.2 & 0.8 \end{bmatrix}$

Verify whether matrix is Irreducible or not?

Sol: Given that the transition Probability matrix of a markov chain is

$$P = \begin{bmatrix} 0.3 & 0.7 & 0 \\ 0.1 & 0.4 & 0.5 \\ 0 & 0.2 & 0.8 \end{bmatrix}$$

First encode of matrix as + (or) 0 and call the encoded matrix Q

$$Q = \begin{bmatrix} + & + & 0 \\ + & + & + \\ 0 & + & + \end{bmatrix} \qquad Q^2 = \begin{bmatrix} + & + & 0 \\ + & + & + \\ 0 & + & + \end{bmatrix} \begin{bmatrix} + & + & 0 \\ + & + & + \\ 0 & + & + \end{bmatrix}$$

$$= \begin{bmatrix} + & + & + \\ + & + & + \\ + & + & + \end{bmatrix}$$

$\therefore$ All entries of Q^2 are strictly positive

$\therefore$ Markov matrix is Irreducible

$\therefore$ P is Irreducible.

(3) Check whether the following markov chain is regular and Ergodic.

$$\begin{bmatrix} 1 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \\ 0 & 1/2 & 1/2 & 1/2 \end{bmatrix}$$

Sol :- Given Markov Chain

$$P = \begin{bmatrix} 1 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \\ 0 & 1/2 & 1/2 & 1/2 \end{bmatrix}$$ Let us denote the given markov chain as 'p'

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

$$= \begin{bmatrix} 1+1/4+1/4+0 & 1/2+0+0+0 & 1/2+0+0+0 & 0+1/4+1/4+0 \\ 1/2+0+0+0 & 1/4+0+0+1/4 & 1/4+0+0+1/4 & 0+0+0+1/4 \\ 1/2+0+0+0 & 1/4+0+0+1/4 & 1/4+0+0+1/4 & 0+0+0+1/4 \\ 0+1/4+1/4+0 & 0+0+0+1/4 & 0+0+0+\frac{1}{4} & 0+1/4+1/4+1/4 \end{bmatrix}$$

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

Thus all the entries of P^n are positive.

$\therefore$ Markov chain is Regular & Ergodic.

12

Markov Chain

A random Process in which the occurence of future state depends on the immediately Preceeding state and only on it is Known as Markov Chain (or) Markov Process.

(next state depends on Current state)

Uses :- (1) Behaviour of Consumers in ~~to~~ terms of their brand loyality and switching pattern.

(2) Machine use to manufacture a Product.

[two state → working or not working at any point]

State :- A state is a Condition (or) Location of an Object in the system at a Particular time.

Assumptions :- (1) Finite number of State

(2) State are mutually exclusive

(3) State are Collectively Exhaustive

(4) Probability of moving from one state to other state is Constant over time

Transition Probability :-

The Probability of moving from one state to another state or remaining in the same state during a single time period is called the Transition Probability

Mathematically

$$P_{ij} = P\left(\text{Next state } S_j \text{ at } t=1 \,\middle/\, \begin{array}{l}\text{initial state}\\ S_i \text{ at } t=0\end{array}\right)$$

(i) initial state (j) next state

Transition Probability Matrix :- (TPM)

With the help of transition Probability matrix (TPM) we Predict the movement of system from One state to the next state.

(next state) $(n=1)$

$$P = \text{Initial state (i)} \; (n=0) \quad \begin{array}{c|ccc} & S_1 & S_2 & S_3 \\ \hline S_1 & P_{11} & P_{12} & P_{13} \\ S_2 & P_{21} & P_{22} & P_{23} \\ S_3 & P_{31} & P_{32} & P_{33} \end{array}$$

$$P_{11} = P\left[\text{in state } S_1 \text{ in next state at } t=1 \,\middle/\, \text{in state } S_1 \text{ in initial state at } t=0\right]$$

13

$P_{11} = P[S_1 \text{ at time } t=1 / S_1 \text{ at time } t=0]$

$P_{12} = P[S_2 \text{ at time } t=1 / S_1 \text{ at time } t=0]$

Ily $P_{21} = P[S_1 \text{ at time } t=1 / S_2 \text{ at time } t=0]$

[One-step Transition Probability]

$P_{11}^{(2)} = P(S_1 \text{ at time } t=2 / S_2 \text{ at time } t=0)$

[2-step Transition Probability]

$$P^{(2)} = \begin{array}{cc} & \text{State}(j) \\ \begin{array}{c} \text{State} \\ (i) \end{array} \begin{array}{c} S_1 \\ S_2 \\ S_3 \end{array} & \begin{bmatrix} P_{11}^{(2)} & P_{12}^{(2)} & P_{13}^{(2)} \\ P_{21}^{(2)} & P_{22}^{(2)} & P_{23}^{(2)} \\ P_{31}^{(2)} & P_{32}^{(2)} & P_{33}^{(2)} \end{bmatrix} \end{array}$$

Ily $P_{11}^{(n)} = P(S_1 \text{ at time } t=n / S_1 \text{ at time } t=0)$

[n-step Transition Probability]

$$P^{(n)} = \begin{array}{cc} & \text{State}(j) \\ \begin{array}{c} \text{State} \\ (i) \end{array} \begin{array}{c} S_1 \\ S_2 \\ S_3 \end{array} & \begin{bmatrix} P_{11}^{(n)} & P_{12}^{(n)} & P_{13}^{(n)} \\ P_{21}^{(n)} & P_{22}^{(n)} & P_{23}^{(n)} \\ P_{31}^{(n)} & P_{32}^{(n)} & P_{33}^{(n)} \end{bmatrix} \end{array}$$

Transition Probability Matrix [TPM]

Assumptions :-

(1) Row Sum = 1

(2) Each element of TPM is Probability

$\therefore 0 \le P_{ij} \le 1$ & non-negative.

(3) Square matrix because.

row show → Initial state.

Column show → Alternate state in next move.

(or) next state.

$$P = \text{Initial state } (i) \; (n=0) \quad \begin{array}{c|ccc} & S_1 & S_2 & S_3 \\ \hline S_1 & P_{11} & P_{12} & P_{13} \\ S_2 & P_{21} & P_{22} & P_{23} \\ S_3 & P_{31} & P_{32} & P_{33} \end{array}$$

Retention

Retention & gain ↓

→ Retention & Loss →

(14)

Markov Analysis

→ Markov Process is derived from Russian Mathematician Andrei Markov (1856 - 1922)

→ This type of Probabilistic models known as Stochastic Process, in which the Current state of a Process depends on all of its Previous state.

→ Uses :-

To examine and Predict the behaviour of Consumers in terms of their brand loyalty and switching Patterns to Other ~~branches~~ brands.

→ Usually Constructed in terms of transition Possibilities

→ Used to study the Stock market Price movements.

Stochastic Process :

A stochastic Process is a family of random variables

$\{ X_\beta : \beta \in A \}$

where β is the index Parameter assumes values in a Certain range A

A - Index set

Andrei Markov (1856 - 1922)

Markov is Particularly remembered for his study of Markov Chains sequences of random variables in which the future Variable is determined by the Present Variable but is independent of the way in which the Present state arose from its Predecessors

This work launched the theory of Stochastic Process.

15

1) What is a Markov Chain?

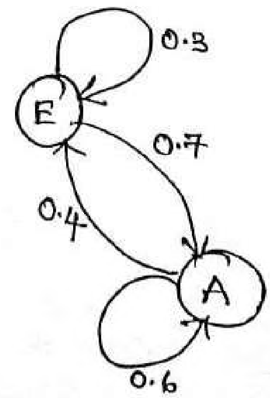
A) ~~A~~ markov Chain is a Mathematical Process that transitions from one state to another within a finite number of possible states.

It is a Collection of different states and Probabilities of a variable.

where its future Condition (or) state is Substantially dependent on its immediate Previous state.

2) What do you mean by Markov Chain?

A) A Markov chain (or) Markov Process is a stochastic model describing a sequence of possible events in which the Probability of each event depends only on the state attained in the Previous event.



(16)

Problems on Markov Chain

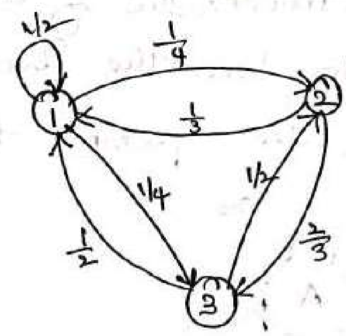
(1) Consider the Markov Chain with 3 states $S = \{1, 2, 3\}$ that has the following transition matrix

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

a) Draw the state transition diagram for this Chain

b) If we know $P(X_1 = 1) = P(X_1 = 2) = \frac{1}{4}$, find $P(X_1 = 3, X_2 = 2, X_3 = 1)$

Sol:- a)



A state transition diagram.

b) First we obtain

$$P(X_1 = 3) = 1 - P(X_1 = 1) - P(X_1 = 2)$$
$$= 1 - \frac{1}{4} - \frac{1}{4}$$
$$= \frac{1}{2}$$

we can now write $P(X_1 = 3, X_2 = 2, X_3 = 1) =$

$$= P(X_1 = 3) \cdot P_{32} \cdot P_{21}$$
$$= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{12}.$$

3) what is the First Order Markov Chain

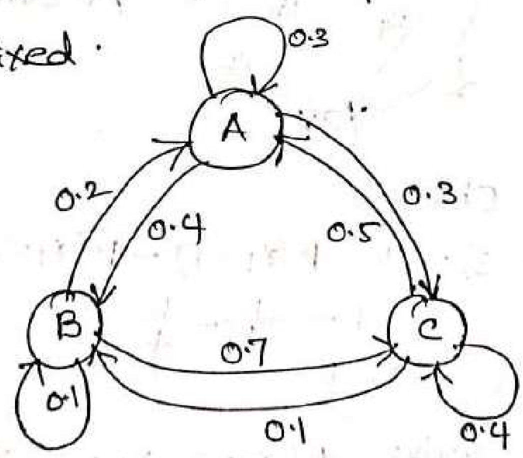
A) The Markov Chain of the First Order is One for which each subsequent state depends Only on the immediately Preceding one.

Markov Chains of second (or) higher Orders are the Processes in which the next state depends on two or more Preceding Ones.

(4) How do Markov Chains work?

A) A Markov Chain is a Mathematical system that experiences transitions from one state to another according to Certain Probabilistic rules.

The defining characteristic of a Markov Chain is that no matter how the Process arrived at its Present state, the Possible future states are fixed.



(17)

(5) What are Markov Processes Used for?

A) They are stochastic Processes for which the description of the Present state fully Captures all the information that Could influence the future evolution of the Process.

Predicting traffic flows,

Communications networks,

genetic issues and ~~gue~~ queues are examples where markov Chains Can be used to model Performance.

18

Markov Chain

Markov Chain :-

(1) 1- step Transition Probability

Probability of going from state i to j in step-1

$$P_{ij} = P(X_{n+1} = j / X_n = i)$$

(2) Transition Probability Matrix :-

Given step-1 transition Probabilities, we can write it in the Matrix form as

$$P = (P_{ij})_{i,j \in S} = \begin{bmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ P_{20} & P_{21} & P_{22} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

Note :-

(1) Each $P_{ij} \geq 0 \ \forall \ i, j$ and each row sum has to be 1

(2) The Matrix whose each row sum is 1 is called stochastic Matrix

(3) And If both row-sum & column-sum is 1, then it is called doubly stochastic Matrix.

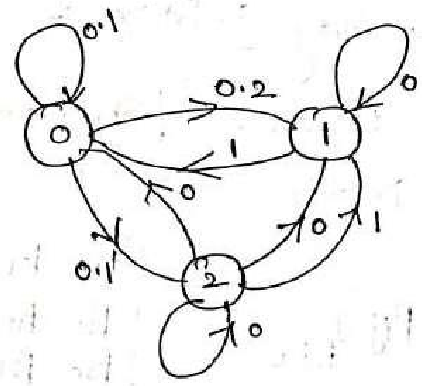
Ex :- (1) Let state space

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

One step transition matrix is given as

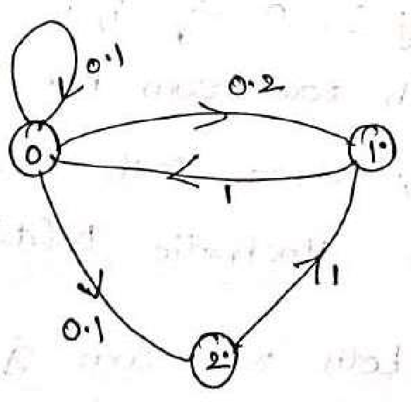
$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 0.1 & 0.2 & 0.1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{array}$$

State transition diagram is given as



(or)

If '0' then that should not be considered.



(19)

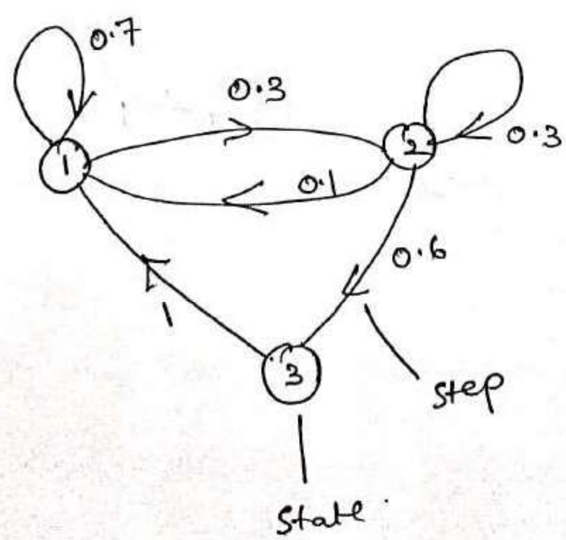
State Transition Matrix :-

Suppose we have given matrix P. we can draw a state transition diagram and vice-versa.

The diagram consists of circle, showing the state and edge from one circle to other circle, showing whether it is possible to go from that state to other.

Ex :- (2)

$$P = \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & 0.7 & 0.3 & 0 \\ 2 & 0.1 & 0.3 & 0.6 \\ 3 & 1 & 0 & 0 \end{array}$$



$\therefore$ Sum of each row = 1

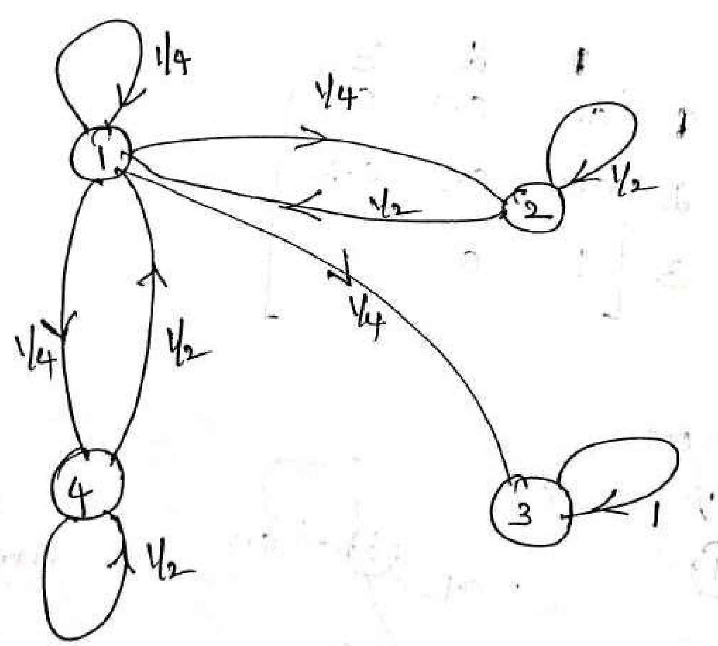
Hence stochastic matrix

ex :- 3

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

Sum of each row = 1

Hence Stochastic Matrix.



20

(3) <u>m- step transition Probability</u> :-

m-step transition Probability gives you the Probability of going from i^{th} state to j^{th} in m-steps

<u>ie</u> :- $$P_{ij}^{(m)} = P\left(X_{\underset{\uparrow \text{ m-steps}}{n+m}} = j \,/\, X_n = i\right) \quad \forall\ i, j \in S$$

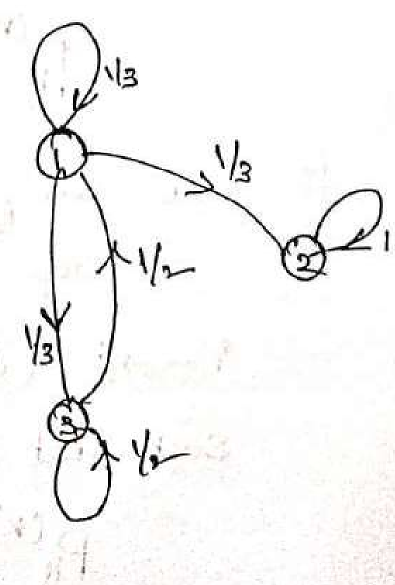
and m-step transition matrix is denoted by

$$P^{(m)} = P^m$$

If $\boxed{n=0}$ $$P_{ij}^{(m)} = P(X_m = j \,/\, X_0 = i) \quad \forall\ i, j \in S$$

<u>Example - 4</u> :-

$$P^m = \begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix} \end{array}$$



$$P_{13}^{(2)} = 1 \longrightarrow 1 \longrightarrow 3 \quad (\text{or})$$
$$1 \longrightarrow 3 \longrightarrow 3$$
$$= \left(\frac{1}{3} \times \frac{1}{3}\right) + \left(\frac{1}{3} \times \frac{1}{2}\right)$$
$$= \frac{1}{9} + \frac{1}{6} = \frac{6+9}{54} = \frac{15}{54} = \frac{5}{18}$$

$$P_{ij}^{(0)} = P\left(X_n = j \,/\, X_n = i\right) \quad (\text{when } m = 0)$$

$$= \begin{cases} 1 & ; \text{ when } i = j \\ 0 & ; \text{ when } i \neq j \end{cases}$$

(4) Accessibility :- We say, state j is accessible from the state i if $\exists$ $K > 0$ such that

$$P_{ij}^{(K)} > 0 \quad \text{for some } K.$$

ie :- We can go from state i to state j in K-step.

$i \longrightarrow j$ (j is accessible from i) if $\exists$ $K \in N$ such that $P_{ij}^{(K)} > 0$

$$P_{ij}^{(K)} = P\left(X_n = j \,/\, X_0 = i\right)$$

21

(5) Communicating state ($i \leftrightarrow j$)

Two states i & j are said to be Communicating

If $i \rightarrow j$ & $j \rightarrow i$

ie :- $i \leftrightarrow j$

$i \leftrightarrow j$ are Communicating Iff $\exists$ $m, n > 0$

such that $P_{ij}^{n} > 0$ &

$P_{ji}^{m} > 0$

$$P(x_n = j / x_n = i) = \begin{cases} 1 & ; i = j \\ 0 & ; i \neq j \end{cases}$$

ie :- If it is Possible to go from $i \rightarrow j$ in n-steps and $j \rightarrow i$ in m-steps for some $m, n > 0$.

(6) Equivalence Relation :-

(i) $i \leftrightarrow i$ Reflexive

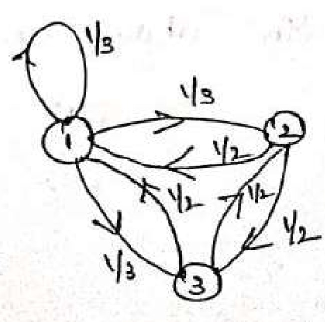
(ii) If $i \leftrightarrow j$ then $j \leftrightarrow i$ symmetry

(iii) If $i \leftrightarrow j$, $j \leftrightarrow k$ then $i \leftrightarrow k$.

Transitive

Ex:-

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix} \end{array} \longrightarrow$$



(7) Communicating Class :-

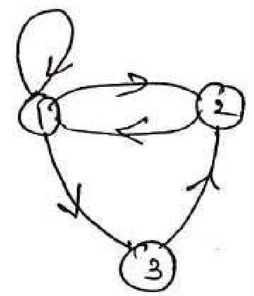
$$C(i) = \{ j \in S \,/\, i \longleftrightarrow j \} \quad i \in S$$

$= \{$ set of all states that Communicate with i $\}$

$i \in C(i)$, means i Communicate with i in 0-steps

Ex :-

(a)



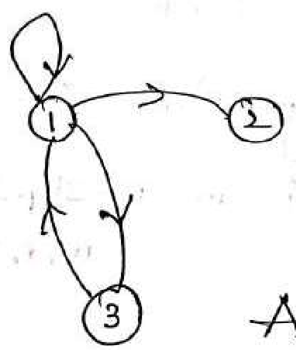
$C(1) = \{1, 2, 3\}$

$= C(2)$

$= C(3)$

$\Rightarrow \underline{\underline{C(1) = C(2) = C(3)}}$

(b)



$C(1) = \{1, 3\}$

$C(2) = \{2\}$

$C(3) = \{1, 3\}$

Also $S = \{1, 2, 3\} = \underbrace{\{1, 3\} \cup \{2\}}_{\text{Proper Partition}}$

Note :- If we get Only One class & it is equal to 'S', then it is said to be Irreducible Chain.

$\longrightarrow$ In above example

(a) is Irreducible.

(b) is not Irreducible.

22

(8) Irreducible :-

A markov Chain is said to be Irreducible if every state (vertex) Communicate with each other means every other state (vertex)

ie :- $C(i) = S$

Otherwise Chain is Called Reducible.

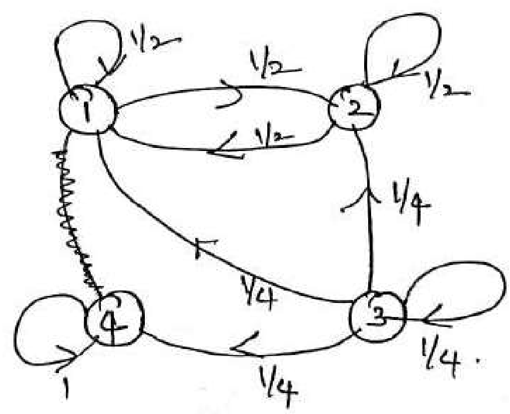
Ex :-

$$P = \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 1/2 & 1/2 & 0 & 0 \\ 2 & 1/2 & 1/2 & 0 & 0 \\ 3 & 1/4 & 1/4 & 1/4 & 1/4 \\ 4 & 0 & 0 & 0 & 1 \end{array}$$

$\therefore C(1) = \{1, 2\} = C(2)$

$C(3) = \{3\}$

$C(4) = \{4\}$



$\left(\begin{array}{l} \because 3 \to 1 \text{ but } 1 \not\to 3 \\ 3 \to 2 \text{ but } 2 \not\to 3 \\ 3 \to 4 \text{ but } 4 \not\to 3 \end{array} \right)$

$\therefore$ Total class $= 3$ $\{1, 2\}, \{3\} \; \{4\}$

Hence Chain is Reducible

(23)

Transition Probability Matrices

Transition Matrix :-

The Transition Probabilities

P_{jk} satisfy $P_{jk} \geq 0$,

$$\sum P_{jk} = 1 \quad \forall j$$

These Probabilities may be written in the matrix form as

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} & \cdots \\ P_{21} & P_{22} & P_{23} & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

This is Called the Transition Probability matrix.

