SREENIVASA INSTITUTE OF TECHNOLOGY AND MANAGEMENT STUDIES
(AUTONOMOUS)
MCA DEPARTMENT

LECTURE NOTES

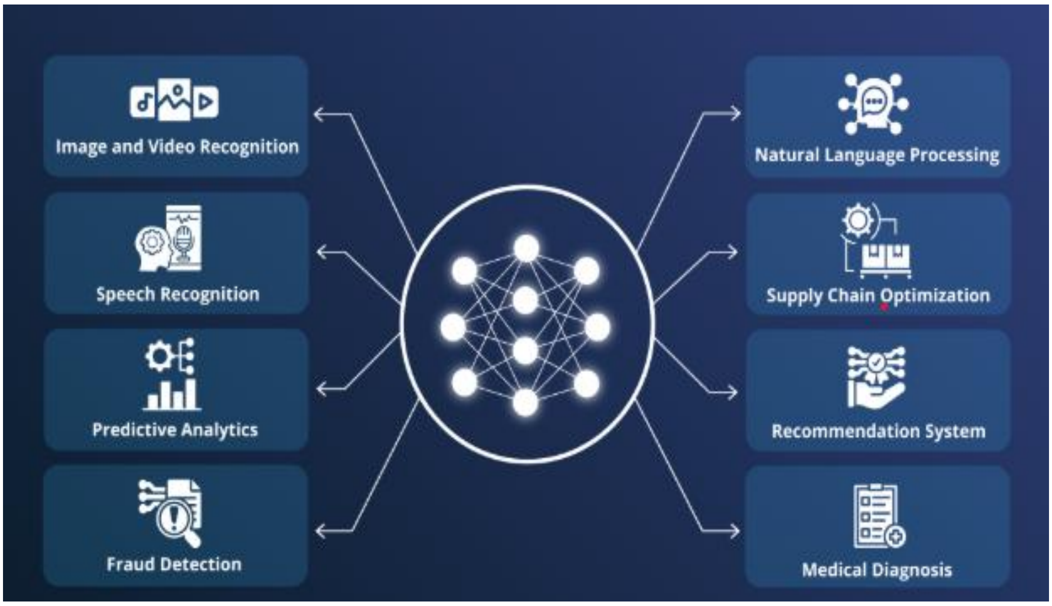
Subject Name: DEEP LEARNING

Year / Branch: II MCA – II SEMESTER

Regulation: R24

Prepared By: Mr. N P Gangadhar,

Assistant Professor



Deep learning is the key technology behind many recent breakthroughs in artificial intelligence
—Yann LeCun

Syllabus:

II MCA – II SEMESTER			
COURSE CODE:	**24MCA221**	**CREDITS:**	**4**
COURSE TITLE:	**DEEP LEARNING**	**L-T-P:**	**4-0-0**

PREREQUISITES: *A Course on "Machine learning", Artificial Neural Networks and Knowledge on basic linear algebra may be helpful.*

COURSE EDUCATIONAL OBJECTIVES:

CEO 1: Toacquireknowledgeaboutdifferentmathematicaltoolsfor DeepLearning.
CEO 2: Tounderstandfundamentalsofartificialneural networks.
CEO 3: ToExplore various optimizers and Regularization Techniques
CEO 4: ToLearn about ConvolutionalNeural Networks.
CEO 5:Tocomprehend Object Detection, Autoencoder and GAN Architectures

UNIT-I: MATHEMATICAL TOOLS FOR DEEP LEARNING	**Lecture Hrs:12**

Linear Algebra :Matrix, Vector,Transpose,Tensor, operations on elements, Systems of Linear equations,Rank, Norm,expressing a Matrix, Determinant, Trace, Eigen values and Eigen Vectors, Singular Value Decomposition(SVD). **Statistics** – Probability, Random Variable, Binomial Distribution, Poisson Distribution, Normal Distribution, Sampling, Central limit Theorem. **Calculus**- Derivatives, rules forderivatives,Partial derivatives.

UNIT-II: FUNDAMENTALSOFARTIFICIALNEURALNETWORK(ANN)	**Lecture Hrs:12**

Understanding the Biological Neuron, Exploring the Artificial Neuron, Early Implementation of ANN – RmcCulloch-Pits mosel of Neuron, Rosenblatt's Perceptron,Types of Activation Functions-linear function, non-linear function, softmax function. Architectures of neural network- Single layer feed forward network, Multi-layer feed forward ANN, Recurrent Neural Network, Convolutional Networks. Learning Process in ANN – Weight of Interconnection between neurons, Gradient Descent and Backpropagation.