Problems.

(1) Consider the matrix of transition Probabilities of a Product available in the market in two brands A & B

$$\begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{bmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{bmatrix} \end{array}$$

Determine the market share of each brand in equilibrium Position.

Sol :- Transition Probability matrix

$$T = \begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{bmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{bmatrix} \end{array}$$

At equilibrium $(A\ B)T = (A\ B)$ where $A+B=1$

$$(A\ B)\begin{pmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{pmatrix} = (A\ B)$$

$$0.9A + 0.3B = A$$

$$0.9A + 0.3(1-A) = A$$

$$0.9A - 0.3A + 0.3 = A$$

$$0.6A + 0.3 = A$$

$$0.4A = 0.3$$

$$A = \frac{0.3}{0.4} = \frac{3}{4}$$

$$B = 1 - A$$

$$B = 1 - \frac{3}{4} = \frac{1}{4}$$

Hence the Market share of brand A is 75% and the Market share of brand B is 25%.

(24)

(8) Parithi is either sad (S) or happy (H) each day. If he is happy in one day, he is sad on the next day by four times out of five. If he is sad on one day, he is happy on the next day by two times out of three. Over a long run, what are the chances that Parithi is happy on any given day?

Sol:- The Transition Probability matrix is

$$T = \begin{pmatrix} \frac{4}{5} & \frac{1}{5} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

At equilibrium $(S, H)\begin{pmatrix} \frac{4}{5} & \frac{1}{5} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix} = (S\ H)$ where $S+H=1$

$$\frac{4}{5}S + \frac{2}{3}H = S$$

$$\frac{4}{5}S + \frac{2}{3}(1-S) = S$$

On solving this, we get

$$\underline{\underline{S = \frac{10}{13}}} \quad \& \quad \underline{\underline{H = \frac{3}{13}}}.$$

In the long run, on a randomly selected day his chances of being happy is $\underline{\underline{10/13}}$.

(3) Akash bats according to the following traits

If he makes a hit (S), there is a 25% chance that he will make a hit his next time at bat

If he fails to hit (F) there is a 35% chance that he will make a hit his next time at bat.

Find the transition probability matrix for the data and determine Akash's long-range batting average.

Sol: The Transition Probability matrix is

$$T = \begin{pmatrix} 0.25 & 0.75 \\ 0.35 & 0.65 \end{pmatrix}$$

At equilibrium $(S\ F)\begin{pmatrix} 0.25 & 0.75 \\ 0.35 & 0.65 \end{pmatrix} = (S\ F)$

where $S+F=1$

$$0.25S + 0.35F = S$$

$$0.25S + 0.35(1-S) = S$$

On solving this we get

$$S = \frac{0.35}{1.10} \Longrightarrow S = 0.318$$

$$F = 0.682$$

$\therefore$ Akash's batting average is 31.8%

28

(4) 80% of students who do maths work during One study Period, will do the maths work at the next study Period.

30% of students who do english work during One study period, will do the english work at the next study Period.

Initially there were 60 students do maths and 40 students do english work. Calculate,

(i) The Transition Probability matrix.

(ii) The number of students who do maths work & english work for the next subsequent 2 study Periods.

Sol:

$$\begin{array}{cc} M & E \\ (60 & 40) \end{array} \begin{array}{c} \\ M \\ E \end{array} \begin{array}{c} \begin{array}{cc} M & E \end{array} \\ \begin{pmatrix} 0.8 & 0.2 \\ 0.7 & 0.3 \end{pmatrix}^2 \end{array}$$

$$\begin{array}{cc} M & E \\ (60 & 40) \end{array} \begin{array}{c} \\ M \\ E \end{array} \begin{array}{c} \begin{array}{cc} M & E \end{array} \\ \begin{pmatrix} 0.78 & 0.22 \\ 0.77 & 0.23 \end{pmatrix} \end{array}$$

$$= (46.8 + 30.8 \quad 13.2 + 9.2)$$

$$= (77.6 \quad 22.4)$$

Sol: Transition Probability matrix

$$T = \begin{array}{c c} & \begin{array}{cc} M & E \end{array} \\ \begin{array}{c} M \\ E \end{array} & \begin{pmatrix} 0.8 & 0.2 \\ 0.7 & 0.3 \end{pmatrix} \end{array}$$

After One study Period

$$\begin{array}{c} \begin{array}{cc} M & E \end{array} \\ \begin{pmatrix} 60 & 40 \end{pmatrix} \end{array} \begin{array}{c c} & \begin{array}{cc} M & E \end{array} \\ \begin{array}{c} M \\ E \end{array} & \begin{pmatrix} 0.8 & 0.2 \\ 0.7 & 0.3 \end{pmatrix} \end{array} = \begin{array}{c} \begin{array}{cc} M & E \end{array} \\ \begin{pmatrix} 76 & 24 \end{pmatrix} \end{array}$$

So in the very next study Period, there will be 76 students do maths work and 24 students do the English work

After two study Periods

$$\begin{array}{c} \begin{array}{cc} M & E \end{array} \\ \begin{pmatrix} 76 & 24 \end{pmatrix} \end{array} \begin{array}{c c} & \begin{array}{cc} M & E \end{array} \\ \begin{array}{c} M \\ E \end{array} & \begin{pmatrix} 0.8 & 0.2 \\ 0.7 & 0.3 \end{pmatrix} \end{array} = \begin{pmatrix} 60.8 + 16.8 & 15.2 + 7.2 \end{pmatrix}$$

$$= \underline{\underline{\begin{pmatrix} 77.6 & 22.4 \end{pmatrix}}}$$

After two study Periods there will be 78 (Approx) students do maths work and 22 (Approx) students do English work.

(26)

- **Stochastic Process :-**

Stochastic Process is a set of Random Variables ~~depending~~. depending on some real Parameter like time t. (based on time t)

Markov Process :- (or) Markov Chain :-

A random Process in which the occurence of future state depends on the immediately ~~Preceding~~. Preceeding state and only on it is known as the Markov Process (or) Markov Chain

Uses :- (1) Behaviour of Consumers in terms of their brand loyality and switting Pattern.

(2) Machine uses and manufacture a Product. (two States working or not working any Point)

(3) State is a Condition (or) Location of an Object in the system at Particular time.

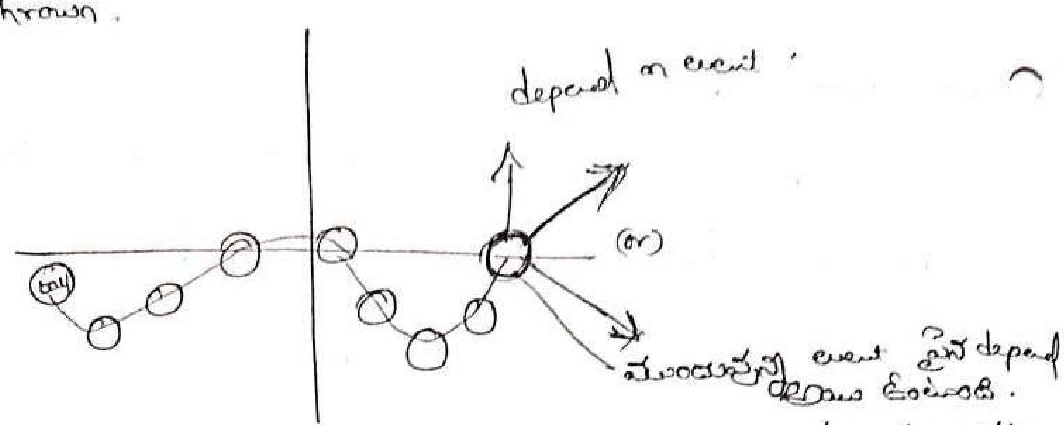
Assumption :- 1) Finite no of state.

2) State are mutually exclusive (working / Not working)

3) State are Collectively exclusive (possible states of brand loyalty)

4) Probability of moving from one State to other state is Constant over time.

ball thrown.



ball goes to other direction does not depend on the Curve, it depends to the event (ముందున్న event మీద base అయ్యి ఉంటుంది.)

Uses - (1)

Working / Not working. ⟶ 2 states.

Samsung, Moto, Apple.

Person Samsung. after seeing reviews Switching to one to another.

90%. going to one brand to another.

70%. 20%. 10%. brand loyality.

60%. 10%. 30%.

Working / Not working.

↓ New cell.

↓ Second hand Cell.

(27)

Transition Probability:-

The Probability of moving from one state to another State (or) are remaining in the same state during a Single time Period is Called the Transition Probability.

Mathematically :-

$$P_{ij} = P\left(\text{Next state } S_j \text{ at } t=1 \;\middle|\; \begin{array}{l}\text{initial state} \\ S_i \text{ at } t=0\end{array}\right)$$

initial next state

Transition Probability matrix :- (TPM)

with the Transition Probability matrix (TPM) we Predict the movement of system from One state to the next state.

Next State (j) (n=1)

$$P = \text{Initial State } (i)\ [n=0] \quad \begin{array}{c|ccc} & S_1 & S_2 & S_3 \\ \hline S_1 & P_{11} & P_{12} & P_{13} \\ S_2 & P_{21} & P_{22} & P_{23} \\ S_3 & P_{31} & P_{32} & P_{33} \end{array}$$

1 step Probability Transition

$P_{11} = P[\text{In state } S_1 \text{ next state at } t=1 \,/\, \text{In state } S_1 \text{ initial State at } t=0]$

(or)

$P(S_1 \text{ at time } t=1 \,/\, S_1 \text{ at time } t=0]$

$P_{12} = P[S_2 \text{ at time } t=1 \,/\, S_1 \text{ at time } t=0)$

$|||^{ly}$ $P_{21} = P[S_1 \text{ at time } t=1 \,/\, S_2 \text{ at time } t=0)$

One step Transition Probability Problem

2 Step Transition Probability.

$P_{11}^{2} = P\,(s_1 \text{ at time } t=2 \,/\, s_1 \text{ at time } t=0)$

$$P^{2} = \text{State } (i)\ \text{Initial} \begin{array}{c|ccc} & s_1 & s_2 & s_3 \\ \hline s_1 & P_{11}^{2} & P_{12}^{2} & P_{13}^{2} \\ s_2 & P_{21}^{2} & P_{22}^{2} & P_{23}^{2} \\ s_3 & P_{31}^{2} & P_{32}^{2} & P_{33}^{2} \end{array} \quad (\text{state } j) \text{ next state.}$$

$$P^{n} = \text{State } (i) \begin{array}{c|ccc} & s_1 & s_2 & s_3 \\ \hline s_1 & P_{11}^{n} & P_{12}^{n} & P_{13}^{n} \\ s_2 & P_{21}^{n} & P_{22}^{n} & P_{23}^{n} \\ s_3 & P_{31}^{n} & P_{32}^{n} & P_{33}^{n} \end{array} \quad \text{State } (j)$$

Ily $P_{11}^{n} = P\,[s_1 \text{ at time } t=n \,/\, s_1 \text{ at time } t=0]$

Properties of TPM :- (or) ~~Stochastic~~ Stochastic matrix.

1) It is a Square matrix $(0 \le x \le 1)$
2) All the entries are between 0 & 1
3) The sum of entries in any row must be 1 (non-negative)

ex:
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{array}{l} = 1+0+0=1 \\ = 0+0+1=1 \\ = \frac{1}{2}+\frac{1}{2}+0=1 \end{array}$$

28

- which of the following matrices are stochastic

① $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}_{2\times3}$ ✗ It is not a stochastic matrix since it is not square matrix.

not square so not stochastic.

② $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{2\times2}$

1) ✓ 2×2
2) ✓ between 0 & 1
3) non-negative.

$\left(\begin{matrix} 1+0=1 \\ 0+1=1 \end{matrix}\right)$ Sum of any row $= 1$ ✓

∴ It is stochastic matrix

③ $\begin{bmatrix} 0 & 1 \\ \frac{1}{3} & \frac{1}{4} \end{bmatrix}_{2\times2}$

1) ✓
2) ✓
3) $\frac{7}{12}$ row not equal to 1 ✗ each row.

$\frac{1}{3}+\frac{1}{4}=\frac{7}{12}$ ✗

It is not a stochastic matrix.

④ $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}_{2\times2}$

1) ✓
2) ✓
3) non-negative ✓.

∴ It is stochastic matrix

⑤ $\begin{bmatrix} 0 & 2 \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}_{2\times2}$ $\begin{matrix} = 2 \\ = \frac{2}{4} = \frac{1}{2} \end{matrix}$

1) ✓
2) ~~0-1~~ ✓
3) ✗ Sum of row $\neq 1$

∴ It is not stochastic.

LN4

Unit - 4

Multivariate Normal Distribution

→ **The Univariate Normal distribution**

$$X \sim N(\mu, \sigma^2)$$

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

→ **The Bi-variate Normal distribution**

$$(X, Y) \sim N(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho_{xy})$$

The Bi-variate Normal distribution is

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \cdot e^{\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{x-\mu_y}{\sigma_y}\right) + \left(\frac{x-\mu_y}{\sigma_y}\right)^2\right]}$$

→ **The Multi-variate Normal distribution**

The K-variate Normal distribution is given by

$$f(x_1, x_2 \cdots x_K) = f(x) = \frac{1}{(2\pi)^{K/2} \cdot |\Sigma|^{1/2}} \cdot e^{-\frac{1}{2}(x-\mu)' \Sigma^{-1} (x-\mu)}$$

where

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_K \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1K} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2K} \\ \vdots & \vdots & & \vdots \\ \sigma_{1K} & \sigma_{2K} & \cdots & \sigma_{KK} \end{bmatrix}$$

Multivariate Normal distribution :-

By Bi-variate Normal distribution we have

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; \quad \begin{array}{l} -\infty < x < \infty \\ -\infty < \mu < \infty \\ \sigma^2 > 0 \end{array}$$

(or)

$$f(x) = K \cdot e^{-1/2 (x-\mu)\alpha(x-\mu)} \longrightarrow (1)$$

where $K = \frac{1}{\sigma\sqrt{2\pi}}$ $\quad \alpha = \frac{1}{\sigma^2}$

$K \geq 0$ $\quad \alpha \geq 0$

Let us Consider the Random variable

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \text{ and}$$

Co-variance matrix $E(X) = \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$

Let us take $x - \mu = X - \mu$ &

$\alpha = A$

Equation (1) becomes as

$$f(X) = f(x_1 x_2 \dots x_p)$$

$$= K e^{-1/2 (x-\mu) A (x-\mu)} \longrightarrow (2)$$

we have to find K such that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[f(x_1 x_2 \ldots x_p)\,dx_p \cdots dx_2 \cdot dx_1\right] = 1$$

(By Using joint Pdf Properties)

(or)

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} K \cdot e^{-\frac{1}{2}(x-\mu)' A (x-\mu)}\, dx_p \cdots dx_2 dx_1 = 1 \longrightarrow (3)$$

Let A is a Positive definite

(ie:- Non-negative definite)

If A is a Positive definite matrix then $\exists$ a non-singular matrix 'C'

$x > 0 \longrightarrow$ Negative
$x \geq 0 \longrightarrow$ Non-negative
$|A| \geq 0 \longrightarrow$ Non-negative
$|A| < 0 \longrightarrow$ Negative

Such that $C' A C = I$

where I = Identity matrix

Let $x - \mu = C y$; where $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix}$

$$(x-\mu)' A (x-\mu) = (Cy)' A (Cy)$$
$$= y'C' A C y$$
$$= y' I y$$
$$A = y'y$$

$$\therefore (x-\mu)' A (x-\mu) = y'y$$

Let us Consider the Jacobian matrix

$$|J| = |C|$$

$$|J| = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} , & \frac{\partial y_1}{\partial x_2} \\ \vdots & \vdots \end{vmatrix}$$

From eqn (3)

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[K e^{-1/2\, y'y} \cdots dy_p \cdots dy_2 \cdot dy_1 \right] \quad |C| = 1$$

$\longrightarrow$ (4)

$$\therefore e^{-1/2\, y'y} = e^{-\frac{1}{2}(y_1^2 + y_2^2 + \cdots + y_p^2)}$$
$$= \prod_{i=1}^{p} e^{-1/2 \cdot y_i^2}$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} \qquad y' = [y_1\, y_2 \cdots y_p]$$

$$y'y = y_1^2 + y_2^2 + \cdots + y_p^2$$

Then eqn (4) becomes as

$$|c| \cdot k \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[e^{-1/2 y_1^2} \cdot e^{-1/2 y_2^2} \cdots e^{-1/2 y_p^2} \right] dy_p \cdot dy_1] = 1$$

$$|c| \cdot k \cdot \left\{ \prod_{i=1}^{p} \int_{-\infty}^{\infty} e^{-1/2 y_i^2} \, dy_i = 1 \right\}$$

$$|c| \cdot k \prod_{i=1}^{p} \sqrt{2\pi} = 1 \qquad \left(\because \int_{-\infty}^{\infty} e^{-1/2 y_i^2} dy_i = \sqrt{2\pi} \right)$$

$$|c| \cdot k \left(\sqrt{2\pi}\right)^p = 1$$

$$|c| \cdot k \, (2\pi)^{p/2} = 1$$

$$k = \frac{1}{|c| \, (2\pi)^{p/2}}$$

$$|c'Ac| = 1$$

$$|c'|\,|A|\,|c| = 1$$

$$|A|\,|c'|\,|c| = 1 \quad (\because \text{Associative Law})$$

$$|A|\,|c|\,|c| = 1$$

$$|A| = |c|^2 = 1$$

$$|c|^2 = \frac{1}{|A|}$$

$$c'Ac = 1$$

$$|c'Ac| = |I| = 1$$

$$|c'| = |c|$$

$$\implies |c| = \frac{1}{|A|^{1/2}}$$

$$\therefore \quad K = \frac{1}{(2\pi)^{p/2}} \cdot |A|^{-1/2}$$

Substitute 'K' value from eqn (2) becomes

$$\boxed{f(x) = \frac{1}{(2\pi)^{p/2}} \cdot |A|^{-1/2}\, e^{-1/2\,(x-\mu)'A\,(x-\mu)}}$$

This is called the Probability density function of Multi-Variate Normal distribution.

Def :- A P-dimensional vector of Random variables

$$X = x_1 x_2 \cdots x_p \quad ; \ -\infty < x_i < \infty$$

for $i = 1, 2 \cdots p$

is said to have a multi-variate normal distribution if its density function $f(x)$ is of the form

$$f(x) = f(x_1, x_2, x_3 \cdots x_p)$$

$$= \left(\frac{1}{2\pi}\right)^{p/2} \cdot |\Sigma|^{-1/2} \cdot e^{-\frac{1}{2}(x-M)' \Sigma^{-1} (x-M)}$$

where $m = (m_1 m_2 \cdots m_p)$ is the vector of means and Σ is the Variance Co-variance matrix of the multi variate normal distribution

The shortcut notation for this density is

$$X = N_p(m, \Sigma)$$

Properties of Multi-variate Normal distribution

(1) Joint density :- The multi-variate Normal distribution $MN(\mu, \Sigma)$ has Joint density

$$f_y(Y/\mu, \Sigma) = \frac{1}{(2\pi)^{n/2}} \cdot \frac{1}{(\det \Sigma)^{1/2}} \cdot e^{-\frac{1}{2}(y-\mu)'\Sigma^{-1}(y-\mu)}$$

(2) Shape :- The Contours of the Joint distributions are n- dimensional ellipsoids.

(3) Mean & Covariance specify the distribution :-

The $MN(\mu, \Sigma)$ Joint distribution is specified by μ and Σ only.

(4) Moment Generating function :-

The $MN(\mu, \Sigma)$ distribution has MGF

$$m(t) = e^{(\mu^T t + \frac{1}{2} t^T \Sigma t)}$$

where t is a real $n \times 1$ vector

5) Characteristic function :-

The $MN(\mu, \Sigma)$ distribution has CF

$$\psi(t) = e^{(i\mu^T t - \frac{1}{2} t^T \Sigma t)}$$

where t is a real $n \times 1$ vector

(6) Linear Combinations :-

(a) Let a be $n \times 1$

Y is $MN(\mu, \Sigma) \Longleftrightarrow$ Any Linear Combination $a^T y$ has a Univariate normal distribution.

(b) Let a be $n \times 1$

The distribution of $a^T Y$ is $N(a^T \mu, a^T \Sigma a)$

(c) Let a be $m \times 1$ &

B be $m \times n$

The distribution of 'm' random variables $a + BY$ is $MN(a + B\mu, B \Sigma B^T)$

(d) Let Z be 'n' Independent standard normal random variables

Then $Y = \mu + LZ$ with $LL^T = \Sigma$ has a $MN(\mu, \Sigma)$ distribution

(e) Again Let $LL^T = \Sigma$ then

$Z = L^{-1}(Y - \mu)$ has a $MN(0, I)$ distribution.

(7) Independence :-

(a) Y_i & Y_j are Independent $\Longleftrightarrow$ $\Sigma_{ij} = 0$

(b) Pairwise Independence of Y_i & Y_j for all $i \neq j \Longrightarrow$ Complete Independence.

(8) Marginal distribution :-

The m- dimensional marginal distribution of Y_1 is $MN(\mu_1, \Sigma_{11})$

(9) Conditional distribution :-

The m- dimensional distribution of Y_1 Conditional on Y_2 is

$$MN\left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(Y_2 - \mu_2),\ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$$

LAY

Distribution of Linear Combination

Distribution of $\boxed{Y = aX_1 + bX_2}$

where $X_i \sim N(\mu_i, \sigma_i)$

x_i is Normally distributed with Parameters μ_i, σ_i

x_i is independent.

by. Moment generating function

$$M_{X_i}(t) = e^{\mu_i t + \frac{t^2\sigma_i^2}{2}}$$

$$M_{X_1}(t) = e^{\mu_1 t + \frac{t^2\sigma_1^2}{2}}$$

$$M_{X_2}(t) = e^{\mu_2 t + \frac{t^2\sigma_2^2}{2}}$$

$$M_Y(t) = M_{aX_1+bX_2}(t) \quad (\because \text{moment generaty function.})$$

$$= M_{aX_1}(t) \cdot M_{bX_2}(t) \quad (\because x_1 \text{ indep } x_2)$$

$$= M_{X_1}(at) \cdot M_{X_2}(bt)$$

$$= e^{\mu_1 at + \frac{a^2t^2\sigma_1^2}{2}} \cdot e^{\mu_2 bt + \frac{b^2t^2\sigma_2^2}{2}}$$

$$= e^{\mu_1 at + \mu_2 bt + \frac{a^2t^2\sigma_1^2}{2} + \frac{b^2t^2\sigma_2^2}{2}}$$

$$= e^{(\mu_1+\mu_2)at + \frac{t^2}{2}(a^2\sigma_1^2 + b^2\sigma_2^2)}$$

$$\Rightarrow e^{\mu t + \frac{t^2\sigma^2}{2}} \quad X \sim N(\mu, \sigma^2)$$

$$Y \sim N\left(a(\mu_1+\mu_2),\ a^2\sigma_1^2 + b^2\sigma_2^2\right)$$

mean. variance

$$\boxed{Y = aX_1 + bX_2}$$

$X_1 \sim N(2,1)$

$X_2 \sim N(2,3)$

$X_1 + X_2 \sim N(4,4)$

$2X_1 + 3X_2 \sim N(2\cdot 2 + 3\cdot 2,\ 2^2\cdot 1 + 3^2\cdot 3)$

$\sim N(10, 31)$

- **Linear Combination :-**

Let x_1 & x_2 be two Random Variables with a, b as Constants.

Mean of ax_1 is $E(ax_1)$

$$= aE(x_1)$$

Variance of ax_1 is

$$Var(ax_1) = a^2 V(x_1)$$
$$= E[ax_1 - E(ax_1)]^2$$
$$= E[a(x_1 - E(x_1)]^2$$
$$= a^2 E[(x_1 - E(x_1)]^2$$
$$= a^2 V[x_1]$$
$$= a^2 \sigma_{11}$$

$\therefore$ Variance of $ax_1 = a^2\sigma_{11}$

Covariance between ax_1 & ax_1

$$Cov(ax_1, ax_1) = E[\{ax_1 - E(ax_1)\}\{ax_1 - E(ax_1)\}]$$
$$= E[a\{x_1 - E(x_1)\} a\{x_1 - E(x_1)\}]$$
$$= a^2 E[x_1 - E(x_1)]^2$$
$$= a^2 V(x_1)$$
$$= a^2\sigma_{11}$$

$$\therefore Cov(ax_1, ax_1) = \sigma^2 V(x_1)$$
$$= a^2\sigma_{11}$$

Mean of Linear Combination $ax_1 + bx_2$ is :-

$$E(ax_1+bx_2) = E[ax_1] + E[bx_2]$$
$$= aE(x_1) + bE(x_2)$$
$$= (a \; b)_{1\times 2} \begin{bmatrix} E(x_1) \\ E(x_2) \end{bmatrix}_{2\times 1}$$
$$\therefore E(\alpha' x) = \alpha' E(x)$$

Let $\alpha' = [a \; b]$

& $\alpha = \begin{bmatrix} a \\ b \end{bmatrix}$

$$E[x] = \begin{bmatrix} E(x_1) \\ E(x_2) \end{bmatrix}$$

Variance of $ax_1 + bx_2$ is :-

$$V(ax_1+bx_2) = a^2 V(x_1) + b^2 V(x_2) + 2ab\, Cov(x_1 x_2)$$
$$= a^2 \sigma_{11} + b^2 \sigma_{22} + 2ab\,\sigma_{12}$$

where $\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$ $= [a \; b]_{1\times 2} \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}_{2\times 2} \begin{bmatrix} a \\ b \end{bmatrix}_{2\times 1}$

$$= [a\sigma_{11} + b\sigma_{21} + a\sigma_{12} + b\sigma_{22}]_{1\times 2} \begin{bmatrix} a \\ b \end{bmatrix}_{2\times 1}$$

$$= \begin{bmatrix} a^2\sigma_{11} + ab\sigma_{21} + a^2\sigma_{12} + ab\sigma_{22} \\ ab\sigma_{11} + b^2\sigma_{21} + ab\sigma_{12} + b^2\sigma_{22} \end{bmatrix}$$

$$= [a^2\sigma_{11} + b^2\sigma_{22} + 2ab\,\sigma_{12}]$$
$$= \alpha' \Sigma \alpha$$

- Let X be a Random variable

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}_{p \times 1}$$

$$\left.\begin{aligned}
Y_1 &= C_{11} x_1 + C_{12} x_2 + \cdots + C_{1p} x_p \\
Y_2 &= C_{21} x_1 + C_{22} x_2 + \cdots + C_{2p} x_p \\
Y_3 &= C_{31} x_1 + C_{32} x_2 + \cdots + C_{3p} x_p \\
&\vdots \\
Y_q &= C_{q_1} x_1 + C_{q_2} x_2 + \cdots + C_{qp} x_p
\end{aligned}\right\} \text{Linear Combination.}$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_q \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1p} \\ C_{21} & C_{22} & \cdots & C_{2p} \\ \vdots & \vdots & & \vdots \\ C_{q1} & C_{q2} & \cdots & C_{qp} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

$$Y = CX$$

The mean of the Linear Combination is

$$E(Y) = E(CX) = CE(X)$$

The variance of the Linear Combination is

$$V(Y) = V(CX) = C^2 V(X)$$

$$\mathrm{Cov}(Y) - \mathrm{Cov}(Y,Y) = E\left[\{Y - E(Y)\}\{Y - E(Y)\}'\right]$$

$$= E\left[\{X - E(CX)\}\{CX - E(X)\}'\right]$$

$$= E\left[\{CX - CE(X)\}\{CX - CE(X)\}'\right]$$

$$= E\left[C\{X - E(X)\}C'\{C - E(X)\}\right]$$

$$= C \Sigma_X C'$$

Joint Probability mass function :- (Descrete)

Let X, Y be two dimensional Random variables then their Joint Probability mass function of X & Y is denoted by $P(X, Y)$ (or) $P(X = x, Y = y)$ (or) $P_{XY}(x, y)$

If it satisfies the following Conditions

(1) $P(x, y) \geq 0 \; \forall \; (x = X, \; y = Y)$

(2) $\sum_{x} \sum_{y} p(x, y) = 1$

Joint Probability density function :- (Continuous)

Let X, Y be two dimensional Random variables taking values x, y where $a \leq x \leq b, \; c \leq y \leq d$.

The function $f(X, Y)$ (or) $f(X = x, Y = y)$ (or) $f_{XY}(x, y)$ is said to be Joint Probability density function If it satisfies the following Conditions

(1) $f(x, y) \geq 0 \; \forall \; X = x, \; Y = y$

(2) $$\int_{a}^{b} \int_{c}^{d} f(x, y) \, dx \, dy = 1$$

Marginal Probability function :-

→ The Marginal Probability function of X is defined as

$$P_X(x_i) = \sum_{j=1}^{m} P(X=x_i, Y=y_j) \quad (\text{for descrete})$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy \quad (\text{for Continious})$$

→ The marginal Probability function of Y is defined as

$$P_Y(y_j) = \sum_{i=1}^{n} P(X=x_i, Y=y_j) \quad (\text{for descrete})$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx \quad (\text{for Continious})$$

$X \backslash Y$	y_1	y_2	---	y_j	---	y_m	Total
x_1	P_{11}	P_{12}	---	P_{1j}	---	P_{1m}	$P(x_1)$
x_2	P_{21}	P_{22}	---	P_{2j}	---	P_{2m}	$P(x_2)$
⋮	⋮	⋮				⋮	⋮
x_i	P_{i1}	P_{i2}	---	P_{ij}	---	P_{im}	$P(x_i)$
⋮	⋮	⋮		⋮		⋮	⋮
x_n	P_{n1}	P_{n2}	---	P_{nj}	---	P_{nm}	$P(x_n)$
Total	$P(y_1)$	$P(y_2)$	---	$P(y_j)$	---	$P(y_m)$	

Bi-Variate Probability distribution table.

Here $P_{ij} = P(X = x_i, Y = y_j)$

$$P(x_i) = \sum_{j=1}^{m} P(X = x_i, Y = y_j)$$

$$P(y_j) = \sum_{i=1}^{n} P(X = x_i, Y = y_j)$$

The Conditional Probability density function of X is given that $Y = y$ is defined as

$$\boxed{f_{X/Y}(x/y) = \frac{f_{XY}(x,y)}{f_Y(y)}}$$

$\|^{ly}$ The Conditional Probability density function of Y is given that $X = x$ is defined as

$$\boxed{f_{Y/X}(y/x) = \frac{f_{XY}(x,y)}{f_X(x)}}$$

Moment Generating Function : (MGF).

The MGF of a Random Variable X about origin having the Probability function $F(x)$ is defined as $M_X(t)$ and defined as

$$M_X(t) = E\left[e^{tx}\right]$$

$$= \int_{-\infty}^{\infty} e^{tx} f(x)\, dx \quad \text{(for Continious)}$$

$$= \sum_{x} e^{tx} f(x) \quad \text{(for descrete)}$$

where t is Real Parameter.

Characteristic function (C.F) :-

The Characteristic function of a Random Variable X is denoted by $\phi_x(t)$ and defined as

$$\phi_x(t) = E\left[e^{itx}\right]$$

$$= \int_{-\infty}^{\infty} e^{itx} f(x)\,dx \quad (\text{for Continuous})$$

$$= \sum_{x} e^{itx} p(x)\,dx \quad (\text{for descrete})$$

where t is a Real Parameter.

Note :-

→ If X is a Random Variable, and C is a Constant then

$$M_{cx}(t) = M_x(ct)$$

→ If X is a Random variable and C is a Constant then

$$\phi_{cx}(t) = \phi_x(ct)$$

Problem.

1) The two Random variables X & Y have the following Probability density function.

$$f(x,y) = \begin{cases} 2-x-y & ; \quad 0 \le (x,y) \le 1 \\ 0 & ; \quad \text{Otherwise} \end{cases}$$ then

Find (1) Marginal density function of x & y

(2) Conditional density functions

(3) Variance of x & Variance of y.

(4) Covariance of (x,y)

Sol: Given $f(x,y) = \begin{cases} 2-x-y & ; \; 0 \le (x,y) \le 1 \\ 0 & ; \; \text{Otherwise}. \end{cases}$

(1) The Marginal density function of X is given as

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy.$$

$$\Rightarrow \underbrace{\int_{-\infty}^{0} f(x,y)\,dx}_{\to 0} + \int_{0}^{1} f(x,y)\,dy + \underbrace{\int_{1}^{\infty} f(x,y)\,dy}_{\to 0}.$$

$$\Rightarrow 0 + \int_{0}^{1} (2-x-y)\,dy + 0$$

$$\Rightarrow \int_{0}^{1} (2-x-y)\,dy$$

$$\Rightarrow \int_0^1 (2-x-y)\,dy$$

$$\Rightarrow \left[2y - xy - \frac{y^2}{2}\right]_0^1$$

$$\Rightarrow \left[2(1) - x(1) - \frac{(1)^2}{2}\right] - [0-0-0]$$

$$\Rightarrow 2 - x - \frac{1}{2}$$

$$\Rightarrow \frac{3}{2} - x.$$

$$\therefore \boxed{f_X(x) = \frac{3}{2} - x.}$$

$$\therefore f_X(x) = \begin{cases} \frac{3}{2} - x & ; \text{ where } 0 \le x \le 1 \\ 0 & ; \text{ Otherwise.} \end{cases}$$

||ly <u>Marginal Probability density of Y is given by</u>

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx.$$

$$\Rightarrow \int_{-\infty}^{0} \underbrace{f(x,y)}_{0}\,dx + \int_0^1 f(x,y)\,dx + \int_1^{\infty} \underbrace{f(x,y)}_{0}\,dx.$$

$$\Rightarrow \int_0^1 (2-x-y)\,dx = \left[2x - \frac{x^2}{2} - xy\right]_0^1$$

$$\Rightarrow \left[2(1) - \frac{(1)^2}{2} - (1)y\right] - [0-0-0]$$

$$\Rightarrow \left[2 - \frac{1}{2} - y\right] = \underline{\underline{\frac{3}{2} - y.}}$$

$$\boxed{f_Y(y) = \frac{3}{2} - y}$$

$$\therefore f_Y(y) = \begin{cases} \frac{3}{2} - y & ; 0 \le x \le 1 \\ 0 & ; \text{Otherwise.} \end{cases}$$

(2) Conditional density Functions :-

The Conditional density function of X given $Y=y$ is

$$f_{X/Y}(x/y) = \frac{f(x,y)}{f_Y(y)} = \frac{2-x-y}{\frac{3}{2}-y} = \frac{2(2-x-y)}{3-2y}$$

$$= \frac{4-2x-2y}{3-2y}$$

$$0 \le (x,y) \le 1$$

The Conditional density function of Y given $X=x$ is

$$f_{Y/X}(y/x) = \frac{f(x,y)}{f_X(x)} = \frac{2-x-y}{\frac{3}{2}-x} = \frac{2(2-x-y)}{3-2y}$$

$$= \frac{4-2x-2y}{3-2x}$$

$$0 \le (x,y) \le 1$$

(3) Variance of X :-

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

$$E(x) = \mu = \int_{-\infty}^{\infty} x \, f_x(x) \, dx$$

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

$$\Rightarrow \int_{0}^{1} x \left(\frac{3}{2} - x\right) dx = \int_{0}^{1} \left(\frac{3}{2}x - x^2\right) dx$$

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

$$\Rightarrow \left[\frac{3}{2}\left(\frac{1}{2}\right) - \left(\frac{1}{3}\right)\right]$$

$$\Rightarrow \frac{3}{4} - \frac{1}{3} = \frac{9-4}{12} = \frac{5}{12}$$

$$\therefore \boxed{E(x) = \frac{5}{12}}$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx$$

$$\Rightarrow \underbrace{\int_{-\infty}^{0} x^2 f_X(x)\,dx}_{\to 0} + \int_{0}^{1} x^2 f_X(x)\,dx + \underbrace{\int_{1}^{\infty} x^2 f_X(x)\,dx}_{\to 0}$$

$$\Rightarrow \int_0^1 x^2 f_X(x)\,dx \Rightarrow \int_0^1 x^2\left(\frac{3}{2} - x\right)dx$$

$$\Rightarrow \int_0^1 \left(\frac{3}{2}x^2 - x^3\right)dx = \left[\frac{3}{2}\left(\frac{x^3}{3}\right) - \left(\frac{x^4}{4}\right)\right]_0^1$$

$$\Rightarrow \left[\frac{3}{2}\left(\frac{1}{3}\right) - \left(\frac{1}{4}\right)\right] - \left[0(0) - 0\right]$$

$$\Rightarrow \frac{3}{6} - \frac{1}{4} = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

$$\therefore \boxed{E(x^2) = \frac{1}{4}}$$

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

$$= \frac{1}{4} - \left(\frac{5}{12}\right)^2$$

$$= \frac{1}{4} - \frac{25}{144}$$

$$= \frac{36 - 25}{144} = \frac{11}{144}$$

Variance of Y :-

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

$$E(Y) = \int_{-\infty}^{\infty} y \, f_Y(y) \, dy$$

$$= \underbrace{\int_{-\infty}^{0} y \, f_Y(y) \, dy}_{\to 0} + \int_{0}^{1} y \, f_Y(y) \, dy + \underbrace{\int_{1}^{\infty} y \, f_Y(y) \, dy}_{\to 0}$$

$$= \int_{0}^{1} y \, f_Y(y) \, dy = \int_{0}^{1} y \left(\frac{3}{2} - y\right) dy$$

$$= \int_{0}^{1} \left(\frac{3}{2}y - y^2\right) dy$$

$$= \left[\frac{3}{2}\left(\frac{y^2}{2}\right) - \frac{y^3}{3}\right]_0^1$$

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

$$= \frac{5}{12}$$

$$\therefore \boxed{E(Y) = \frac{5}{12}}$$

- $E(Y^2) \Rightarrow \int_{-\infty}^{\infty} y^2 f_Y(y)\,dy$

$$\Rightarrow \int_{-\infty}^{0} y^2 \underbrace{f_Y(y)}_{0}\,dy + \int_{0}^{1} y^2 f_Y(y)\,dy + \int_{1}^{\infty} y^2 \underbrace{f_Y(y)}_{0}\,dy$$

$$\Rightarrow \int_{0}^{1} y^2 f_Y(y)\,dy = \int_{0}^{1} y^2 \left(\frac{3}{2} - y\right) dy$$

$$\Rightarrow \int_{0}^{1} \frac{3}{2}(y^2) - y^3)\,dy$$

$$\Rightarrow \left[\frac{3}{2}\left(\frac{y^3}{3}\right) - \frac{y^4}{4}\right]_0^1$$

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

$$\therefore \boxed{E(Y^2) = \frac{1}{4}}$$

$$\therefore V(Y) = E[Y^2] - [E(Y)]^2$$

$$= \frac{1}{4} - \left(\frac{5}{12}\right)^2$$

$$= \frac{1}{4} - \frac{25}{144} = \frac{36-25}{144} = \frac{11}{144}$$

$$\therefore \boxed{V(Y) = \frac{11}{144}}$$

(4) Covariance of $(X\,Y)$:-

$$Cov(X\,Y) = E[XY] - E[X]\cdot E[Y]$$

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\,dx\,dy$$

$$= \int_0^1\int_0^1 xy(2-x-y)\,dx\,dy$$

$$= \int_0^1\int_0^1 (2xy - x^2y - xy^2)\,dx\,dy$$

$$= \int_0^1 \left[\cancel{2}y\cdot\frac{x^2}{\cancel{2}} - y\cdot\frac{x^3}{3} - y^2\cdot\frac{x^2}{2}\right]_0^1 dy$$

$$\Rightarrow \left[\frac{y^2}{2} - \frac{1}{3}\left(\frac{y^2}{2}\right) - \frac{y^3}{6}\right]_0^1$$

$$\Rightarrow \frac{1}{2} - \frac{1}{6} - \frac{1}{6} = \frac{3-2}{6} = \frac{1}{6}$$

$$\therefore Cov(X\,Y) = \frac{1}{6} - \left(\frac{5}{12}\right)\left(\frac{5}{12}\right)$$

$$Cov(XY) = \frac{-1}{144}$$

(2) Two Random variables have the following Probability density function

$$f(x,y) = \begin{cases} K(4-x-y) & ; 0 \le (x,y) \le 2 \\ 0 & ; \text{otherwise.} \end{cases}$$

Find (1) Value of K

(2) Marginal density function of x & y

(3) Covariance of (X,Y)

Sol :- 1) Given $f(x,y) = \begin{cases} K(4-x-y) & ; 0 \le (x,y) \le 2 \\ 0 & ; \text{otherwise.} \end{cases}$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_0^2\int_0^2 f(x,y)\,dx\,dy = 1$$

$$\Rightarrow \int_0^2\int_0^2 K(4-x-y)\,dx\,dy = 1 \Rightarrow \int_0^2\int_0^2 (4K - Kx - Ky)\,dx\,dy = 1$$

$$\Rightarrow \int_0^2 \left[4Ky - Kxy - K\frac{y^2}{2}\right]_0^2 dx = 1$$

$$\Rightarrow \int_0^2 [8K - 2Kx - 2K]\,dx = 1 \Rightarrow \left[8Kx - 2K\frac{x^2}{2} - 2Kx\right]_0^2 dx = 1$$

$$\Rightarrow \left[6Kx - 2K\frac{x^2}{2}\right]_0^2 dx = [12K - 4K] = 1 \Rightarrow 8K = 1$$

$$\boxed{K = \frac{1}{8}}$$

(2) Marginal density function of x is given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy.$$

$$\Rightarrow \underbrace{\int_{-\infty}^{0} f(x,y)\,dy}_{\to 0} + \int_{0}^{2} f(x,y)\,dy + \underbrace{\int_{2}^{\infty} f(x,y)\,dy}_{\to 0}$$

$$\Rightarrow \int_{0}^{2} f(x,y)\,dy \Rightarrow \int_{0}^{2} \frac{1}{8}(4-x-y)\,dy$$

$$\Rightarrow \left[\frac{1}{2}y - \frac{1}{8}xy - \frac{y^2}{2} \times \frac{1}{8}\right]_0^2$$

$$\Rightarrow \left[\frac{1}{\cancel{2}}(\cancel{2}) - \frac{1}{8}xy - \frac{y^2}{16}\right]_0^2$$

$$\Rightarrow \left[1 - \frac{1}{\cancel{8}_4}x(\cancel{2}) - \frac{4}{16}) - (0)\right]$$

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

Marginal density function of y is given as

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx.$$

$$\Rightarrow \int_{0}^{2} f(x,y)\,dx \Rightarrow \int_{0}^{2} \frac{1}{8}(4-x-y)\,dx.$$

$$\Rightarrow \left[\frac{1}{8}(4x) - \frac{1}{8}\left(\frac{x^2}{2}\right) - \frac{1}{8}xy\right]_0^2$$

$$\Rightarrow \left[\frac{4(2)}{8} - \frac{1}{\cancel{8}_{2}}\left(\frac{\cancel{4}^{1}}{2}\right) - \frac{1}{\cancel{8}_{4}}(\cancel{2})\,y\right] - [0]$$

$$\Rightarrow 1 - \frac{1}{4} - \frac{1}{4}y$$

$$\Rightarrow \frac{3}{4} - \frac{1}{4}y \Rightarrow \frac{3-y}{4}$$

(3) <u>Covariance</u> $(x, y) = E[xy] - E[x]E[y]$

$$E[xy] = \int_0^2\int_0^2 \frac{1}{8}(4-x-y)\,dx\,dy$$

$$\Rightarrow \int_0^2\int_0^2 \left(\frac{1}{2} - \frac{x}{8} - \frac{y}{8}\right) dx\,dy$$

$$\Rightarrow \int_0^2 \left[\frac{1}{2}(x) - \frac{x^2}{2}\left(\frac{1}{8}\right) - \frac{xy}{8}\right]_0^2 dy$$

$$\Rightarrow \int_0^2 \left[1 - \frac{1}{4} - \frac{y}{4}\right] dy$$

$$\Rightarrow \left[y - \frac{1}{4}y - \frac{y^2}{2}\left(\frac{1}{4}\right)\right]_0^2$$

$$\Rightarrow \left[2 - \frac{2}{4} - \frac{\cancel{4}}{2} \times \frac{1}{\cancel{4}}\right]$$

$$\Rightarrow \left[2 - \frac{2}{4} - \frac{1}{2}\right] = \left[\frac{8-2-2}{4}\right] = \underline{1}.$$

$$\therefore \underline{E(xy) = 1}.$$

$$E(X) = \int_{-\infty}^{\infty} x \, f_X(x)\, dx$$

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

$$\Rightarrow \left[\frac{3}{4}\left(\frac{x^2}{2}\right) - \frac{x^3}{3\times 4}\right]_0^2$$

$$\Rightarrow \left[\frac{\cancel{4}}{2}\left(\frac{3}{\cancel{4}}\right) - \frac{8^2}{3\times \cancel{4}}\right] = \frac{3}{2} - \frac{2}{3} = \frac{9-4}{6} = \frac{5}{6}$$

$$E(Y) = \int_{-\infty}^{\infty} y \, f_Y(y)\, dy$$

$$\Rightarrow \int_0^2 y\left(\frac{3-y}{4}\right)dy = \int_0^2 \frac{3}{4}y - \frac{y^2}{4}\, dy$$

$$\Rightarrow \left[\frac{3}{4}\left(\frac{y^2}{2}\right) - \frac{y^3}{3\times 4}\right]_0^2 = \frac{5}{6}$$

$$\therefore \; Cov(XY) = E[XY] - E(X)\,E(Y)$$

$$= 1 - \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)$$

$$= 1 - \frac{25}{36}$$

$$= \frac{36-25}{36} = \frac{11}{36}$$

UNIT-4

Marginal distribution & Conditional distributions

Marginal distribution :- (descrete Random variable)

1) $f_X(x) = \sum_y f_{X,Y}(x, y)$ (descrete Random variable)

2) $f_Y(y) = \sum_x f_{X,Y}(x, y)$ (descrete Random variable)

Properties :-

(1) $f_X(x) \geq 0, \quad f_Y(y) \geq 0$

(2) $\sum_i f_X(x_i) = 1, \quad \sum_i f_Y(y_i) = 1$

Marginal distribution :- (Continious Random variable)

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

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

Properties :-

(1) $f_X(x) \geq 0, \quad f_Y(y) \geq 0$

2) $\int_{-\infty}^{\infty} f_X(x)\, dx = 1, \quad \int_{-\infty}^{\infty} f_Y(y)\, dy = 1$

Independent Random Variable :-

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

Conditional Distribution :-

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

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

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

- **Properties of Conditional :-**

(1) $f_{X/Y}(x/y) \geq 0$

$f_{Y/X}(y/x) \geq 0$

(2) $\sum_x f_{X/Y}(x/y) = \frac{\sum_x f_{X,Y}(x,y)}{f_Y(y)} = \frac{\cancel{f_Y(y)}}{\cancel{f_Y(y)}} = 1$

Independent :-

(1) $f_{X/Y}(x/y) = \frac{f_X(x)\cdot \cancel{f_Y(y)}}{\cancel{f_Y(y)}} = f_X(x)$

2) $f_{Y/X}(y/x) = \frac{\cancel{f_X(x)}\cdot f_Y(y)}{\cancel{f_X(x)}} = f_Y(y)$

Conditional Probability function :-

Let (X, Y) be a descrete two dimensional random variable, Then the Conditional Probability mass function of X, given $Y = y$

denoted by

$P_{X/Y}(x/y)$ and defined as

$$\boxed{P_{X/Y}(x/y) = \frac{P(X=x, Y=y)}{P(Y=y)}}$$ where $P(Y=y) \neq 0$

Mly the Conditional Probability mass function of Y given $X = x$

denoted by $P_{Y/X}(y/x)$ and defined as

$$\boxed{P_{Y/X}(y/x) = \frac{P(X=x, Y=y)}{P(X=x)}}$$ where $P(X=x) \neq 0$

ex:- Suppose $X = 1, 2, 3$
$Y = 1, 2, 3, 4$

X = Condition at $Y = 2$ fixed.

$$P_{X/Y}(x/2) = \frac{P(X=x, Y=2)}{P(Y=2)}$$

Here Y is fixed as $Y=2$ then we can change value of X Conditional Prob of X.