UNIT-III: TRAININGDEEPNEURALNETWORK	**Lecture Hrs:12**

Initializing Weights- He/ Kaiming initialization, Xavier initialization. Batch, Mini-batch and stochastic gradient descent. Regularization-L1/ L2 regularization, Early stopping, Dropout regularization, data augmentation. Normalization of inputs-Batch Normalization, Batch Normas regularizer.

UNIT-IV: CONVOLUTIONAL NETWORKS	**Lecture Hrs:10**

Building blocks of CNN, Building a Convolution Neural Network, Popular CNN Architectures-LeNet-5, AlexNet, ZFNET, VGG-16, GoogleNet and ResNet Object Detection- one stage deection techniques – YOLO, SSD,Two stage object detection techniques – R-CNN, fast R-CNN, faster R-CNN, Mask R-CNN, Applications of Object Detection

UNIT-V: SEQUENCE-BASEDMODELS ANDOTHERDLARCHITECTURES	**Lecture Hrs:12**

Recurrent Neural Network – Data Preparation for RNN Vanishing Gradient problem and RNN, Applications of RNN, Types of RNN, Limitations of RNN, Longshort-termmemory (LSTM), Gated RNNs, Bidirectional RNNs. Other DL Architectures- Autoencoder, Architecture and its applications, GAN, GAN Architecture and its applications,

TEXTBOOKS:

1. *AmitKumarDas,SaptarsiGoswami,PabitraMitra,AmlanChakrabarti,"DeepLearning", Pearson Paperback,First Edition,2021.*

2. IanGoodfellow,YoshuaBengio,AaronCourville, "DeepLearning",MITPress,2016.

REFERENCE BOOKS:

1. JoshPattersonandAdamGibson,"Deeplearning:Apractitioner'sapproach",O'ReillyMedia,First Edition,2017.
2. NikhilBuduma, "Fundamentals of Deep Learning, Designing next-generationmachineintelligence algorithms",O'Reilly,ShroffPublishers,2019.

COURSE OUTCOMES: *On successful completion of this course, students will be able to:*		POs related to COs
CO1	Understand the mathematical background required for Deep learning.	**PO1,PO2**
CO2	Analyze the fundamental concepts of Artificial neural networks.	**PO1,PO2,PO3**
CO3	Understand and apply the deep neural networks	**PO1,PO2,PO3,PO4**
CO4	Explore the Purpose of Convolution Neural Network, Popular Architectures of CNN	**PO1,PO2,PO3,PO4,PO8**
CO5	To learn about sequence-based models like RNN, LSTM and Gated RNN and other DL architectures and use for real-world problems	**PO1,PO2,PO3,PO4,PO8**

CO-PO MAPPING(DETAILED; HIGH:3; MEDIUM:2; LOW:1)

Course	**POs / COs**	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**
C401: Deep Learning	**C401.1**	**3**	**3**	-	-	-	-	-	-
	C401.2	**3**	**3**	**2**	-	-	-	-	-
	C401.3	**3**	**3**	**2**	**3**	-	-	-	-
	C401.4	**3**	**3**	**3**	**2**	-	-	-	**3**
	C401.5	**3**	**3**	**3**	**3**	-	-	-	**3**
	C401	**3**	**3**	**2.5**	**2.67**				**3**

Unit Overview

This unit builds the mathematical foundation for deep learning by explaining how data is represented using vectors, matrices, and tensors, how uncertainty is handled through probability and statistics, and how models are optimized using calculus. It also introduces important concepts such as eigenvalues, probability distributions, and derivatives, which are essential for understanding how neural networks learn and improve. These mathematical tools are widely used in designing, training, and optimizing deep learning algorithms, making this unit crucial for developing a strong base in artificial intelligence and machine learning.

Objectives of the Unit

After studying this unit, students will be able to:

- To understand how **data is represented mathematically** using vectors and tensors.
- To learn how **linear algebra supports neural network computations**.
- To understand **uncertainty and randomness** using probability theory.
- To study how **calculus helps in model optimization**.
- To apply mathematical concepts in **real-world deep learning problems**.

Learning Outcomes

After completing this unit, students will be able to:

- To perform **matrix operations** such as addition, multiplication, and transpose.
- To represent real-world data as **vectors and tensors**.
- To compute **determinant, rank, norm**, and solve linear equations.
- To analyze data using **mean, variance, and probability distributions**.
- To apply **Central Limit Theorem** in data analysis.
- To calculate **gradients using derivatives and partial derivatives**.
- To relate mathematical concepts to **training neural networks**.