Supp $x=1$ $P_{X/Y}(1/2) = \dfrac{P(X=1, Y=2)}{P(Y=2)}$

(1) For the joint Probability distribution of two random variable X & Y given below

X \ Y	1	2	3	4	Total
1	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{10}{36}$
2	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{9}{36}$
3	$\frac{5}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{8}{36}$
4	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{5}{36}$	$\frac{9}{36}$
Total	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	1 Total

Find (1) Marginal distribution of X & Y &.

(2) Conditional distribution of X given the value of $Y=1$ and that of Y given the value of $X=2$.

Sol :-

X	1	2	3	4	Total
P(X)	$\frac{10}{36}$	$\frac{9}{36}$	$\frac{8}{36}$	$\frac{9}{36}$	$\Sigma P(x) = 1$

The Marginal dist of X is given as

$$P(X=x) = \sum_{y} P(X=x, Y=y)$$

$$P(X=1) = \sum_{y} P(X=1, Y=y)$$

$$= P(X=1, Y=1) + P(X=1, Y=2) + P(X=1, Y=3) + P(X=1, Y=4)$$

$$= \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \underline{\underline{\frac{10}{36}}}$$

||ly $P(X=2) = \frac{1}{36} + \frac{3}{36} + \frac{3}{36} + \frac{2}{36} = \frac{9}{36}$

$$P(X=3) = \frac{5}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{8}{36}$$

$$P(X=4) = \frac{1}{36} + \frac{2}{36} + \frac{1}{36} + \frac{5}{36} = \frac{9}{36}$$

$\therefore$ The Marginal dist of X

value of X	1	2	3	4	Total
$P(X=x)$	$\frac{10}{36}$	$\frac{9}{36}$	$\frac{8}{36}$	$\frac{9}{36}$	$\frac{36}{36} = 1$

- |ly The Marginal dist of Y is defined as

$$p(Y=y) = \sum_{x} p(X=x, Y=y)$$

$$p(Y=1) = \sum_{i} p(X=x_i, Y=1)$$

$$= p(X=1, Y=1) + p(X=2, Y=1) + p(X=3, Y=1) + p(X=4, Y=1)$$

$$= \frac{4}{36} + \frac{1}{36} + \frac{5}{36} + \frac{1}{36} = \underline{\underline{\frac{11}{36}}}$$

|ly $p(Y=2) = \frac{9}{36}$

$p(Y=3) = \frac{7}{36}$

$p(Y=4) = \frac{9}{36}$

The marginal dist of Y

Values of $Y=y$	1	2	3	4	$\sum p(y)$
$p(Y=y)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$=\underline{\underline{1}}$

(2) The Conditional Probability function of X given the value of Y=1 is defined as follows

$$P(X=x / Y=1) = \frac{P(X=x, Y=1)}{P(Y=1)} \quad \text{(Intersection)}$$

$$P(X=1/Y=1) = \frac{P(X=1, Y=1)}{P(Y=1)} = \frac{\frac{4}{36}}{\frac{11}{36}} = \frac{4}{11}$$

$$P(X=2/Y=1) = \frac{P(X=2, Y=1)}{P(Y=1)} = \frac{\frac{1}{36}}{\frac{11}{36}} = \frac{1}{11}$$

$$P(X=3/Y=1) = \frac{P(X=3, Y=1)}{P(Y=1)} = \frac{\frac{5}{36}}{\frac{11}{36}} = \frac{5}{11}$$

$$P(X=4/Y=1) = \frac{P(X=4, Y=1)}{P(Y=1)} = \frac{\frac{1}{36}}{\frac{11}{36}} = \frac{1}{11}$$

∴ The Conditional dist of X given Y=1 is

x	1	2	3	4	
$P(X=x/Y=1)$	$\frac{4}{11}$	$\frac{1}{11}$	$\frac{5}{11}$	$\frac{1}{11}$	

llly $P(Y=1/X=2) = \frac{P(Y=1, X=2)}{P(X=2)} = \frac{1/36}{\frac{9}{36}} = \frac{1}{9}$ //

Moment Generating Function :-

Moments :-

The r^{th} moment of a random variable X about any point A is defined as

$$E\left[(X-A)^r\right]$$

$\therefore$ The r^{th} moment of random variable X about point A =

$$E\left[(X-A)^r\right] = \begin{cases} \sum_x (x-A)^r \, p(x) & ; \text{descrete case} \\ \int_{-\infty}^{\infty} (x-A)^r . f(x)\, dx & ; \text{Continious case} \end{cases}$$

The r^{th} moment of a random variable X about origin is denoted by μ_r' and defined as

$$\mu_r' = E(X^r) = \begin{cases} \sum_x x^r \, p(x) & ; \text{descrete case} \\ \int_{-\infty}^{\infty} x^r f(x)\, dx & ; \text{Continious case} \end{cases}$$

$$E(X) = \sum_x x \, p(x)$$

$$\int_{-\infty}^{\infty} x f(x)\, dx .$$

(At origin)
$\boxed{A = 0}$

Clearly $\boxed{\mu_1' = E(X)}$ It is also called mean of random variable X and denoted by μ or $\overline{X}$

Also $\boxed{\mu_2' = E(X^2)}$

we Know that

$$Var(X) = E(X^2) - \{E(X)\}^2$$

$$Var(X) = \mu_2' - (\mu_1')^2$$

The r^{th} moment of random variable X about the mean $\overline{X}$ (or μ),

Usually denoted by μ_r is given by

$$\mu_r = E[(x-\bar{x})^r] = \begin{cases} \sum_x (x-\bar{x})^r p(x) & ; \text{ deserete case} \\ \int_{-\infty}^{\infty} (x-\bar{x})^r f(x)\,dx & ; \text{ Continuous Case} \end{cases}$$

$$\therefore \mu_2 = E[(x-\bar{x})^2] = Var(X) = \mu_2' - \mu_1'^2$$

(Second moment is var of X)

Moment Generating function :- (MGF)

The moment Generating function (MGF) of a Random variable X (about Origin) having the Probability function $f(x)$ is given by

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} f(x) & ; \text{ for descrete} \\ \int e^{tx} f(x)\,dx & ; \text{ for Continuous} \end{cases}$$

The Integration (or) Summation being extended to the entire range of x,

t be the real Parameter and it is being assumed that the R.H.S is absolutely Convergent

Thus $M_X(t) = E(e^{tx}) = E\left(1 + tx + \frac{t^2x^2}{\lfloor 2} + \cdots \frac{t^r \cdot x^r}{\lfloor r} + \cdots\right)$

$$\left(e^x = 1 + x + \frac{x^2}{2!} + \cdots\right)$$

e^x expansion

(Constant $E(f) = 1$)

$$M_X(t) = E(1) + tE(x) + \frac{t^2}{\lfloor 2} E(x^2) + \cdots \frac{t^r}{\lfloor r} E(x^r) + \cdots$$

$$M_X(t) = 1 + t\mu_1' + \frac{t^2}{\lfloor 2} \cdot \mu_2' + \cdots + \frac{t^r}{\lfloor r} \mu_r' + \cdots \quad \text{(1)}$$

$$\boxed{M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{\lfloor r} \cdot \mu_r'}$$

In general the moment generating function of random variable x about the point 'a' is defined as

$$M_x(t)\ (\text{about } a) = E\left[e^{t(x-a)}\right]$$

$$\left(M_x(t)\ (\underset{a=0}{\text{about Origin}})\right) = E\left(e^{tx}\right)$$

$$M_x(\theta) = E\left(e^{\theta x}\right)$$

(1) Let X & Y be two random variables with Joint Prob density function

$$f(x,y) = Axy \quad ; \; 0 < x < y < 1$$
$$= 0 \quad ; \; \text{Otherwise}$$

Find A, Also find the marginal density function of X & Y.

We know that

$$\iint f(x,y)\,dx\,dy = 1$$

Sol :-

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

$$\Rightarrow \int_0^1 \int_0^y Axy \; dx\,dy = 1$$

$$\Rightarrow \int_0^1 A \cdot y \left(\frac{x^2}{2}\right)_0^y dy = 1$$

$$\cancel{A \cdot \int_0^1 \frac{y^3}{2} dy = 1 \Rightarrow A = 8}$$

$$\Rightarrow \int_0^1 \frac{A}{2} \, y(y^2 - 0)\,dy = 1$$

$$\Rightarrow \frac{A}{2} \int_0^1 y^3\,dy = 1$$

$$\Rightarrow \frac{A}{2}\left(\frac{y^4}{4}\right)_0^1 = 1$$

$$\Rightarrow \frac{A}{8}(1 - 0^4) = 1 \Rightarrow \boxed{A = 8}$$

$$\therefore f(x,y) = \begin{cases} 8xy & ; 0 < x < y < 1 \\ 0 & ; \text{Otherwise} \end{cases}$$

<u>Marginal density function of X</u>.

$$f(x) = \int_x^1 f(x,y)\,dy =$$

$$= \int_x^1 8xy\,dy = 8x\left(\frac{y^2}{2}\right)_x^1$$

$$= 8x\left(\frac{1}{2} - \frac{x^2}{2}\right)$$

$$f(x) = \begin{cases} 4x(1-x^2) & ; 0 < x < 1 \\ 0 & ; \text{Otherwise} \end{cases}$$

<u>Marginal density function of Y</u> :-

$$f(y) = \int_0^y f(x,y)\,dx = \int_0^y 8xy\,dx$$

$$= 8y\left(\frac{x^2}{2}\right)_0^y$$

$$= 4y(y^2 - 0^2) = \underline{\underline{4y^3}}.$$

$$\therefore f(y) = \begin{cases} 4y^3 & ; 0 < y < 1 \\ 0 & \text{Otherwise} \end{cases}$$

(2) If $f(x,y) = \begin{cases} \frac{1}{8}(6-x-y) & ; \ 0<x<2 \ \& \ 2<y<4 \\ 0 & ; \ \text{Otherwise} \end{cases}$

is a Joint Probability density function then Find $P(X<1 \cap Y<3)$, $P(X+Y<3)$ & $P(X<1 / Y<3)$

Sol :- $P(X<1 \cap Y<3) =$

$$= \int_0^1 \int_2^3 f(x\,y)\, dy\, dx = \int_0^1 \int_2^3 \frac{1}{8}(6-x-y)\, dy\, dx$$

$$= \frac{1}{8} \int_0^1 6(y)_2^3 - x(y)_2^3 - \left(\frac{y^2}{2}\right)_2^3 dx$$

$$= \frac{1}{8} \int_0^1 6(3-2) - x(3-2) - \frac{1}{2}(9-4)\, dx$$

$$= \frac{1}{8} \int_0^1 \left(6 - x - \frac{5}{2}\right) dx$$

$$= \frac{1}{8} \int_0^1 \left(\frac{7}{2} - x\right) dx = \frac{1}{8}\left[\frac{7}{2}(x)_0^1 - \left(\frac{x^2}{2}\right)_0^1\right]$$

$$= \frac{1}{8}\left[\frac{7}{2}(1-0) - \frac{1}{2}(1-0)\right]$$

$$= \frac{1}{8}\left[\frac{7}{2} - \frac{1}{2}\right] = \frac{1}{8}\left(\frac{6}{2}\right) = \frac{\cancel{6}\,3}{\cancel{16}\,8} = \underline{\underline{\frac{3}{8}}}$$

(ii) $P(x+y<3)$

$x+y<3$
$y<3-x$
$x<3-y$

$$= \int_0^2 \int_2^{3-x} f(x,y)\,dy\,dx$$

$$= \int_0^2 \int_2^{3-x} \frac{1}{8}(6-x-y)\,dy\,dx$$

$$= \frac{1}{8}\int_0^2 6(y)_2^{3-x} - x(y)_2^{3-x} - \left(\frac{y^2}{2}\right)_2^{3-x}\,dx$$

$$= \frac{1}{8}\int_0^2 6[3-x-2] - x[3-x-2] - \frac{1}{2}\left[(3-x)^2-4\right]dx$$

$$= \frac{1}{8}\int_0^2 6(1-x) - x(1-x) - \frac{1}{2}\left[9+x^2-6x-4\right]dx$$

$$= \frac{1}{8}\int_0^2 6-6x-x+x^2-\frac{5}{2}-\frac{x^2}{2}+3x)\,dx$$

$6-\frac{5}{2} = \frac{12-5}{2} = \frac{7}{2}$

$$= \frac{1}{8}\int_0^2\left(\frac{7}{2}-4x+\frac{x^2}{2}\right)dx$$

$$= \frac{1}{8}\left[\frac{7}{2}(x)_0^2 - 4\left(\frac{x^2}{2}\right)_0^2 + \frac{1}{2}\left(\frac{x^3}{3}\right)_0^2\right]$$

$$= \frac{1}{8}\left[\frac{7}{2}(2-0) - 2(4-0) + \frac{1}{6}(8-0)\right]$$

$$= \frac{1}{8}\left[7-8+\frac{8}{6}\right] = \frac{1}{8}\left[-1+\frac{8}{6}\right] = \frac{1}{8}\left[\frac{-6+8}{6}\right]$$

$$= \frac{1}{8}\left(\frac{2}{6}\right) = \frac{1}{24}$$

• (iii) $P(X<1 / Y<3)$

$$= \frac{P(X<1 \cap Y<3)}{P(Y<3)}$$

$$= \frac{3/8}{P(Y<3)}$$

<u>marginal density function of Y</u>

$$f(y) = \int_0^2 f(x,y)\,dx$$

$$= \int_0^2 \frac{1}{8}(6-x-y)\,dx$$

$$= \frac{1}{8}\left[6(x)_0^2 - \left(\frac{x^2}{2}\right)_0^2 - y\,(x)_0^2\right]$$

$$= \frac{1}{8}\left[6(2-0) - \frac{1}{2}(4-0) - y(2-0)\right]$$

$$= \frac{1}{8}\left[12-2-2y\right] = \frac{1}{8}\left[10-2y\right]$$

$$\therefore f(y) = \begin{cases} \dfrac{10-2y}{8} & ; 2<y<4 \\ 0 & ; \text{Otherwise.} \end{cases}$$

$$P(Y<3) = \int_2^3 f(y)\,dy$$

$$= \int_2^3 \frac{10-2y}{8}\,dy$$

$$= \frac{1}{8}\left[10(y)_2^3 - \cancel{2}\left(\frac{y^2}{\cancel{2}}\right)_2^3\right]$$

$$= \frac{1}{8}\left[10(3-2) - (9-4)\right]$$

$$= \frac{1}{8}\left[10-5\right] = \frac{1}{8}[5] = \frac{5}{8} //$$

$$\therefore\ P(X<1 / Y<3) = \frac{P(X<1 \cap Y<3)}{P(Y<3)}$$

$$= \frac{3/8}{P(Y<3)} = \frac{3/8}{5/8} = \frac{3}{5} //.$$

• Marginal distributions & Expected Values

1) Two Random variables have joint Probability distributions

$P(X=x, Y=y)$ given :

X \ Y	0	1	2
0	3/24	2/24	1/24
1	2/24	5/24	2/24
2	7/24	1/24	2/24

what are the Marginal Probability distributions of X & Y and Expected values E(X) & E(Y)

Sol :- By the above table we have :

x	0	1	2
$P(X=x)$	6/24	9/24	10/24

y	0	1	2
$P(Y=y)$	12/24	8/24	5/24

These are Marginal Probability distribution

$$E(X) = (0)\left(\frac{6}{24}\right) + (1)\left(\frac{9}{24}\right) + (2)\left(\frac{10}{24}\right)$$

$$E(X) = \frac{29}{24}$$

$$E(Y) = (0)\left(\frac{12}{24}\right) + (1)\left(\frac{8}{24}\right) + (2)\left(\frac{5}{24}\right)$$

$$E(Y) = \frac{18}{24} = \frac{3}{4}$$

Conditional Probability distribution, Independence.

Find the Conditional Probability distribution of $(X = x / Y = 1)$ state If X & Y are Independent.

X \ Y	0	1	2
0	3/24	2/24	1/24
1	2/24	5/24	2/24
2	7/24	1/24	2/24

x	0	1	2
$P(X=x / Y=1)$	2/8	5/8	1/8

X & Y are not Independent because Y=1 actually changes the Probability distribution we have

x	0	1	2
$P(X=x)$	$6/24$	$9/24$	$10/24$

$\neq \quad \neq \quad \neq$

x	0	1	2
$P(X=x)/(Y=1)$	$2/8$	$5/8$	$1/8$

X & Y are not Independent.

X & Y are not Independent, because $Y=1$ actually changes the Probability distribution of X.

Marginal distribution VS. Conditional distribution.

1) The table below shows the Color Preference for boys & girls.

	Red	Bluee	Green	Total.
Boys.	1	5	8	14
girls.	6	4	3	13
Total	7	9	11	27

Sol:- a) Find the marginal distribution for Gender Variable.

$$\text{Boys} = \frac{14}{27} = 0.5185 = 51.85\%$$

$$\text{Girls} = \frac{13}{27} = 0.4815 = 48.15\%$$

b) Find the marginal distribution for the Colour Preference Variable

Red.

$$\frac{7}{27} = 0.2593$$
$$= 25.93\%$$

Blue

$$\frac{9}{27} = 0.3333$$
$$= 33.33\%$$

Green

$$\frac{11}{27} = 0.4074$$
$$= 40.74\%$$

(c) Find the Conditional distribution for the Gender Variable.

		Red	Blue	Green
Gender	Boys	1/14 = 7.14%	5/14 = 35.71%	8/14 = 57.14%
	Girls	6/13 = 46.15%	4/13 = 30.77%	3/13 = 23.08%

Look at each Cell Compare to total row

(d) Find the Conditional distribution for Colour Variable.

	Red	Blue	Green
Boys	1/7 = 14.29%	5/9 = 55.56	8/11 = 72.73%
Girls	6/7 = 85.71%	4/9 = 44.44%	3/11 = 27.27%
	100%	100%	100%

Look at each Cell Compare to total Column.

e) Based on these Calculations, Is there an association between Colour Reference & gender explain.

Sol: Yes, there is an association between the Colour Reference and gender 14.29% of boys Refer red Compared to 85.71% of girls.

55.56% of boys Refer blue Compared to 44.44% of girls and.

72.73% of boys Refer green. Compared to 27.27% of girls.

444.

7) Partial & Multiple Correlation Coefficient

→ Partial Correlation is Called Net Correlation
It is a study of relationship between one dependent variable and One independent Variable by Keeping the Other independent Variable Constant.

→ Simple Correlation between two variables is Called Zero Order Coefficient, here no factors held Constant

Ex :- r_{12}, r_{13}, r_{21}

→ If a Partial Correlation is studied between two variables by keeping a third variable Constant. it would be Called First Order Correlation Coefficient. here One Variable is Kept Constant.

Ex :- $r_{12\cdot3}$, $r_{23\cdot1}$, $r_{13\cdot2}$

→ If a Partial Correlation is studied between two variables by Keeping two Order variables Constant it would be Called Second Order Correlation Coefficient

Ex :- $r_{12\cdot34}$

→ Partial Correlation by the following Correlation

$$\boxed{r_{12.3} = \frac{r_{12} \cdot r_{13} \cdot r_{23}}{\sqrt{1-r_{13}^2}\ \sqrt{1-r_{23}^2}}}$$

Multiple Correlation :-

In multiple Correlation we study Correlation between 3 or more Variables at a time.

In Case of multiple Correlation the effect of all independent on a dependent factors is studied

The Coefficient of multiple Correlation is represented by 'R'

→ If there are 3 variables X_1, X_2, X_3 then $R_{1.23}$ is multiple Correlation Coefficient with X_1 as dependent variable, X_2, X_3 are Independent Variables

→ $R_{2.31}$ is multiple Correlation Coefficient with X_2 as dependent Variable, X_1, X_3 are Independent Variables.

→ $R_{3.21}$ is multiple Correlation Coefficient with X_3 as dependent Variable X_2, X_1 are Independent Variables.

Multiple Correlation Coeffient

$$R_{1\cdot 23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2 r_{12} r_{13} \cdot r_{23}}{1 - r_{23}^2}}$$

$$R_{2\cdot 13} = \sqrt{\frac{r_{12}^2 + r_{23}^2 - 2 r_{12} r_{23} \cdot r_{13}}{1 - r_{13}^2}}$$

$$R_{3\cdot 12} = \sqrt{\frac{r_{13}^2 + r_{23}^2 - 2 r_{13} r_{23} \cdot r_{12}}{1 - r_{12}^2}}$$

Problem

1) The following Zero Order Correlation Coefficent are given as

$r_{12} = 0.98 \quad r_{13} = 0.44 \quad r_{23} = 0.54$

Calculate multiple Correlation Coefficient treating 1st Variable is dependent Variable and Second and Third as Independent Variables.

Sol :-

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{23}^2}}$$

$$= \sqrt{\frac{(0.98)^2 + (0.44)^2 - 2(0.98)(0.44)(0.54)}{1 - (0.54)^2}}$$

$$= \sqrt{\frac{0.9604 + 0.1936 - 0.4657}{0.7084}}$$

$$= \frac{0.6883}{0.7084} = \sqrt{0.9716}$$

$$= 0.986 \simeq 0.99$$

(2) X_1, X_2, X_3 are three variates measured from their means with $N=10$, $\Sigma x_1^2 = 90$, $\Sigma x_2^2 = 160$, $\Sigma x_3^2 = 40$, $\Sigma x_1 x_2 = 60$, $\Sigma x_2 x_3 = 60$, $\Sigma x_3 x_1 = 40$

Calculate the multiple Correlation Coefficient $R_{1.23}$

- Sol:- First we Calculate r_{12}, r_{13} & r_{23}

$$r_{12} = \frac{\sum x_1 x_2}{\sqrt{\sum x_1^2 \cdot \sum x_2^2}} = \frac{60}{\sqrt{(90)(160)}} = \underline{\underline{0.5}}.$$

$$r_{13} = \frac{\sum x_1 x_3}{\sqrt{\sum x_1^2 \cdot \sum x_3^2}} = \underline{\underline{0.67}}.$$

$$r_{23} = \frac{\sum x_2 x_3}{\sqrt{\sum x_2^2 \cdot \sum x_3^2}} = \underline{\underline{0.75}}$$

$$R_{1.23} = \sqrt{\frac{(r_{12})^2 + r_{13}^2 - 2 r_{12} r_{13} r_{23}}{1 - r_{23}^2}}$$

$$= \sqrt{\frac{(0.5)^2 + (0.67)^2 - 2(0.5)(0.67)(0.75)}{1 - (0.75)^2}}$$

$$= \sqrt{\frac{0.6989 - 0.5025}{0.4375}}$$

$$= \sqrt{\frac{0.1964}{0.4375}} = \sqrt{0.4489} = \underline{\underline{0.67}}.$$

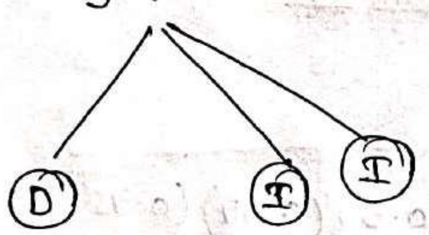
Zero Order Coefficients :- Correlation between two variables only, r_{12}, r_{13}, r_{23}, No variable are held constant.

First Order Coefficients :- One variable has been held constant $r_{12\cdot3}$, $r_{13\cdot2}$, $r_{23\cdot1}$

Second Order Coefficients :- Two variables has been held constant $r_{13\cdot24}$, $r_{23\cdot14}$ $\underline{r_{12\cdot34}}$.

First Order

3 Variables



$r_{12\cdot3}$

1 = D.V
2,3 = I.V.

$r_{13\cdot2}$

1 = D.V
3,2 = I.V

$r_{23\cdot1}$

2 D.V
3, 1, I.V.

$r_{12\cdot34}$

1 → D.V
2,3,4 → I.V

$r_{23\cdot14}$

2 — D.V
3,1,4 → I.V

$r_{31\cdot24}$

3 → D.V
1,2,4 → I.V.

LM4

Difference between Multiple Correlation and Partial Correlation.

Multiple Correlation

1) The Relationship between a Variable and a Combined Variable is Called as multiple Correlation

Ex :-

$R_{1.23}$

a Variable

Combined variable

$R_{4.12}$

a variable.

Combined Variable.

Partial Correlation

1) The Relationship between any two variables by neglecting the effect of Other variable is called a Partial Correlation.

Ex :-

$r_{23.1}$

two variable

Other variable.

Calculating of Coefficient of Partial Correlation

Problem: Given data

x_1	2	5	7	8	5
x_2	8	8	6	5	3
x_3	0	1	1	3	4

$n = 5$

Calculate the Co-efficient of Partial Correlation.

Sol: we know that

(1) $$\boxed{r_{12\cdot3} = \frac{r_{12} - r_{13}\cdot r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}}}$$

(2) $$\boxed{r_{13\cdot2} = \frac{r_{13} - r_{12}\, r_{23}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{23}^2}}}$$

(3) $$\boxed{r_{23\cdot1} = \frac{r_{23} - r_{12}\, r_{13}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{13}^2}}}$$

Partial Correlation :-

The relationship between any two variables by neglecting the effect of other variable is called as Partial Correlation.

x_1	x_2	x_3	x_1^2	x_2^2	x_3^2	x_1x_3	x_1x_2	x_2x_3
2	8	0	4	64	0	0	16	0
5	8	1	25	64	1	5	40	8
7	6	1	49	36	1	7	42	6
8	5	3	64	25	9	24	40	15
5	3	4	25	9	16	20	15	12
$\Sigma x_1 =$ 27	Σx_2 30	Σx_3 9	Σx_1^2 167	Σx_2^2 198	Σx_3^2 27	Σx_1x_3 56	$\Sigma x_1x_2 =$ 153	$\Sigma x_2x_3 =$ 41

1 ⟶ shows x_1

2 ⟶ shows x_2

3 ⟶ show x_3

we have

$\Sigma x_1 = 27$ $\qquad \Sigma x_1^2 = 167$ $\qquad \Sigma x_1x_3 = 56$

$\Sigma x_2 = 30$ $\qquad \Sigma x_2^2 = 198$ $\qquad \Sigma x_1x_2 = 153$

$\Sigma x_3 = 9$ $\qquad \Sigma x_3^2 = 27$ $\qquad \Sigma x_2x_3 = 41$

(1) $r_{12} = ?$

$r_{13} = ?$

$r_{23} = ?$

(i) $$r_{12\cdot3} = \frac{r_{12} - r_{13}\, r_{23}}{\sqrt{1-r_{13}^2}\,\sqrt{1-r_{23}^2}}$$

$r_{12} = ?$ $r_{13} = ?$ $r_{23} = ?$

$1 \longrightarrow x_1$
$2 \longrightarrow x_2$
$3 \longrightarrow x_3$

we Know that

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

$n = 5$

$$r_{12} = \frac{n\sum x_1 x_2 - (\sum x_1)(\sum x_2)}{\sqrt{\left[\left(n\sum x_1^2 - (\sum x_1)^2\right]\left[n\sum x_2^2 - (\sum x_2)^2\right]\right.}}$$

$$= \frac{(5)(153) - (27)(30)}{\sqrt{\left[\left[(5)(167) - (27)^2\right]\left[(5)(198) - (30)^2\right]\right.}}$$

$$= \frac{765 - 810}{\sqrt{\left[(835 - 729)(990 - 900)\right.}}$$

$$\boxed{r_{12} = -0.46}$$

$$= \frac{-45}{\sqrt{(106)(90)}} = \frac{-45}{\sqrt{9540}} = \frac{-45}{97.67} = \underline{\underline{-0.46}}$$

- $$r_{xy} = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[(n\sum x^2 - (\sum x)^2)][n\sum y^2 - (\sum y)^2]}}$$

$$r_{13} = \frac{n\sum x_1 x_3 - (\sum x_1)(\sum x_3)}{\sqrt{[(n\sum x_1^2 - (\sum x_1)^2][n\sum x_3^2 - (\sum x_3)^2]}}$$

$$= \frac{(5)(56) - (27)(9)}{\sqrt{[(5)(167) - (27)^2][(5)(27) - (9)^2]}}$$

$$= \frac{280 - 243}{\sqrt{(835 - 729)(135 - 81)}}$$

$$= \frac{37}{\sqrt{(106)(54)}} = \frac{37}{\sqrt{5724}}$$

$$= \frac{37}{75.6} = \underline{\underline{0.48}}$$

$$\boxed{r_{13} = 0.48}$$

$$r_{xy} = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[(n\sum x^2 - (\sum x)^2][(n\sum y^2 - (\sum y)^2]}}$$

$$r_{23} = \frac{n\sum x_2 x_3 - (\sum x_2)(\sum x_3)}{\sqrt{[(n\sum x_2^2 - (\sum x_2)^2][(n\sum x_3^2 - (\sum x_3)^2]}}$$

$$= \frac{(5)(41) - (30)(9)}{\sqrt{[(5)(198) - (30)^2][(5)(27) - (9)^2]}}$$

$$= \frac{205 - 270}{\sqrt{(990 - 900)(135 - 81)}}$$

$$= \frac{-65}{\sqrt{(90)(54)}} = \frac{-65}{\sqrt{4860}}$$

$$= \frac{-65}{69.71} = \underline{\underline{-0.93}}$$

$$\therefore \boxed{r_{23} = -0.93}$$

$\therefore$ we have

$$\boxed{r_{12} = -0.46}$$

$$\boxed{r_{13} = 0.48}$$

$$\boxed{r_{23} = -0.93}$$

(1) $$r_{12\cdot3} = \frac{r_{12} - r_{13}\cdot r_{23}}{\sqrt{1-r_{13}^2}\,\sqrt{1-r_{23}^2}}$$

$$= \frac{(-0.46)-(0.48)(-0.93)}{\sqrt{1-(0.48)^2}\,\sqrt{1-(-0.93)^2}}$$

$$= \frac{-0.46+0.4464}{\sqrt{(1-0.2304)(1-0.8679)}}$$

$$= \frac{-0.0136}{\sqrt{(0.7696)(0.1351)}}$$

$$= \frac{-0.0136}{\sqrt{0.1039}} = \frac{-0.0136}{0.322}$$

$$\boxed{r_{12\cdot3} = -0.004.}$$

(2) $$r_{13.2} = \frac{r_{13} - r_{12}\, r_{23}}{\sqrt{1 - r_{12}^2}\,\sqrt{1 - r_{23}^2}}$$

$$= \frac{0.48 - (-0.46)(-0.93)}{\sqrt{1 - (-0.46)^2}\,\sqrt{1 - (-0.93)^2}}$$

$$= \frac{0.48 - 0.4278}{\sqrt{1 - 0.2116}\,\sqrt{1 - 0.8649}}$$

$$= \frac{0.0522}{\sqrt{(0.7884)}\,(0.1351)}$$

$$= \frac{0.0522}{\sqrt{(0.7884)(0.1351)}}$$

$$= \frac{0.0522}{\sqrt{0.1065}} = \frac{0.0522}{0.3263}$$

$$= 0.159$$

$$\therefore \boxed{r_{13.2} = 0.159}$$

(3) $$r_{23.1} = \frac{r_{23} - r_{12}\, r_{13}}{\sqrt{1 - r_{12}^2}\,\sqrt{1 - r_{13}^2}}$$

$$= \frac{(-0.93) - (-0.46)(0.48)}{\sqrt{1 - (-0.46)^2}\,\sqrt{1 - (0.48)^2}}$$

$$= \frac{-0.93 + 0.2208}{\sqrt{1 - 0.2116}\,\sqrt{1 - 0.2304}}$$

$$= \frac{-0.7092}{\sqrt{(0.7884)(0.7696)}}$$

$$= \frac{-0.7092}{\sqrt{0.606}}$$

$$= \frac{-0.7092}{0.7789}$$

$$r_{23.1} = -0.9105$$

(4) In a Tri-variate distribution it is found that $r_{12} = 0.70$

$r_{13} = 0.61$ &

$r_{23} = -0.40$

Find value of $r_{23.1}$ & $r_{13.2}$

Sol: $r_{23.1} = \dfrac{r_{23} - r_{12} \times r_{13}}{\sqrt{(1-(r_{12})^2)(1-(r_{13})^2}}$

$= \dfrac{0.4 - 0.7 \times 0.61}{\sqrt{(1-(0.7)^2(1-(0.61)^2}}$

$= \dfrac{0.4 - 0.427}{\sqrt{(1-0.49)(1-0.3721)}}$

$= \dfrac{-0.027}{\sqrt{0.51 \times 0.6279}}$

$= \dfrac{-0.027}{0.566}$

$= -0.048$

$$\gamma_{13.2} = \frac{\gamma_{13} - \gamma_{12} \times \gamma_{23}}{\sqrt{(1-(r_{12})^2)(1-(r_{23})^2}}$$

$$= \frac{0.61 - 0.70 \times 0.4}{\sqrt{(1-(0.7)^2 (1-(0.4)^2}}$$

$$= \frac{0.61 - 0.28}{\sqrt{(1-0.49)(1-(0.16)}}$$

$$= \frac{0.33}{\sqrt{0.51 \times 0.84}}$$

$$= \frac{0.33}{\sqrt{0.4284}} = \frac{0.33}{0.65}$$

$$= 0.504$$

(6) From the following data.

x_1	3	5	6	8	12
x_2	16	10	7	4	3
x_3	9	7	5	4	8

Find all the Possible multiple Correlation Coefficient

ie :- $R_{1.23}$, $R_{2.13}$, $R_{3.12}$

Sol :-

x_1	x_2	x_3	x_1^2	x_2^2	x_3^2	x_1x_2	x_1x_3	x_2x_3
3	16	9	9	256	81	48	27	144
5	10	7	25	100	49	50	35	70
6	7	5	36	49	25	42	30	35
8	4	4	84	16	16	32	32	16
12	3	8	144	9	64	36	96	24
34	40	33	278	430	235	208	220	289

$\Sigma x_1 = 34$ $\quad \Sigma x_1^2 = 278$ $\quad \Sigma x_1 x_2 = 208$

$\Sigma x_2 = 40$ $\quad \Sigma x_2^2 = 430$ $\quad \Sigma x_1 x_3 = 220$

$\Sigma x_3 = 33$ $\quad \Sigma x_3^2 = 235$ $\quad \Sigma x_2 x_3 = 289$

(a) The Correlation between X_1 & X_2

$$r_{12} = \frac{n\sum x_1x_2 - (\sum x_1)(\sum x_2)}{\sqrt{n\sum x_1^2 - (\sum x_1)^2}\sqrt{n\sum x_2^2 - (\sum x_2)^2}}$$

$$= \frac{(5)(208) - (34)(40)}{\sqrt{(5)(278) - (34)^2}\sqrt{(5)(430) - (40)^2}}$$

$$= -0.892$$

(b) The Correlation between X_1 & X_3

$$r_{13} = \frac{n\sum x_1x_3 - (\sum x_1)(\sum x_3)}{\sqrt{n\sum x_1^2 - (\sum x_1)^2}\sqrt{n\sum x_3^2 - (\sum x_3)^2}}$$

$$= \frac{(5)(220) - (34)(33)}{\sqrt{(5)(278) - (34)^2}\sqrt{(5)(235) - (33)^2}}$$

$$= -0.1551$$

(c) The Correlation between X_2 & X_3

$$r_{23} = \frac{n\sum X_2X_3 - (\sum X_2)(\sum X_3)}{\sqrt{n\sum X_2^2 - (\sum X_2)^2}\sqrt{n\sum X_3^2 - (\sum X_3)^2}}$$

$$= \frac{(5)(289) - (40)(33)}{\sqrt{(5)(430) - (40)^2}\sqrt{(5)(235) - (33)^2}}$$

$$= \underline{\underline{0.5748}}$$

Now we have

(a) $$R_{1\cdot 23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{23}^2}}$$

$$= \sqrt{\frac{(-0.892)^2 + (-0.1551)^2 - 2(-0.892)(0.1551)(0.5748)}{1 - (0.5748)^2}}$$

$$= \sqrt{\frac{0.66067}{0.6696}}$$

$$= \underline{\underline{0.9933}}$$

(b) $$R_{2.13} = \sqrt{\frac{r_{12}^2 + r_{23}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{13}^2}}$$

$$= \sqrt{\frac{(-0.892)^2 + (0.5748)^2 - 2(-0.892)(0.1551)(0.5748)}{1 - (-0.1551)^2}}$$

$$= \sqrt{\frac{0.28510}{0.9759}}$$

$$= \underline{\underline{0.1475}}$$

(c) $$R_{3.12} = \sqrt{\frac{r_{13}^2 + r_{23}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{12}^2}}$$

$$= \sqrt{\frac{(-0.1551)^2 + (0.5748)^2 - 2(-0.892)(0.1551)(0.5748)}{1 - (-0.892)^2}}$$

$$= \sqrt{\frac{0.51349}{0.204336}}$$

$$= \underline{\underline{0.5852}}$$

(7) Find the value of $R_{1.23}$ & $R_{2.13}$ from the following given information

$b_{12} = 0.75 \quad b_{13} = 0.58 \quad b_{21} = 0.88$

$b_{23} = 0.53 \quad b_{31} = 1.68 \quad b_{32} = 1.30$

Sol: we know that

(1) $\gamma_{12} = \sqrt{b_{12} \times b_{21}} = \sqrt{(0.75)(0.88)} = \underline{\underline{0.81}}$

(2) $\gamma_{13} = \sqrt{b_{13} \times b_{31}} = \sqrt{(0.58)(1.68)} = \underline{\underline{0.99}}$

(3) $\gamma_{23} = \sqrt{b_{23} \times b_{32}} = \sqrt{(0.53)(1.30)} = \underline{\underline{0.83}}$

(a) $$R_{1.23} = \sqrt{\frac{\gamma_{12}^2 + \gamma_{13}^2 - 2\gamma_{12}\gamma_{13}\gamma_{23}}{1 - \gamma_{23}^2}}$$

$$= \sqrt{\frac{(0.81)^2 + (0.99)^2 - 2(0.81)(0.99)(0.83)}{1 - (0.83)^2}}$$

$$= \sqrt{\frac{0.305046}{0.3111}} = \underline{\underline{0.99}}$$

(b) $$R_{2.13} = \sqrt{\frac{r_{12}^2 + r_{23}^2 - 2 r_{12} r_{13} r_{23}}{1 - r_{13}^2}}$$

$$= \sqrt{\frac{(0.81)^2 + (0.83)^2 - 2(0.81)(0.99)(0.83)}{1 - (0.99)^2}}$$

$$= \sqrt{\frac{0.013846}{0.0199}} = \underline{\underline{0.83}}$$

Partial Correlation Coefficient

$$r_{12\cdot3} = \frac{r_{12} - r_{13}\, r_{23}}{\sqrt{(1-r_{13}^2)}\sqrt{1-r_{23}^2}}$$

$$r_{13\cdot2} = \frac{r_{13} - r_{12}\, r_{23}}{\sqrt{(1-r_{12}^2)}\sqrt{(1-r_{23}^2)}}$$

$$r_{23\cdot1} = \frac{r_{23} - r_{12}\, r_{13}}{\sqrt{(1-r_{12}^2)}\sqrt{(1-r_{13}^2)}}$$

Problem.

8) From the given data

x_1	3	5	6	8	12
x_2	16	10	7	4	3
x_3	90	72	54	42	30

Find all the Partial Correlation Coefficients.

Sol :-

x_1	x_2	x_3	x_1^2	x_2^2	x_3^2	x_1x_2	x_1x_3	x_2x_3
3	16	90	9	256	8100	48	270	1440
5	10	72	25	100	5184	50	360	720
6	7	54	36	49	2916	42	324	378
8	4	42	64	16	1764	32	336	168
12	3	30	144	9	90	36	360	90
34	40	288	278	430	18864	208	1650	2796

$$\gamma_{12} = \frac{n\sum x_1x_2 - (\sum x_1)(\sum x_2)}{\sqrt{n\sum x_1^2 - (\sum x_1)^2}\sqrt{n\sum x_2^2 - (\sum x_2)^2}}$$

$$= \frac{(5)(208) - (34)(40)}{\sqrt{(5)(278) - (34)^2}\sqrt{(5)(430) - (40)^2}} = -0.892$$

$$\gamma_{13} = \frac{n\sum x_1x_3 - (\sum x_1)(\sum x_3)}{\sqrt{n\sum x_1^2 - (\sum x_1)^2}\sqrt{n\sum x_3^2 - (\sum x_3)^2}}$$

$$= \frac{(5)(1650) - (34)(288)}{\sqrt{(5)(278) - (34)^2}\sqrt{(5)(18864) - (288)^2}} = -0.9451$$

$$r_{23} = \frac{n\sum X_2 X_3 - (\sum X_2)(\sum X_3)}{\sqrt{n\sum X_2^2 - (\sum X_2)^2}\sqrt{n\sum X_3^2 - (\sum X_3)^2}}$$

$$= \frac{(5)(2796) - (40)(288)}{\sqrt{(5)(430) - (40)^2}\sqrt{(5)(18864) - (288)^2}}$$

$$= 0.9835$$

Now we have to find Partial Correlation Coefficients:

$r_{12\cdot3}$, $r_{13\cdot2}$ & $r_{23\cdot1}$

a) $$r_{12\cdot3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1-r_{13}^2)}\sqrt{(1-r_{23}^2)}}$$

$$= \frac{-0.892 - (-0.945)(0.984)}{\sqrt{(1-(-0.945)^2}\sqrt{(1-(0.984)^2)}}$$

$$= \frac{0.03788}{0.05827}$$

$$= 0.6500$$

(b) $$\gamma_{13\cdot 2} = \frac{\gamma_{13} - \gamma_{12}\cdot\gamma_{23}}{\sqrt{(1-r_{12}^2)}\sqrt{(1-r_{23}^2)}}$$

$$= \frac{-0.945 - (-0.892)(0.9835)}{\sqrt{1-(-0.892)^2}\sqrt{1-(0.9835)^2}}$$

$$= \frac{-0.067718}{0.081770}$$

$$= \underline{\underline{-0.8281}}$$

(c) $$\gamma_{23\cdot 1} = \frac{\gamma_{23} - \gamma_{12}\cdot\gamma_{13}}{\sqrt{(1-\gamma_{12}^2)}\sqrt{(1-\gamma_{13}^2)}}$$

$$= \frac{0.9835 - (-0.892)(-0.945)}{\sqrt{1-(-0.892)^2}\sqrt{1-(-0.945)^2}}$$

$$= \frac{0.05551}{0.14783}$$

$$= \underline{\underline{0.3754}}$$

Optimization

Unit-5

(1)

Optimization :- Optimization Problem requires us to determine maximum and Minimum value of a function.