Importance of studying this Unit:

- Provides the **foundation for all machine learning and deep learning models**.
- Helps in understanding how **weights and biases are updated** in neural networks.
- Essential for implementing **optimization algorithms** like gradient descent.
- Used in **feature extraction, dimensionality reduction (SVD)**, and data preprocessing.
- Improves analytical and problem-solving skills required in **AI and Data Science**.

Key Concepts:

1. Linear Algebra

- **Vector:** A list of numbers representing data (e.g., features of an image).
- **Matrix:** A 2D array used to store datasets.
- **Tensor:** Multi-dimensional array (used in deep learning frameworks).
- **Transpose:** Flipping rows into columns.
- **Determinant:** Value that shows if a matrix is invertible.
- **Rank:** Number of independent rows/columns.
- **Eigenvalues & Eigenvectors:** Help in understanding transformations.
- **SVD:** Decomposes matrix into simpler components (used in compression).

2. Statistics & Probability

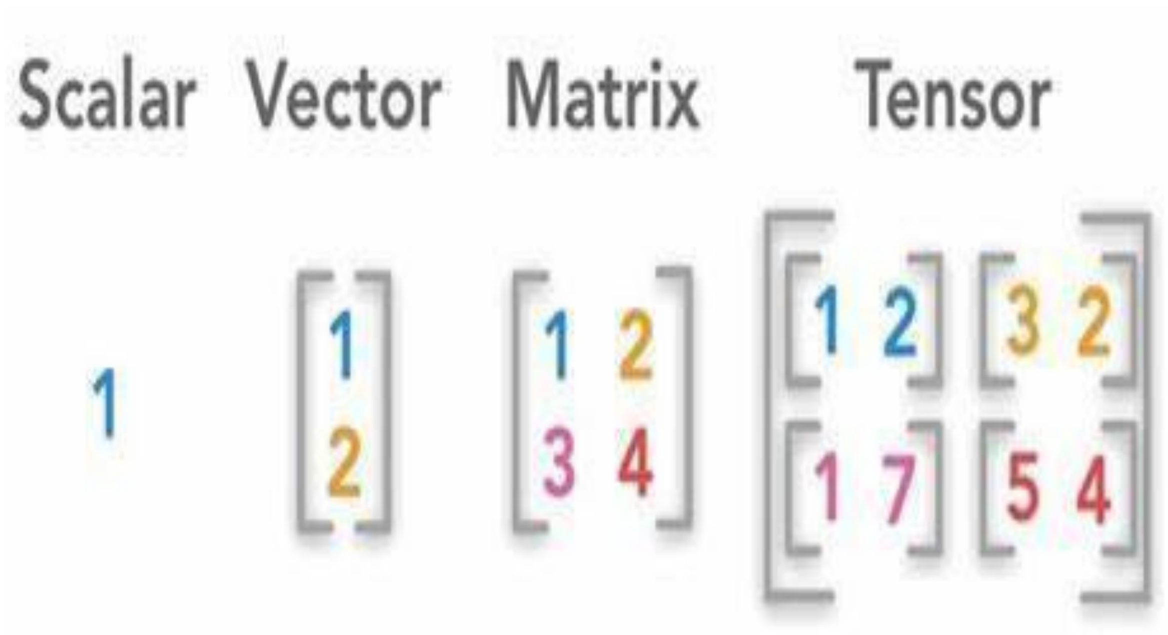
- **Probability:** Measure of likelihood of an event.
- **Random Variable:** Variable with random values.
- **Binomial Distribution:** Used for binary outcomes.
- **Poisson Distribution:** Used for counting events.
- **Normal Distribution:** Bell-shaped curve (most important).
- **Sampling:** Selecting subset from population.
- **Central Limit Theorem:** Sample mean tends to normal distribution.

3. Calculus

- **Derivative:** Rate of change of a function.
- **Rules of Derivatives:** Product, quotient, chain rules.
- **Partial Derivative:** Derivative with respect to one variable.
- Used in gradient descent and backpropagation**.**

Linear Algebra

- The understanding of **Linear Algebra** is a must for **better comprehension and effective implementation of Neural Networks.**
- Scalars, Vectors, Matrices and Tensors are the most important mathematical concepts of Linear Algebra.
- **Scalars** are single numbers and are an example of a 0th-order tensor.
- **Vectors** are single dimensional arrays and are an example of 1st-order tensor.
- **Matrices** are Two Dimensional Arrays are an example of 2nd-order tensors
- **Tensors** are n-dimensional Arrays with n greater than 2 and are an example of n-order tensors.