There are Two types of Optimization

(1) Constrained optimization

(2) UnConstrained optimization

(1) **UnConstrained Optimization** :-

There are two types of Problems

They are (1) Profit maximization

(2) Cost minimization

Practical Computational task of finding maxima (or) minima of a function of many variables

Method :-

Step (1) Find the derivative of a function with respect to x and y then put it equals to Zero to find the values of x and y Points which is Called as stationary points.

$$\frac{df}{dx} = 0 \longrightarrow (1)$$

$$\frac{df}{dy} = 0 \longrightarrow (2)$$

By solving (1) & (2) we have Points (x, y) $\longrightarrow$ stationary points.

Step-2 :- Second step is to find Second Order derivative of a function with respect to x and y and xy, then solve the equation $AC - B^2$.

w.r.to $x \longrightarrow A$ put
w.r.to $y \longrightarrow C$ put
w.r.to to $xy \longrightarrow B$ put

Step-3 :- If $AC - B^2 > 0$, then its the Case of extreme Points

$$\frac{\partial^2 f}{\partial x^2} = A$$

$$\frac{\partial^2 f}{\partial y^2} = C$$

$$\frac{\partial^2 f}{\partial x \partial y} = B$$

(1) $AC - B^2 > 0$ ✓ (extreme points)

(2) $AC - B^2 < 0$
(3) $AC - B^2 = 0$
$\Big\}$ No extreme ✗ Points.

Step-4 :- we Can find maxima (or) minima, according to the Value of A.

Step-5 :- Last step is to find the Maximum (or) Minimum value of that function through the extreme Value.

Problem

(2)

(1) Find the Extreme Value of

$f(x) = x^3 + y^3 - 6xy$ and determine

whether they are Maximum (or) Minimum.

Sol: given $f(x) = x^3 + y^3 - 6xy$

$$\frac{\partial f}{\partial x} = 3x^2 + 0 - 6y = 0$$

$$= \cancel{3(x^2} = 3x^2 - 6y = 0$$

$$= 3(x^2 - 2y) = 0$$

$$x^2 - 2y = 0$$

$$-2y = -x^2$$

$$\Rightarrow \boxed{y = \frac{x^2}{2}} \longrightarrow (1).$$

$$\frac{\partial f}{\partial y} = 3y^2 - 6x = 0$$

$$= 3(y^2 - 2x) = 0$$

$$= y^2 - 2x = 0 \longrightarrow (2)$$

By Solving (1) & (2)

Putting y in (2)

$$y^2 - 2x = 0$$

$$\left(\frac{x^2}{2}\right)^2 - 2x = 0$$

$$\frac{x^4}{4} - 2x = 0$$

$$= \frac{x^4 - 8x}{4} = 0$$

$$x^4 - 8x = 0$$

$$x(x^3 - 8) = 0$$

$\boxed{x=0}$ $\quad x^3 - 8 = 0$

$$x^3 = 8$$

$$x = \sqrt[3]{8} = 2^3$$

$\boxed{x=2}$ $\quad x^3 = 2^3$

$\boxed{x=2}$

$\therefore$ Putting x in ①

$$y = \frac{x^2}{2}$$

when $\boxed{x=0} \Rightarrow y = \frac{0}{2} = 0$

when $\boxed{x=2} \Rightarrow y = \frac{2^2}{2} = \frac{4}{2} = 2$

$\therefore$ Stationary Points are $(0,0)\ (2,2)$.

Second Order Condition :-

$$\frac{\partial f}{\partial x} = 3x^2 - 6y \qquad \frac{\partial f}{\partial y} = 3y^2 - 6x$$

$$A = \frac{\partial^2 f}{\partial x^2} = 6x$$

$$B = \frac{\partial^2 f}{\partial x \partial y} = -6$$

$$C = \frac{\partial^2 f}{\partial y^2} = 6y$$

3

$A = 6x$, $B = -6$, $C = 6y$

a) At Point $(\underset{x}{0}, \underset{y}{0})$:-

$A = 6(0) = 0 \Rightarrow \boxed{A = 0}$
$B = -6 \Rightarrow \boxed{B = -6}$
$C = 6(0) = 0 \Rightarrow \boxed{C = 0}$

$AC - B^2$

$\Rightarrow (0)(0) - (-6)^2 = -(36) = -36 < 0$

$\therefore$ No Extreme Point

b) At Point $(\underset{x}{2}, \underset{y}{2})$:-

$A = 6(2) = 12 \Rightarrow \boxed{A = 12}$
$B = -6 \Rightarrow \boxed{B = -6}$
$C = 6(2) = 12 \Rightarrow \boxed{C = 12}$

$AC - B^2$

$\Rightarrow (12)(12) - (-6)^2 = 144 - 36 = 108 > 0$

It is an Extreme Point.

$A = 12 > 0$, Positive minima Point.

Extreme value at $f(\underset{x}{2}, \underset{y}{2})$

$f(x,y) = x^3 + y^3 - 6xy$
$= 2^3 + 2^3 - 6(2)(2)$
$= 8 + 8 - 6(4) = 16 - 24 = -8$

So, -8 is Minimum value at $x = 2$ & $y = 2$
So, function is Minimum

Constrained Optimization

→ The general Constrained Optimization task is to Maximize (or) Minimize a function $f(x)$ by varying x, given certain Constraints on x.

→ for example :-

Find Minimum of

$$f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + x_3^2$$

where $\|x\|_2 \geq 1$

→ Very Common to encounter this in engineering Practice

for example :- designing the fastest Vehicle with a Constraint on fuel efficiency.

→ All Constraints Can be Converted to two types of Constraints

→ Equality Constraints :-

Ex :- Minimize $f(x_1, x_2, x_3)$

Subject to $x_1 + x_2 + x_3 = 1 \longrightarrow (1)$

→ Inequality Constraints :-

Ex :- Minimize $f(x_1, x_2, x_3)$

Subject to $x_1 + x_2 + x_3 < 1 \longrightarrow (2)$

→ Canonical form :-

All optimization Problems Can be written as

Minimize $f(x)$ subject to Constraint that $x \in S$ → feasible points

$$S = \{x / \forall i,\ g^{(i)}(x) = 0 \text{ and } \forall j,\ h^{(j)}(x) \leq 0\}$$

multiple equality Constraints . multiple Inequality Constraints

from eqn ①

⟹ Equality Constraints Can be written as

equality Constraint

$$g(x) = x_1 + x_2 + x_3 - 1 = 0$$

from eqn ② Inequality Constraints Can be written as

$$h(x) = x_1 + x_2 + x_3 - 1 < 0$$

feasibility set is a Combination of equality Constraint & Inequality Constraints.

- **Generalized Lagrange function** :-

→ The Constrained Optimization Problem requires us to minimize the function $f(x)$ while ensuring that the point discovered belongs to the feasible set.

→ There are several techniques that acheive this but it is, in general, a difficult Problem.

→ A very common approach is to define a new function called the generalized Lagrangian.

$$\underbrace{L}_{\text{Lagrangian}}(x, \underbrace{\lambda, \alpha}_{\substack{\text{two vectors} \\ (\text{variables } \lambda, \alpha)}}) = \underbrace{f(x)}_{\text{original function}} + \sum_i \lambda_i g^{i}(x) + \sum_j \alpha_j h^{(j)}(x)$$

→ Then the Constrained minimum is given by

$$\min_{x \in S} f(x) = \min_{x} \max_{\lambda} \max_{\alpha, \alpha \geq 0} L(x, \lambda, \alpha)$$

→ we will the Proof and details of this when we come to later weeks.

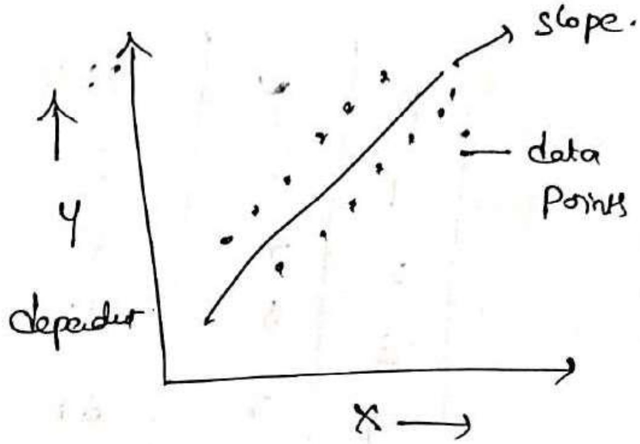
Linear Regression

→ Dependent Variable is Continious in nature.

Simple Linear Regression

$$\boxed{y = \alpha_0 + \alpha_1 x_1}$$

$y = c + mx$



Linear in nature.

x_1 is '1' Independent Variable
y is dependent Variable
α values are Co-efficients of Regression

This Can be easily written as

$$\boxed{y = c + mx}$$

x is independent Variable
m is a slope.
c is Intercept
Y is dependent Variable.

multiple linear regression

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m$$

α_i = Regression Co-eff
x_i = Independent Variables
Y = dependent Variables

$\alpha_3 \geq \alpha_4$

ex:- $y = 0.9 + 1.2x_1 + 2x_2 + 4x_3 + 1x_4$

α_3 is more
X_3 is more

(1) Given $x = 1, 2, 3, 4$. $\boxed{Y = bx + a}$

$y = 3, 4, 5, 7$

x	y	xy	x^2
1	3	3	1
2	4	8	4
3	5	15	9
4	7	28	16
$\Sigma x = 10$	$\Sigma y = 19$	Σxy 54	$\Sigma x^2 = 30$.

$$\boxed{a = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}}$$

$$a = \frac{(19)(30) - (10)(54)}{(4)(30) - (100)}$$

a & b are Unknown.

$$= \frac{570 - 540}{120 - 100} = \frac{30}{20} = \frac{3}{2} = \underline{1.5}$$

$\therefore \boxed{a = 1.5}$

$\therefore \boxed{a = 1.5}$ $\boxed{b = 1.3}$

$$b = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$\therefore \boxed{Y = 1.3x + 1.5}$

$$b = \frac{(4)(54) - (10)(19)}{(4)(30) - 100} = \frac{216 - 190}{120 - 100} = \frac{26}{20} = \frac{13}{10}$$

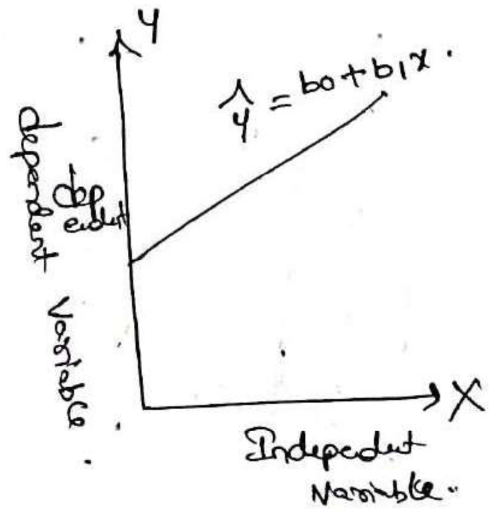
$= \underline{1.3}$

Linear Regression Analysis

→ This analysis is Used in Understanding the relationship between two (or) more Variables (Multiple regression.)

→ when in Case of understanding two variables One is Independent Variable (Input) and the Other variable is dependent variable (Predicted variable)

$$\boxed{\hat{y} = b_0 + b_1 x}$$



→ In this analysis we try to find a straight line which fits most of the Observations

b_0 – y-Intercept
b_1 – slope

→ This line is Called as <u>Linear Regression Line</u> and it is Obtained by <u>Least Square method</u>

→ when $x \uparrow y \uparrow$ slope is +ve $(\hat{y} = b_0 + b_1 x)$
→ when $x \uparrow y \downarrow$ slope is −ve $(\hat{y} = b_0 - b_1 x)$

Problem.

x	1	2	3	4	5
y	2	4	5	4	5

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y})$	$\hat{y}$	$\hat{y}-\bar{y}$	$(\hat{y}-\bar{y})^2$
1	2	-2	-2	4	4	2.8	0.8	0.64
2	4	-1	0	1	0	3.4	0.6	0.36
3	5	0	1	0	0	4	-1	1
4	4	1	0	1	0	4.6	0.6	0.36
5	5	2	1	4	2	5.2	0.2	0.04
15	20			10	6			2.4

$\bar{x} = \frac{15}{5} = \underline{\underline{3}}$ $\quad \therefore \bar{x} = 3$

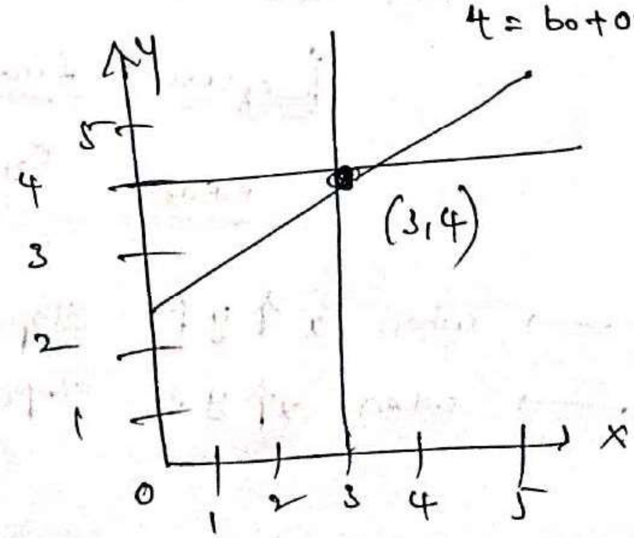
$\bar{y} = \frac{20}{5} = \underline{\underline{4}}$ $\quad \bar{y} = 4$

$\hat{y} = b_0 + b_1 x$

$4 = b_0 + 0.6 \times 3$

$$b_1 = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}$$

$$= \frac{6}{10} = \underline{\underline{0.6}}$$



$$\therefore \; b_1 = \frac{\Sigma (x-\bar{x})(y-\bar{y})}{\Sigma (x-\bar{x})^2} = \frac{6}{10} = \underline{\underline{0.6}}$$

b_0 is Calculated Using the mean Coordinate $(3, 4)$

$$y' = b_0 + b_1 x$$

$$4 = b_0 + (0.6)3$$

$$b_0 = 4 - (0.6)3$$

$$\boxed{b_0 = 2.2}$$

$$\therefore \; \boxed{\hat{y} = 2.2 + 0.6x}$$ (Regression Line)

Standard error

$$= \sqrt{\frac{\Sigma (\hat{y} - \bar{y})^2}{n-2}} = \sqrt{\frac{2.4}{5-2}} = \underline{\underline{0.89}}$$

Calculates the difference between actual + estimated.

Assumptions of Linear Regression :-

1) Linear Relation
2) very low / No multi Collinearity
3) Heterostochestaycity
4) No Auto Correlation of errors.
5) Normal distribution of errors.
6) All the Observations are Independent to each other

Linear Regression

1) A simple example of a Regression equation to Predict the Glucose level given the age

Subject	Age(X)	Glucose level (Y)
1	43	99
2	21	65
3	25	79
4	42	75
5	57	87
6	59	81
7	55	?

X is Independent
Y is dependent
(Here 6 data points are given & 7th one is asked)

Sol:-

The Simple Linear Regression Equation Provides an Estimate of the Population Regression Line.

$$\boxed{\hat{Y}_i = b_0 + b_1 X_i}$$

Value of X for Observation i

Estimated (or Predicted) Y value for ~~Obse~~ Observation i

Estimate of the Regression Intercept

Estimate of the Regression slope.

Y is dependent
X is Independent.

A simple example of a Regression equation to Predict the glucose level given the age.

$$b_0 = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$b_1 = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

Subject	Age (X)	Glucose level (Y)	XY	X^2	Y^2
1	43	99	4257	1849	9801
2	21	65	1365	441	4225
3	25	79	1975	625	6241
4	42	75	3150	1764	5625
5	57	87	4959	3249	7569
6	59	81	4779	3481	6561
	ΣX = 247	ΣY = 486	ΣXY = 20485	ΣX^2 = 11409	ΣY^2 = 40022

$$\boxed{b_0 = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}}$$

$$b_0 = \frac{(486)(11409) - (247)(20485)}{6(11409) - (247)^2}$$

$$b_0 = \frac{4848979}{7445} = \underline{\underline{65.14}}$$

$$\therefore \boxed{b_0 = 65.14}$$

$$\boxed{b_1 = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}}$$

$$b_1 = \frac{6(20485) - (247)(486)}{6(11409) - (247)^2}$$

$$b_1 = \frac{2868}{7445} = \underline{\underline{0.385335}}$$

$$\therefore \boxed{b_1 = 0.385335}$$

$$\boxed{\hat{y} = b_0 + b_1 x}$$

$$\therefore \boxed{\hat{y} = b_0 + b_1 x}$$

$$\boxed{\hat{y} = 65.14 + 0.385225\,x}$$

The value of Y for given value of X = 55

$$\hat{y} = 65.14 + (0.385225)(55)$$

$$\boxed{\hat{y} = 86.327}$$

Hence the glucose level for the given age 55 is <u>86.327</u>.

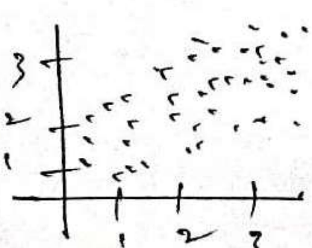
→ what is the Problem of Heteroscedasticity?

<u>Sol</u>: The Problem of Heteroscedasticity refers to a situation when the residuals in a Regression <u>do not have Uniform Variance</u>.

→ Arise when Variation is Uneven across Observations

→ Tends to give inefficient regression results

— Incorrect standard error.



→ Heteroscedasticity.

LA4

Unit-5

→ In this Chapter we shall Concern ourselves with the Classical theory of Optimization.

→ This theory deals with the use of differential Calculus to determine the Points of Maxima & Minima for both Unconstrained and Constrained Continuous functions.

→ In this Chapter the topics include the development of necessary and sufficient Conditions for locating extreme points for Unconstrained Problems

→ The treatement of the Constrained Problems Using the Lagrangian methods and the development of the Kuhn-Tucker Conditions for the general Problem with Inequality Constraints.

Un-Constrained Problems of Maxima & Minima.

→ Here we shall discuss the Problem of determining the extreme points (the Points of Maxima & Minima) of an Unconstrained type of Continious functions.

Mathematically

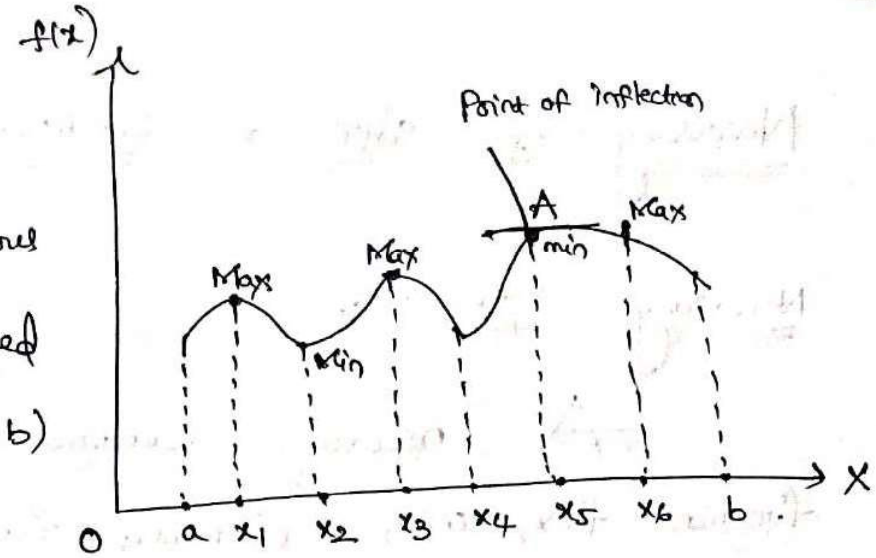
→ A function $f(x)$ has a Maximum at a Point 'x_0' If

$$\boxed{f(x_0+h) - f(x_0) < 0}$$

→ Illy a function $f(x)$ has a Minimum at a point 'x_0' If

$$\boxed{f(x_0+h) - f(x_0) > 0}$$

Consider a Continious function $f(x)$ defined on in interval (a, b)



Here the points x_1, x_2, x_3, x_4 & x_6 (not x_5) represent all the Points of Maxima & Minima.

(Called the stationary (or) Critical points) of $f(x)$

These includes $x_1, x_3,$ & x_6 as Points of Maxima &

x_2 & x_4 as Points of Minima.

Global (absolute) Maximum :-

Since $f(x_6) = \max\{f(x_1), f(x_3), f(x_6)\}$,

$f(x_6)$ is Called a global (or) absolute maximum.

Local (relative) maxima :-

On the Other hand $f(x_1)$ & $f(x_3)$ are Called local (or) relative maxima.

Illy $f(x_4)$ is a local Minimum while,

$f(x_2)$ is a global minimum.

→ It should be noted that the point 'A' Corresponding to $f(x_5)$ is Called Point of Inflection.

Necessary & Sufficient Conditions for Optima.

Necessary Condition :-

A necessary Condition for a Continious function $f(x)$ with Continious first and second Partial derivatives to have an extreme Point at 'x_0' is that each first Partial derivative of $f(x)$ evaluate at x_0, Vanish that is

$$\boxed{\nabla f(x_0) = 0}$$

where $\nabla \equiv \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \cdots \frac{\partial f}{\partial x_n}\right)$ is the <u>gradient vector</u>

Sufficient Condition :-

A sufficient Condition for a Stationary Point 'x_0' to be an extreme Point is that the <u>Hession matrix H</u> evaluated at 'x_0' is

(1) Negative - definite when 'x_0' is a <u>Maximum Point</u> &

(2) Positive - definite when 'x_0' is a <u>Minimum Point</u>.