Scalars are single numbers and are an example of a 0th-order tensor.	**Vectors** are single dimensional arrays and are an example of 1st-order tensor.	**Tensors** are n-dimensional Arrays with n greater than 2 and are an example of n-order tensors.	**Matrices** are Two Dimensional Arrays are an example of 2nd-order tensors.

VECTOR

- Vectors are fundamental objects in linear algebra and are used to represent various physical quantities such as force, velocity, displacement, etc.
- Example : $\mathbf{v} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$
- They can be

1) added together,

1) multiplied by scalars (real numbers)

2) Perform dot product, and more

VECTOR ADDITION

Suppose we have two vectors $\mathbf{v}$ and $\mathbf{w}$ given by:

$$\mathbf{v} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$
$$\mathbf{w} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

To add these two vectors, we simply add their corresponding components. So, the result $\mathbf{v} + \mathbf{w}$ would be:

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 2+3 \\ -1+4 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

So, the resulting vector after adding $\mathbf{v}$ and $\mathbf{w}$ is $\begin{pmatrix} 5 \\ 3 \end{pmatrix}$.

VECTOR MULTIPLIED BY SCALAR

Let's consider a vector $\mathbf{v}$ and a scalar k:

$$\mathbf{v} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$$
$$k = 4$$

To find the result of multiplying the vector $\mathbf{v}$ by the scalar k, we simply multiply each component of $\mathbf{v}$ by k:

$$k\mathbf{v} = 4\begin{pmatrix} 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 4 \times 2 \\ 4 \times (-3) \end{pmatrix} = \begin{pmatrix} 8 \\ -12 \end{pmatrix}$$

So, when we multiply the vector $\mathbf{v}$ by the scalar $k = 4$, we get $\begin{pmatrix} 8 \\ -12 \end{pmatrix}$.

Vector Properties

For all vectors u, v, and w, and for all scalars b and c:

1. u + v = v + u.

2. u + (v + w) = (u + v) + w.

3. v + O = v.

4 1.v = v; 0.v = O.

5. v + (- v) = O.

6. b(cv) = (bc)v.

7. (b + c)v = bv + cv.

8. b(u + v) = bu + bv.

MATRIX

- A matrix is a way of representing the numbers in the form of rows and columns and all the numbers are represented in the cells of this matrix.
- We represent the matrix as [A] m×n and m represents the number of rows and n represents the number of columns.

Types of Matrix

a) Singleton Matrix

b) Null Matrix

c) Row Matrix

d) Column Matrix

e) Horizontal Matrix

f) Vertical Matrix

g) Rectangular Matrix

h) Square Matrix

i) Diagonal Matrix

j) Scalar Matrix

k) Identity Matrix

l) Triangular Matrix

   a. Upper Triangular Matrix

   b. Lower Triangular Matrix

m) Singular Matrix

n) Non Singular Matrix

o) Symmetric Matrix

p) Skew Symmetric Matrix

q) Orthogonal Matrix

r) Idempotent Matrix

1) Singleton Matrix

- A matrix that has only one element is called a singleton matrix.
- In this type of matrix number of columns and the number of rows is equal to 1.
- A singleton matrix is represented as $[a]_{1\times1}$.
- Example of Singleton Matrix

$$[5]_{1\times1}$$

2) Null Matrix

- A matrix whose all elements are zero is called a Null Matrix.
- A null matrix is also called a Zero Matrix because its all elements are zero.
- Example of Null Matrix.

$$\begin{bmatrix}0 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{bmatrix}_{3\times3}$$

3) Row Matrix

- A matrix that contains only one row and any number of columns is known as a row matrix.
- A row matrix is represented as $[a]_{1\times n}$ where 1 is

  the number of row and n is the number of columns

  present in a row matrix.

- Example of Row Matrix.

$$\begin{bmatrix}1 & 3 & 7\end{bmatrix}_{1\times3}$$

4) Column Matrix

- A matrix that contains only one Column and any number of Rows is known as a row matrix.
- A row matrix is represented as $[a]_{n*1}$ where n is the number of rows and 1 is the number of column.
- Example of Column Matrix.

$$\begin{bmatrix}1\\15\\4\\5\end{bmatrix}_{4\times1}$$

5) Horizontal Matrix

- A matrix in which the number of rows is lower than the number of columns is called a Horizontal Matrix.
- Example of Horizontal Matrix.