Proof for Conditions :-

1) Proof for necessary Condition :-

By Taylors theorem, for $0 < \theta < 1$

$$f(x_0+h) - f(x_0) = \nabla f(x_0)h + \frac{1}{2} h' H h \Big|_{x_0+\theta h} \longrightarrow (1)$$

where $h = (h_1 h_2 \cdots h_j \cdots h_n)'$ &

$|h_j|$ is small enough $\forall$ $j = 1, 2 \cdots n$.

for small $|h_j|$ the remainder term $\frac{1}{2}(h'Hh)$ is of Order h_j^2 & hence it will tend to Zero as $h_j \to 0$

$$f(x_0+h) - f(x_0) = \nabla f(x_0)h + O(h_j^2) \longrightarrow (2)$$

$$\nabla f(x_0)h = \left[h_1 \frac{\partial f(x)}{\partial x_1} + h_2 \frac{\partial f(x)}{\partial x_2} + \cdots + h_p \frac{\partial f(x)}{\partial x_p} + \cdots + h_n \frac{\partial f(x)}{\partial x_n}\right]_{x = x_0}$$

Suppose that x_0 is an extreme point, now we shall prove the theorem by Contradiction.

If possible, let us suppose that one of the partial derivatives, say p^{th}, does not vanish,

ie:- $\frac{\partial f(x_0)}{\partial x_p} \neq 0$

then eqn (2) becomes $f(x_0+h) - f(x_0) = h_p \frac{\partial f(x_0)}{\partial x_p} \longrightarrow (3)$

Since $\frac{\partial f(x_0)}{\partial x_p} \neq 0$, either

$$\frac{\partial f(x_0)}{\partial x_p} < 0 \quad (\text{or}) \quad \frac{\partial f(x_0)}{\partial x_p} > 0$$

Now suppose $\frac{\partial f(x_0)}{\partial x_p} > 0$ then $f(x_0+h) - f(x_0)$ will have the same sign as h_p

ie: (i) $f(x_0+h) - f(x_0) > 0$ when $h_p > 0$ &

(ii) $f(x_0+h) - f(x_0) < 0$ when $h_p < 0$.

This contradicts the assumption that 'x_0' is an extreme point.

The argument when $\frac{\partial f(x_0)}{\partial x_p} < 0$ is similar to the given above.

Thus we may conclude that when any of the partial derivatives are not identically equal to zero at 'x_0', the point 'x_0' is not an extreme point.

Thus, it follows that for 'x_0' to be an extreme point it is necessary that

$$\boxed{\nabla f(x_0) = 0}$$

This completes the proof of the theorem.

2) Proof for Sufficient Condition :-

Proof :- By Taylor's theorem for $0<\theta<1$

we have

$$f(x_0+h) - f(x_0) = \nabla f(x_0)h + \frac{1}{2}h' Hh \Big|_{x_0+\theta h}.$$

Since 'x_0' is a stationary Point, then by Preceding theorem (necessary Condition theorem) we have

$$\boxed{\nabla f(x_0) = 0}$$

Thus $f(x_0+h) - f(x_0) = \frac{1}{2}h' Hh \Big|_{x_0+\theta h}$

Let x_0 be a Maximum Point then by definition

$$f(x_0+h) < f(x_0)$$

for all non-null h.

This implies that for x_0 ~~be~~ to be a Maximum.

$$\frac{1}{2}h' Hh \Big|_{x_0+\theta h} < 0 \quad \text{(or)}$$

$$h' Hh \Big|_{x_0+\theta h} < 0 \longrightarrow (1)$$

writing the quadratic form $h'Hh$ in expanded form

we have $\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \frac{\partial^2 f(x)}{\partial x_i \partial x_j}\Bigg|_{x=x_0+\theta h} < 0$

However, since the second Partial derivative 5

$\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$ is Continuous in the neighbourhood of 'x_0'

$\left.\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right|_{x=x_0}$ will have the Same Sign as

$$\left|\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right|_{x=x_0+\theta h}.$$

Consequently $h'Hh$ must yield the same sign when evaluated at both x_0 & $x_0+\theta h$.

Thus from eqn (1) we have

$$\left. h'Hh \right|_{x=x_0} < 0$$

Since $\left. h'Hh \right|_{x=x_0}$ defines a quadratic form, this expression (and hence $\left. h'Hh \right|_{x=x_0+\theta h}$) is negative $\Longleftrightarrow$ the Hessian matrix H is negative-definite at x_0

This Completes the Proff for maximization Case.

A similar Proof Can be established for minimization Case to show the Corresponding Hessian matrix H is Positive definite at x_0

Problem

(1) Find the Maximum (or) Minimum of the function

$$f(x) = x_1^2 + x_2^2 + x_3^2 - 4x_1 - 8x_2 - 12x_3 + 56$$

Sol:- Applying the necessary Condition

$$\nabla f(x_0) = 0 \quad \text{(or)}$$

$$\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right) f(x) = (0, 0, 0)$$

this gives

$$\frac{\partial f}{\partial x_1} = 2x_1 - 4 = 0$$

$$\frac{\partial f}{\partial x_2} = 2x_2 - 8 = 0$$

$$\frac{\partial f}{\partial x_3} = 2x_3 - 12 = 0$$

The Solution of these Simultaneous equations is given by $x_0 = (2, 4, 6)$ which is the only point that satisfies the necessary Condition.

Now by checking the sufficient Condition, we must determine whether this point is a Maximum or minimum

The Hessian matrix, evaluated at $(2,4,6)$ is given by

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$\therefore$ The Principal minor determinants of H

$$|2|, \quad \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} \quad \& \quad \begin{vmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix}$$

have the values 2, 4 & 8 respectively

Thus each of the Principal minor determinant is Positive

Hence H is Positive - definite

$\therefore$ the Point $(2,4,6)$ yields a ~~maxi~~ minimum of $f(x)$

Optimization View of Machine Learning

i) why do we need Optimization for Machine learning ?

A) Optimization is an Advanced topic, we still can build many models, experiments etc

a) what do you learn in optimization ?

1) Calculus :- Understanding different types of functions how to find maxima & Minima of functions

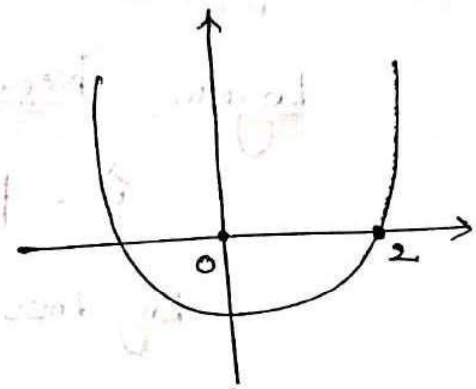
ex :- $y = f(x) = x^2 - 2x$

minima :- $\frac{dy}{dx} = f'(x) = 0$

$= 2x - 2 = 0$

$\Rightarrow 2x = 2$

$\Rightarrow \boxed{x = 1}$



→ functions Could get more Complicated (with "loss functions" of ML Models)

usually no Closed form solutions.

Optimization : Can we Come up with algorithms to find maxima / minima of these functions ? in an efficient and effective way.

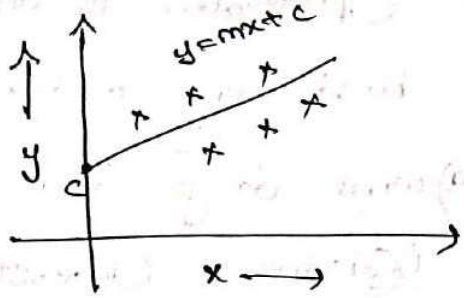
(1) **Understanding Loss functions :-**

most ML algorithms are driven by

→ writing an Objective function / loss function

→ finding the best Parameters that minimize the Loss .

Ex :- Linear Regression :-



$$\boxed{y = mx + c}$$

straight line

$$\sum_{i=1}^{n} [y_i - (mx_i + c)]^2$$

Logistic Regression :-

$$\hat{y} = p(y=1) = \sigma(\omega^T x) = \frac{1}{1+e^{-\omega^T x}}$$

$$\text{Log Loss} = \sum_{i=1}^{N} -y_i \log(\hat{y}_i) - (1-y_i)\log(1-\hat{y}_i)$$

(2) **Understanding what Solver to use :-**

→ Do you want to use

RMSProp vs ADAM vs Momentum ?

→ Do you want to use

Batch Gradient Descent vs Stochastic Gradient Descent vs Mini-Batch Gradient Descent ?

(3) Writing Loss functions :-

You have a new Problem that does not fit in the usual Setting

Easier in the Context of deep Learning :-

You Can write a Custom loss function, leave the Solving to underlying platform.

Optimization : Learn to write Loss functions.

→ Optimization gives the understanding of how you Can write Loss function.

4) Writing updates / solvers for an Optimization Problem :-

Understand writing the updates for different Kinds of solvers

Gradient Descent :

$$W_j^{new} = W_j - \alpha \frac{\partial}{\partial w_j} \text{loss}(w)$$

Summarize :- Optimization for Machine Learning

Why Optimization :-

1) Understanding Loss functions
2) Picking the right Solver
3) writing new Loss functions
4) Implementing solver for Custom loss function.

Non-Linear Programming.

Un Constrained Optimization Techniques :-

(i) Direct Search methods

(ii) Descent methods (or) Gradient methods.

~~(i)~~ Steepest descent (Cauchy) method

(ii) Newton's method

(iii) Fletcher Reeves method

(iv) Marquardt methods

(v) Quasi-Newton Methods.

(i) Steepest descent (Cauchy) method :-

Procedure :-

(1) Start with the arbitrary initial point X_1
set the iteration number $i=1$

(2) Find the search direction S_i as

$$S_i = -\nabla f = -\nabla f(X_i)$$

(3) Find the optimal step length λ_i^* in the direction S_i set

$$X_{i+1} = X_i + \lambda^* S_i = X_i - \lambda_i^* \nabla f_i$$

(4) Test X_{i+1} for Optimality.
If X_{i+1} is Optimum, stop, Otherwise go to step-5

(5) Set the new iteration number $i=i+1$ and go to step-2

This method looks to be a very effective Unconstrained Optimization technique.

But since steepest descent direction is a local property, the method is not very effective in most of the problems.

Problems

(1) Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ and starting from the point $X_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ using descent method.

Sol :- Iteration -1 :-

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}$$ Diff Partial w.r.to x_1 & x_2

$$\nabla f_1 = \nabla f(x_1) = \begin{bmatrix} +1 \\ -1 \end{bmatrix} \longrightarrow \begin{pmatrix} x_1 = 0 \\ x_2 = 0 \end{pmatrix}$$ Substitute.

$$\therefore S_1 = -\nabla f_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

For X_2, we need λ^* (the Optimal step length)

so we minimize

$$f(x_1 + \lambda_1 S_1) = f(-\lambda_1, \lambda_1) = \lambda_1^2 - 2\lambda_1$$

w.r.t from λ_1

Since $\frac{\partial f}{\partial \lambda_1} = 0$ gives $\lambda_1^* = 1$,

we get $X_2 = X_1 + \lambda_1^* S_1$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} + 1\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\nabla f_2 = \nabla f(x_2) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$\Longrightarrow$ X_2 is <u>not</u> Optimal.

<u>Iteration-2</u> :-

$$S_2 = -\nabla f_2 = (1, 1)$$

$$f(x_2 + \lambda_2 S_2) = f(-1 + \lambda_2, \ 1 + \lambda_2)$$

$$= 5\lambda_2^2 - 2\lambda_2 - 1$$

$$\frac{\partial f}{\partial \lambda_2} = 0 \Longrightarrow \lambda_2^* = \frac{1}{3}$$

$$\Longrightarrow X_3 = X_2 + \lambda_2^* S_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \frac{1}{5}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -0.8 \\ 1.2 \end{pmatrix}$$

$$\nabla f_3 = \nabla f(x_3) = \begin{pmatrix} 0.2 \\ -0.2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$\Longrightarrow$ X_3 is <u>not</u> Optimal

<u>Iteration -3</u> :- $S_3 = -\nabla f_3 = \begin{pmatrix} -0.2 \\ 0.2 \end{pmatrix}$

$$f(x_3 + \lambda_3 S_3) = f(-0.8 - 0.2\lambda_3,\ 1.2 + 0.2\lambda_3)$$

$$= 0.04\lambda_3^2 - 0.08\lambda_3 - 1.2$$

$$\frac{\partial f}{\partial \lambda_3} = 0 \implies \lambda_3^* = 1.0$$

$$\implies X_4 = X_3 + \lambda_3^* S_3 = \begin{pmatrix} -1.0 \\ 1.4 \end{pmatrix}$$

$$\nabla f_4 = \nabla f(X_4) = \begin{pmatrix} -0.2 \\ -0.2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$\implies$ X_4 is not Optimal

Iteration - 4

$$S_4 = -\nabla f_4 = \begin{pmatrix} 0.2 \\ 0.2 \end{pmatrix}$$

$$f(x_4 + \lambda_4 S_4) = f(-1 + 0.2\lambda_4,\ 1.4 + 0.2\lambda_4)$$

and so on.

we Continue the Process Until the Optimum Point

$$X^* = \begin{pmatrix} -1.0 \\ 1.5 \end{pmatrix}$$ is Found

(2) Newton's Method :-

$$f(x) = f(x_i) + \nabla f_i^T (x - x_i) + \frac{1}{2} (x - x_i)^T [J_i] (x - x_i) \longrightarrow (1)$$

where $[J_i] = (J)\big|_{x_i}$ is the Hessian matrix (matrix of second Order Partial derivatives) of f evaluated at the Point x_i

for the Minimum of $f(x)$,

evaluate the Partial derivatives of $f(x) = 0$ $f(x)$ to Zero

ie:- $$\frac{\partial f(x)}{\partial x_j} = 0 \quad , \quad j = 1, 2, 3 \ldots n \longrightarrow (2)$$

from eqns (1) & (2) we get

$$\nabla f = \nabla f_i + [J_i][x - x_i] = 0 \longrightarrow (3)$$

If $[J_i]$ is non-singular the above equation (3) Can be used to improve the approximation of x as x_{i+1} given by

$$x_{i+1} = x_i - [J_i]^{-1} \nabla f_i \longrightarrow (4)$$

Problem.

1) Minimize

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$$ by

taking the starting Point $X_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ Use Newtons Method.

Sol :-

$$J_1 = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}_{X_1} = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}$$

$$\therefore J_1^{-1} = \frac{1}{4} \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}$$

$$\nabla f_1 = g_i = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 3x_2 \end{bmatrix}_{(0\ 0)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

by eqn (4) $(x_{i+1} = x_i - [J_i]^{-1} \nabla f_i)$

$$x_2 = x_1 - J^{-1} \nabla f_1$$

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

$$= \begin{bmatrix} -1 \\ 3/2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1.5 \end{bmatrix}$$

$$\nabla f_2 = g_2 = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}_{X_2} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}_{X_2} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Since $\nabla f_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow X_2$ is Optimal Point

- ## Constrained Optimization :-

The general Constrained Optimization task is to maximize (or) Minimize a function $f(x)$ by varying x given certain constraint on X.

→ for example,
find minimum of $f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + x_3^2$
where $\|x\|_2 \geq 1$

→ Very Common to encounter this in engineering Practice.

→ For example, designing the fastest vehicle with a Constraint on fuel efficiency.

→ All Constraints Can be Converted to two types of Constraints
1) Equality Constraints
2) Inequality Constraints

1) **Equality Constraints** :-

Ex :- Minimize $f(x_1, x_2, x_3)$
Subject to $x_1 + x_2 + x_3 = 1 \Rightarrow x_1 + x_2 + x_3 - 1 = 0$
$\Rightarrow g(x) = x_1 + x_2 + x_3 - 1 = 0$.

2) **Inequality Constraints** :-
Minimize $f(x_1, x_2, x_3)$
Subject to $x_1 + x_2 + x_3 < 1 \Rightarrow h(x) = x_1 + x_2 + x_3 - 1 < 0$

Canonical form :-

All optimization problems can be written as

$$S = \{ x \,/\, \forall_i ,\ g^{(i)}(x) = 0 \ \& \ \forall_j \ h^{(j)}(x) < 0 \}$$

minimize $f(x)$ subject to the constraints that $x \in S$ is the feasible point.

Generalized Lagrange function :-

→ The constrained optimization problem requires us to minimize the function $f(x)$, while ensuring that the point discovered belongs to the feasible state

→ There are several techniques that achieve this but it is in general, a difficult problem

→ A very common approach is to define a new function called the generalized Lagrangian

$$L(x, \lambda, \alpha) = f(x) + \sum_i \lambda_i g^i(x) + \sum_j a_j h^{(j)}(x)$$

where

$L(x, \lambda, \alpha)$ Lagrangian

$f(x)$ = given function

→ Then the constrained minimum is given by

$$\min_{x \in S} f(x) = \min_{x} \max_{\lambda} \max_{\alpha > 0} L(x, \lambda, \alpha)$$

- **KKT Conditions / necessary and sufficient Condition for Optima :-**

In Mathematical Optimization, the Karush – Kuhn - Tucker (KKT) Conditions, also known as the Kuhn - Tucker Conditions, are first derivative tests (sometimes called First Order necessary Conditions) for a solution in non-linear Programming to be Optimal, Provided that some regularity Conditions are Satisfied.

The Necessary and Sufficient Conditions for solving the Non - Linear Programming Problem with Inequality Constraints

→ We know that when Non - Linear Programming Problem with equality Constraints

Maximize / Minimize $f(x)$ such that $g_i(x) = b_i$

⟹ We can solve the Problem with Lagrangian multiplier method.

→ Now, Consider the Non - Linear Programming Problem with Inequality Constraints

Maximize / Minimize $f(x)$ such that $g_i(x) \leq b_i$

(or) $g_i(x) \geq b_i$

$\Rightarrow$ We can solve the Problem with KKT Method.

Conditions :-

Maximization Problem :-

Consider the NLPP maximize $f(x)$ such that

$$g_i(x) \leq b_i$$

Convert each ith inequality constraints into equations by adding the non-negative slack variables s_i^2

Slack Variables :- $x_1 + x_2 \leq 1$

$$\Rightarrow x_1 + x_2 + s_1 = 1$$

$$\therefore g_i(x) + s_i^2 = b_i$$

Consider $h_i(x) = g_i(x) + s_i^2 - b_i = 0 \longrightarrow (1)$

Thus given NLPP reduces to maximize $f(x)$

such that $h_i(x) = 0$

$\therefore$ Now equality constraint so we can use Lagrangian method.

Formulate the Lagrangian function as

$$L(x_i, s, \lambda) = f(x) - \sum_i \lambda_i h_i(x)$$

$$= f(x) - \sum_i \lambda_i (g_i(x) + s_i^2 - b_i)$$

$$(\because \text{from } (1))$$

- The necessary Conditions for stationary Points are

$$\frac{\partial L}{\partial x} = 0 \Longrightarrow \frac{\partial f}{\partial x} - \sum_i \lambda_i \frac{\partial g_i}{\partial x} = 0 \longrightarrow (2)$$

$$\frac{\partial L}{\partial \lambda_i} = 0 \Longrightarrow g_i(x) + S_i^2 - b_i = 0 \longrightarrow (3)$$

$$S_i^2 = b_i - g_i(x)$$

$$\frac{\partial L}{\partial S_i} = 0 \Longrightarrow -2\lambda_i S_i = 0 \longrightarrow (4)$$

solve (2) (3) & (4) we get Stationary Points

multiply equations (4) by S_i & get

$$\lambda_i S_i^2 = 0$$

$$\Longrightarrow \lambda_i (b_i - g_i(x)) = 0$$

$$\underline{\lambda_i = 0} \quad (\text{or}) \quad b_i - g_i(x) = 0$$

$$\Longrightarrow \underline{b_i = g_i(x)}$$

Since λ_i measures the rate of variance of f wrto b_i

$$\underline{\text{ie}} : \frac{\partial f}{\partial b_i} = \lambda_i$$

from equation (4) we have either $\boxed{\lambda_i = 0}$ (or) $\boxed{S_i = 0}$ (or) both Vanish at Optimal Conditions.

Case-1 :- when $S_i \neq 0$

It means Constraint is Satisfied as strict Inequality $(\because S_i \lambda_i = 0)$

If we relaxed the Constraint (make b_i Larger) the stationary point will not be affected

$\therefore \boxed{\lambda_i = 0}$

Case-2 :- when $\lambda_i \neq 0$.

This implies $S_i = 0$

ie :- Constraint Satisfy as equality.

ie :- $g_i(x) = b_i$

Let $\lambda_i < 0 \implies \frac{\partial f}{\partial b_i} < 0$

This implies that as b_i is increased, the Objective function decreases

However as b_i increases more space become feasible and the Optimal value of the Objective function $f(x)$, Clearly Cannot decrease.

Hence an Optimal Solution

ie :- $\boxed{\lambda_i \geq 0}$

||ly for Case of minimization as b_i increases $f(x)$ Cannot increase which implies that

$\boxed{\lambda_i \leq 0}$

- **Remarks** :-

If the Constraints are equations

i.e :- $g_i(x) = b_i$

then λ_i becomes Unrestricted in sign.

Conclusion :-

Hence for Non-Linear Programming Problem

Maximize $f(x)$

Such that $g_i(x_i) \leq b_i$

The necessary Conditions

$$\frac{\partial f}{\partial x} - \sum_i \lambda_i \frac{\partial g_i}{\partial x} = 0 \cdot \left(\because \frac{\partial L}{\partial x} = 0\right)$$

$$\lambda_i \left(g_i(x_i) - b_i\right) = 0$$

$$\therefore g_i(x) \leq b_i$$

$$\boxed{\lambda_i \geq 0}$$

Minimize $f(x)$

Such that $g_i(x_i) \leq b_i$

The necessary Conditions

$$\frac{\partial f}{\partial x} - \sum_i \lambda_i \frac{\partial g_i}{\partial x} = 0$$

$$\lambda_i \left(g_i(x_i) - b_i\right) = 0$$

$$g_i(x_i) \leq b_i$$

$$\boxed{\lambda_i \leq 0}$$

- (1) Solve the NLPP

$$\text{Maximize } Z = 3.6x_1 - 0.4x_1^2 + 1.6x_2 - 0.2x_2^2$$

Such that $2x_1 + x_2 \leq 10$

$x_1, x_2 \geq 0$

Sol:- For the KKT Conditions to be necessary and sufficient for Z to a maximum $f(x)$ should be Concave

and $g(x) \leq 0$ is Convex.

for $f(x) = 3.6x_1 - 0.4x_1^2 + 1.6x_2 - 0.2x_2^2$ to be Concave we Construct the Hessian matrix as

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} -0.8 & 0 \\ 0 & -0.4 \end{bmatrix}$$

The Principle minors $D_1 = -0.8 < 0$

$$D_2 = \begin{vmatrix} -0.8 & 0 \\ 0 & -0.4 \end{vmatrix} = \underline{\underline{0.32}}$$

Thus $D_1 < 0$, $D_2 > 0$

ie:- Opposite sign with $<$ & hence it is $\underline{\underline{\text{Concave}}}$

Also the Constraint $2x_1 + x_2 \leq 10$ is Linear form and we know every Linear function is Convex

Hence the KKT Conditions are sufficient Conditions for the maximum.

Define the Lagrangian function as

$$L = f(x) - \lambda g(x)$$

$$= (3.6x_1 - 0.4x_1^2 + 1.6x_2 - 0.2x_2^2) - \lambda(2x_1 + x_2 - 10)$$

The necessary Conditions are

$$\frac{\partial L}{\partial x} = 0, \quad \lambda g = 0, \quad \lambda \geq 0, \quad g \leq 0, \quad x \geq 0$$

$$\frac{\partial L}{\partial x_1} = 0 \Rightarrow 3.6 - 0.8x_1 - 2\lambda = 0 \longrightarrow (1)$$

$$\frac{\partial L}{\partial x_2} = 0 \Rightarrow 1.6 - 0.4x_2 - \lambda = 0 \longrightarrow (2)$$

$$\lambda g = 0 \Rightarrow \lambda(2x_1 + x_2 - 10) = 0 \longrightarrow (3)$$

$$\lambda \geq 0 \Rightarrow \lambda \geq 0 \longrightarrow (4)$$

$$g \leq 0 \Rightarrow 2x_1 + x_2 \leq 10 \longrightarrow (5)$$

$$x \geq 0 \Rightarrow x_1, x_2 \geq 0 \longrightarrow (6)$$

we have the following Cases :

<u>Case-1</u> :- when $\boxed{\lambda = 0}$

from (1) & (2) we have $x_1 = 4.5$, $x_2 = 4$

which does not satisfy eqn (5) &

hence this Case is discarded.

- **Case-II** :- when $\boxed{\lambda \neq 0.}$

from (3) we get

$$2x_1 + x_2 = 10$$

from (1) & (2) $x_1 = \frac{3.6 - 2\lambda}{0.8}$ $x_2 = \frac{1.6 - \lambda}{0.4}$

$$\therefore 2\left(\frac{3.6-2\lambda}{0.8}\right) + \left(\frac{1.6-\lambda}{0.4}\right) = 10$$

$$\boxed{\lambda = 0.4} \quad (\because \lambda \geq 0)$$

Hence $x_1 = 3.5$, $x_2 = 3$ ($\because x_1\ x_2$ sub values in (5) satisfy)

(2) Solve the following NLPP

Minimize $Z = -\log x_1 - \log x_2$

such that $x_1 + x_2 \leq 2$

$x_1\ x_2 \geq 0$

Sol :- For the KKT Conditions to be for Z to be minimum

$f(x)$ should be convex & $g(x) \leq 0$ is convex.

for $f(x)$ to be convex, we construct the Hessian matrix as

$$H = \begin{bmatrix} 1/x_1^2 & 0 \\ 0 & 1/x_2^2 \end{bmatrix} \quad \because H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

Minors are $D_1 = \frac{1}{x_1^2} > 0$

$$D_2 = \begin{vmatrix} 1/x_1^2 & 0 \\ 0 & 1/x_2^2 \end{vmatrix} = \frac{1}{x_1^2 x_2^2} > 0$$

Also the Constraint $x_1 + x_2 \leq 2$ is Linear function and hence it is Convex also

Thus KKT Conditions will be minimum

Define a Lagrangian function as

$$L = (-\log x_1 - \log x_2) - \lambda (x_1 + x_2 - 2)$$

The necessary Conditions are

$$\frac{\partial L}{\partial x} = 0, \quad \lambda g \geq 0, \quad \lambda \leq 0, \quad g \leq 0, \quad x \geq 0.$$

$$\frac{\partial L}{\partial x_1} = 0 \Rightarrow \frac{-1}{x_1} - \lambda = 0 \longrightarrow (1)$$

$$\frac{\partial L}{\partial x_2} = 0 \Rightarrow \frac{-1}{x_2} - \lambda = 0 \longrightarrow (2)$$

$$\lambda g = 0 \Rightarrow \lambda (x_1 + x_2 - 2) = 0 \longrightarrow (3)$$

$$\lambda \leq 0 \Rightarrow \lambda \leq 0 \longrightarrow (4)$$

$$g \leq 0 \Rightarrow x_1 + x_2 \leq 2 \longrightarrow (5)$$

$$x \geq 0 \Rightarrow x_1, x_2 \geq 0 \longrightarrow (6)$$

We have the following Cases.

- Case-1 :- $\boxed{\lambda = 0}$

from ① & ② we have

$$x_1 = \infty, \quad x_2 = \infty$$

which violate eqn ⑤ and hence this case is discarded.

Case-2 :- when $\boxed{\lambda \neq 0}$

from eqn ③, $x_1 + x_2 = 2$

from ① & ② we get

$$x_1 = -\frac{1}{\lambda}, \quad x_2 = -\frac{1}{\lambda}$$

Hence $-\frac{1}{\lambda} - \frac{1}{\lambda} = 2$

$$\boxed{\lambda = -1}$$

Further we have $x_1 = 1, \; x_2 = 1$

which satisfy all the necessary conditions

Hence the stationary point is

$$(x_1, x_2, \lambda) = (1, 1, -1)$$

is the optimal solution and value is

$$Z = -\log 1 - \log 1$$

$$Z = 0$$

(3) Solve the following NLPP

Maximize $Z = 8x_1 + 10x_2 - x_1^2 - x_2^2$

Such that $3x_1 + 2x_2 \leq 6$

$x_1, x_2 \geq 0$

Sol:- For the KKT Conditions to be NC & SC

for Z to be Maximum

$f(x)$ should be Concave &

$g(x) \leq 0$ is Convex.

for $f(x)$ to be Concave we Construct the

Hessian Matrix

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

Its minors are $D_1 = -2 < 0$

$$D_2 = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4 > 0.$$

which is opposite sign starting with < 0

Thus it is Concave.

Also the Constraint $3x_1 + 2x_2 \leq 6$ is a

Linear Inequality function &

hence it is Convex function.

- Thus the KKT Conditions will be Necessary & sufficent for a Minimum

Define a Lagrangian function as

$$L = (8x_1 + 10x_2 - x_1^2 - x_2^2) - \lambda(3x_1 + 2x_2 - 6)$$

The necessary Conditions are

$$\frac{\partial L}{\partial x} = 0, \quad \lambda g = 0, \quad \lambda \geq 0, \quad g \leq 0, \quad x \geq 0.$$

The necessary Conditions are

$$8 - 2x_1 - 9\lambda = 0 \longrightarrow (1)$$
$$10 - 2x_1 - 2\lambda = 0 \longrightarrow (2)$$
$$\lambda(3x_1 + 2x_2 - 6) = 0 \longrightarrow (3)$$
$$\lambda \geq 0 \longrightarrow (4)$$
$$3x_1 + 2x_2 \leq 6 \longrightarrow (5)$$
$$x_1, x_2 \geq 0 \longrightarrow (6)$$

The following Cases arives.

Case-1 :- when $\boxed{\lambda = 0}$

from (1) & (2) we get $x_1 = 4$, $x_2 = 5$

which violates eqn (5) & (4) and

hence this Case is discarded

Case-2 :- when $\boxed{\lambda \neq 0}$

from eqn (3) we get.

$$3x_1 + 2x_2 = 6.$$

from ① & ② we have

$$x_1 = \frac{8-3\lambda}{2}, \qquad x_2 = 5-\lambda$$

Thus $3\left(\frac{8-3\lambda}{2}\right) + 2(5-\lambda) = 6$

$$\boxed{\lambda = \frac{32}{13}}$$

Hence $x_1 = \frac{4}{13}$, $x_2 = \frac{33}{13}$.

which satisfy all the necessary conditions.

Hence the stationary point is

$$(x_1\ x_2, \lambda) = \left(\frac{4}{13}, \frac{33}{13}, \frac{32}{13}\right)$$

is optimal solution & value is

$$Z = 8\left(\frac{4}{13}\right) + 10\left(\frac{33}{13}\right) - \left(\frac{4}{13}\right)^2 - \left(\frac{33}{13}\right)^2$$

$$Z = \frac{277}{13}$$

(4) Solve the NLPP

$$\max f(x) = 4x_1 + 6x_2 - x_1^2 - x_2^2 - x_3^2$$

Such that $x_1 + x_2 \leq 2$,

$$2x_1 + 3x_2 \leq 12$$

$$x_1, x_2 \geq 0$$

- <u>Sol</u>:- For KKT Conditions to be necessary and sufficient for Z to be maximum $f(x)$ should be Concave and $g(x) \leq 0$ is Convex

for $f(x)$ to be Concave

we Construct the Hessian matrix as

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix} \qquad H = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Its minors are

$$D_1 = -1 < 0$$

$$D_2 = \begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix} = 1 > 0$$

$$D_3 = \begin{vmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix} = -1 < 0$$

which is Opposite sign starting with < 0

Thus is Concave.

Also the Constraint $x_1 + x_2 \leq 2$, $2x_1 + 3x_2 \leq 12$ is Linear inequality function

and hence it is <u>Convex</u> function

Thus the KKT Condition will be necessary & sufficient for a Maximum.

Define the Lagrangian function L as

$$L = (4x_1 + 6x_2 - x_1^2 - x_2^2 - x_3^2) - \lambda_1 (x_1 + x_2 - 2) - \lambda_2 (2x_1 + 3x_2 - 12)$$

The necessary Conditions are

$$\frac{\partial L}{\partial x} = 0, \quad \lambda_i g_i = 0, \quad \lambda_i \geq 0, \quad g_i \leq 0, \quad x \geq 0$$

$$\frac{\partial L}{\partial x_1} = 0 \Rightarrow 4 - 2x_1 - \lambda_1 - 2\lambda_2 = 0 \longrightarrow (1)$$

$$\frac{\partial L}{\partial x_2} = 0 \Rightarrow 6 - 2x_2 - \lambda_1 - 3\lambda_2 = 0 \longrightarrow (2)$$

$$\frac{\partial L}{\partial x_3} = 0 \Rightarrow -2x_3 = 0 \longrightarrow (3)$$

$$\lambda_1, g_1 = 0 \Rightarrow \lambda_2 (2x_1 + 3x_2 - (2)) = 0 \longrightarrow (5)$$

$$\lambda_1 \lambda_2 \geq 0 \longrightarrow (6)$$

$$\lambda_1 g_1 = 0 \Rightarrow \cancel{-2x_3 = 0}$$

$$\lambda_1 (x_1 + x_2 - 2) = 0 \longrightarrow (4)$$

$$g_1 \leq 0 \Rightarrow x_1 + x_2 \leq 2 \longrightarrow (7)$$

$$g_2 \leq 0 \Rightarrow 2x_1 + 3x_2 \leq 12 \longrightarrow (8)$$

$$x_1, x_2 \geq 0 \longrightarrow (9)$$

Case-1:- when $\lambda_1 = 0$, $\lambda_2 = 0$

from (1) & (2) we get $x_1 = 2$, $x_2 = 3$

This does not satisfy (7) & (8)

and hence discarded.

- Case-2 when $\lambda_1 \neq 0$, $\lambda_2 = 0$

from (1) & (2)

$$-2x_1 + 4 = \lambda_1$$
$$-2x_2 + 6 = \lambda_1$$

from (4) we get $x_1 + x_2 = 2$

$$\Longrightarrow \frac{4-\lambda_1}{2} + \frac{6-\lambda_1}{2} = 2$$
$$\Longrightarrow \lambda_1 = 3$$

Hence $x_1 = \frac{1}{2}$, $x_2 = \frac{3}{2}$

which also satisfy (7), (8) & (9)

Hence the stationary Point is

$x_1 = \frac{1}{2}$, $x_2 = \frac{3}{2}$, $x_3 = 0$, $\lambda_1 = 3$, $\lambda_2 = 0$

The Optimal solution and value is

$$Z = 4\left(\frac{1}{2}\right) + 6\left(\frac{3}{2}\right) - \left(\frac{1}{2}\right)^2 - \left(\frac{3}{2}\right)^2 - (0)^2$$

$$\boxed{Z = \frac{17}{2}}$$

Using Hessian Matrix H :-

1) $f(x)$ is Convex $\Longleftrightarrow$ H is Positive semi-definite (Convex - minimum point)

$\Rightarrow D_1 > 0,\ D_2 \geq 0,\ D_3 \geq 0$ and atleast one $D_i = 0$

(∵ D = Principle minors)

(2) $f(x)$ is Strictly Convex $\Longleftrightarrow$ H is Positive definite (Convex - minimum point)

$\Rightarrow D_1 > 0,\ D_2 > 0,\ D_3 > 0$

(3) $f(x)$ is Concave $\Longleftrightarrow$ $-f(x)$ is Convex.

(Convex $\Rightarrow$ Maximum Point)

$\Rightarrow D_1 < 0,\ D_2 > 0,\ D_3 < 0 \ldots$

ie: alternative sign, but first Principle minor should be -ve.

GRADIENT METHOD

(or)

Steepest Ascent method

(1) The Maximizing function is

$$f(x_1, x_2) = 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2 \longrightarrow (1)$$

absolute optimum occurs at

$$(x_1^*, x_2^*) = (1/3, 4/3)$$

$\begin{pmatrix} 1/3 & 4/3 \end{pmatrix}$
0.333 1.333

Sol :- Let the initial Point be given by

$$X^0 = (1,1)$$

$$\nabla f(x) = (4 - 4x_1 - 2x_2, 6 - 2x_1 - 4x_2) \longrightarrow (2)$$

(on differentiating $x_1 \cdot x_2$)

First Iteration $x^0 = (1,1)$ in (2)

$$\nabla f(x^0) = (-2, 0)$$

$$x^1 = x^0 + r\nabla f(x^0)$$

$$= (1,1) + r(-2,0)$$

$$= (1,1) + (-2r, 0)$$

$$x^1 = (\underset{x_1}{1-2r}, \underset{x_2}{1}) \longrightarrow (3)$$

$$\boxed{h(r) = f(x^1)}$$

$$= 4(1-2r) + 6(1) - 2(1-2r)^2 + 2(1-2r)(1) - 2(1)^2$$

$$h(r) = 4 - 8r^2 + 4r$$

Let $h'(r) = 0$

$-16r + 4 = 0$

$-16r = -4$

$$\boxed{r = 1/4}$$

Substitute $r = \frac{1}{4}$ in eqn (3)

$x' = (1 - 2r, 1)$

$= (1 - 2(1/4), 1)$

$$\boxed{x' = (1/2, 1)}$$

Second Iteration :-

Sub $x' = (1/2, 1)$ in eqn (2)

$\nabla f(x') = (0, 1)$

$$\boxed{x'' = x' + r \nabla f(x')}$$

$= (1/2, 1) + r(0, 1)$

$= (1/2, 1) + (0, r)$

$$\boxed{x'' = \left(\frac{1}{2}, 1 + r\right)}$$

$h(r) = f(x'')$

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

$= \cancel{2} + 6 + 6r - \frac{1}{2} - 1 - r - \cancel{2} - 2r^2 - 4r$

$= 6 + r - 2r^2 - \frac{1}{2} - 1$

$$= 5 + r - 2r^2 - \frac{1}{2}$$

$$= r - 2r^2 + \frac{9}{2}$$

$$\boxed{h(r) = r - 2r^2 + \frac{9}{2}}$$

Let $h'(r) = 0$

$1 - 4r = 0$

$-4r = -1$

$$\boxed{r = \frac{1}{4}}$$

Sub $r = 1/4$ in x''

$x'' = (1/2, 1 + \frac{1}{4})$

$x'' = (1/2, 5/4)$

Third Iteration :-

Sub $x'' = (1/2, 5/4)$ in eqn (2)

$$\nabla f(x'') = (-1/2, 0)$$

$$\boxed{x''' = x'' + r \nabla f(x'')}$$

$$= (1/2, 5/4) + r(-1/2, 0)$$

$$= (1/2, 5/4) + (-r/2, 0)$$

$$\boxed{x''' = (1/2 - r/2, 5/4)}$$

$$h(r) = f(x^{III})$$

$$= 4\left(1/2 - r/2\right) + 6\left(5/4\right) - 2\left(1/2 - \frac{r}{2}\right)^2 -$$

$$2\left(\frac{1}{2} - \frac{r}{2}\right)\left(\frac{5}{4}\right) - 2\left(\frac{5}{4}\right)^2$$

$$= 4\left(\frac{1-r}{2}\right) + \frac{15}{2} - 2\left(\frac{1-r}{2}\right)^2 - 2\left(\frac{1-r}{2}\right)\left(\frac{5}{4}\right)$$

$$-2\left(\frac{25}{16}\right)$$

$$= 4\left(\frac{1-r}{2}\right) + \frac{15}{2} - 2\left(\frac{1-r}{2}\right)^2 - \frac{2}{2}(1-r)\left(\frac{5}{2}\right)$$

$$-\left(\frac{25}{8}\right)$$

$$= 2(1-r) + \frac{15}{2} - 2\left(\frac{1}{4} + \frac{r^2}{4} - \frac{r}{2}\right) - \left(\frac{5}{2} - \frac{5r}{2}\right) - \frac{25}{8}$$

$$= 2 - 2r + \frac{15}{2} - \frac{1}{2} - \frac{r^2}{2} + r - \frac{5}{2} - \frac{5r}{2} - \frac{25}{8}$$

$$= -\frac{r^2}{2} - r + \frac{5r}{2} + \frac{37}{8}$$

$$\boxed{h(r) = -\frac{r^2}{2} + \frac{1}{4}r + \frac{37}{8}}$$

Let $h'(r) = 0$

$$-\frac{2r}{2} + \frac{1}{4} = 0$$

$$-r + \frac{1}{4} = 0$$

$$\boxed{r = \frac{1}{4}}$$

Sub $r = \frac{1}{4}$ in eqn x^{III}

$$x^{III} = \left(1/2 - \frac{1/4}{2},\ 5/4\right)$$

$$= \left(\frac{1}{2} - \frac{1}{8},\ \frac{5}{4}\right)$$

$$\boxed{x^{III} = \left(3/8,\ 5/4\right)}$$

$\therefore \boxed{r = \frac{1}{4}}$ $\boxed{x^{III} = (3/8, 5/4)}$

Fourth Iteration :-

Sub $x^{III} = (3/8, 5/4)$ in eqn ②

$$\nabla f(x^{III}) = f(0, 1/4)$$

$$\boxed{x^{IV} = x^{III} + r \nabla f(x^{III})}$$

$$= (3/8, 5/4) + r(0, 1/4)$$

$$= (3/8, 5/4) + (0, r/4)$$

$$\boxed{x^{IV} = (\underset{x_1}{3/8}, \underset{x_2}{5/4 + \frac{r}{4}})}$$

$$h(r) = f(x^{IV})$$

$$= 4(3/8) + 6\left(\frac{5}{4} + \frac{r}{4}\right) - 2(3/8)^2 - 2(3/8)\left(\frac{5}{4} + \frac{r}{4}\right) - 2\left(\frac{5}{4} + \frac{r}{4}\right)^2$$

$$= \frac{3}{2} + \frac{30}{4} + \frac{6r}{4} - \frac{18}{64} - \left(\frac{6}{8}\right)\left(\frac{10}{4} + \frac{2r}{4}\right) - 2\left(\frac{25}{16} + \frac{r^2}{16} + \frac{20r}{4}\right)$$

$$= \frac{3}{2} + \frac{30}{4} + \frac{6r}{4} - \frac{18}{64} - \frac{60}{32} + \frac{12r}{32} - \frac{50}{16} - \frac{2r^2}{16} - \frac{20r}{4}$$

$$= -\frac{1}{8}r^2 + \frac{3r}{2} + \frac{3}{8}r - \frac{5r}{4} - \frac{60}{32} + \frac{30}{4} - \frac{50}{16} - \frac{18}{64}$$

$$= -\frac{1}{8}r^2 + \frac{3r}{2} + \frac{3}{8}r - \frac{5r}{4} - \frac{60}{32} + \frac{30}{4} - \frac{50}{16} - \frac{18}{64}$$

$$= -\frac{1}{8}r^2 + \frac{24r + 6r - 20r}{16} - \frac{149}{32}$$

$$\boxed{h(r) = -\frac{1}{8}r^2 + \frac{1}{16}r + \frac{149}{32}}$$

Let $h'(r) = 0$

$$-\frac{1}{4}r + \frac{1}{16} = 0$$

$$-\frac{1}{4}r = -\frac{1}{16}$$

$$\boxed{r = \frac{1}{4}}$$

Sub $r = \frac{1}{4}$ in eqn x^{IV}

$$x^{IV} = \left(3/8,\ 5/4 + \frac{r}{4}\right)$$

$$= \left(3/8,\ \frac{5}{4} + \frac{(1/4)}{4}\right)$$

$$\boxed{x^{IV} = \left(3/8,\ \frac{21}{16}\right)}$$

Fifth Iteration:- Sub $x^{IV} = (3/8, 21/16)$ in eqn (2)

$$\nabla f(x^{IV}) = (-1/8, 0)$$

$$x^{V} = x^{IV} + r\nabla f(x^{IV})$$

$$= (3/8, 21/16) + r(-1/8, 0)$$

$$= (3/8, 21/16) + (-r/8, 0)$$

$$\boxed{x^{V} = \left(\frac{3}{8} - \frac{r}{8}, \frac{21}{16}\right)}$$

$$\therefore \boxed{h(r) = \frac{-2r^2}{64} + \frac{1}{64}r + \frac{597}{38}}$$

$$h(r) = f(x^{V})$$

$$= 4\left(\frac{3-r}{8}\right) + 6\left(\frac{21}{16}\right) - 2\left(\frac{3-r}{8}\right)^2 - 2\left(\frac{3-r}{8}\right)\left(\frac{21}{16}\right) - 2\left(\frac{21}{16}\right)^2$$

$$\boxed{h(r) = \frac{-2r^2}{64} + \frac{1}{64}r + \frac{597}{38}}$$

Let $h'(r) = 0$

$$\frac{-4r}{64} + \frac{1}{64} = 0$$

$$\frac{-r}{16} + \frac{1}{64} = 0$$

$$\frac{-r}{16} = \frac{-1}{64}$$

$$\boxed{r = \frac{1}{4}}$$

Sub $r = \frac{1}{4}$ in eqn X^r

$$X^v = \left(\frac{3-r}{8}, \frac{21}{16}\right)$$

$$= \left(\frac{3-\left(\frac{1}{4}\right)}{8}, \frac{21}{16}\right)$$

$$\boxed{X^v = \left(11/32, \frac{21}{16}\right)}$$

<u>Sixth Iteration</u> :- sub $X^v = \left(\frac{11}{32}, \frac{21}{16}\right)$ in eqn (2)

$$\nabla f(x^v) = \left(0, \frac{1}{16}\right)$$

Because $\nabla f(x^v) \simeq 0$, the Process can be terminated at this point.

∴ The approximate maximum Point is given by

$$X^v = (\underline{0.34375}, 1.3125) \text{ (or)}$$

$$X^v = \underline{\left(\frac{11}{32}, \frac{21}{16}\right)}$$

∴ The <u>exact</u> <u>Optimum</u> is

$$X^* = \underline{(0.3333, 1.3333)}$$