6) Vertical Matrix

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

- A matrix in which the number of rows exceeds the number of columns is called a Vertical Matrix
- Example of Vertical Matrix.

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

7) Rectangular Matrix

- A matrix that does not have an equal number of rows and columns is known as a Rectangular Matrix.
- Example of Rectangular Matrix.

$$\begin{bmatrix} 1 & 3 & 7 & 15 \\ 3 & 4 & 6 & 11 \\ 5 & 2 & 9 & 8 \end{bmatrix}_{3\times 4}$$

8) Square Matrix

- A matrix that has an equal number of rows and an equal number of columns is called a Square Matrix.
- Example of Square Matrix.

$$\begin{bmatrix} 8 & 3 & 2 \\ 6 & 4 & 6 \\ 5 & 7 & 9 \end{bmatrix}_{3\times 3}$$

9) Diagonal Matrix

- A matrix that has all elements as 0 except diagonal elements is known as a diagonal matrix.
- Example of Diagonal Matrix.

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix}_{3\times 3}$$

10) Scalar Matrix

- A diagonal matrix whose all diagonal elements are non-zero and the same is called a Scalar Matrix.
- Scalar Matrix is a kind of diagonal matrix where all diagonal elements are the same.
- Example of Scalar Matrix.

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}_{3\times 3}$$

11) Identity Matrix

- A diagonal matrix where all the diagonal elements are 1

  and all non-diagonal elements are 0 is called an Identity

  Matrix.

- The Identity Matrix is called Unit Matrix.
- The identity matrix or unit matrix always has an equal number of rows and columns.
- Example of Rectangular.

12) Triangular Matrix

- A square matrix in which the non-zero elements form a triangular below and above the diagonal is called a Triangular Matrix.
- Based on the triangle formed below or above the diagonal, the triangular matrix is classified as:
    - **Upper Triangular Matrix**
    - **Lower Triangular Matrix**

a. Upper Triangular Matrix

- A square matrix in which all the elements below the diagonal are zero and the elements from the diagonal and above are non-zero elements is called an Upper Triangular Matrix.
- In an Upper Triangular Matrix, the non-zero elements form a triangular-like shape**.**

b. Lower Triangular Matrix

- A square matrix in which all the elements above the diagonal are zero and the elements from the diagonal and below **are non-zero elements is called a Lower Triangular Matrix**
- In a Lower Triangular Matrix, the non-zero elements form a triangular-like shape from the diagonal and below...

13) Symmetric Matrix

- A square matrix "A" of any order is defined as a symmetric matrix if the transpose of the matrix is equal to the original matrix itself, i.e., $A^T = A$.

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

14) Symmetric Matrix

- A square matrix "A" of any order is defined as a skew-symmetric matrix if the transpose of the matrix is equal to the negative of the original matrix itself, i.e., $A^T = -A$.

$$\begin{bmatrix} 0 & 3 & 5 \\ -3 & 0 & -2 \\ -5 & 2 & 0 \end{bmatrix}$$

15) Orthogonal Matrix

- A square matrix whose transpose is equal to its inverse is called Orthogonal Matrix in an Orthogonal Matrix. $A^T = A$ then $AA^T = I$ where I is the Identity Matrix.

Matrix Operations

- Matrix operations are the operations that are used to combine various matrices to form a single matrix.
- The operations such as addition, subtraction, and multiplication are easily performed on the matrix.
- These matrix operations are very useful to solve matrix problems and to find the transpose and the inverse of the matrix.
- Various matrix operations that are used to solve matrix problems are
  - Addition of Matrix
  - Subtraction of Matrix
  - Scaler Multiplication of Matrix
  - Multiplication of Matrix

Addition of Matrix:

$$A + B = \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix} + \begin{bmatrix} 5 & 9 \\ 0 & 7 \end{bmatrix}$$
$$= \begin{bmatrix} 4+5 & 1+9 \\ 3+0 & 2+7 \end{bmatrix}$$
$$= \begin{bmatrix} 9 & 10 \\ 3 & 9 \end{bmatrix}$$

Properties of Matrix Addition:

There are various properties associated with matrix addition that are, for matrices A, B, and C of the same order, then

1) **Commutative Law:** A + B = B + A
2) **Associative Law:** (A + B) + C = A + (B + C)
3) **Identity of Matrix:** A + O = O + A = A, where O is a zero matrix which is the Additive Identity of Matrix.
4) **Additive Inverse:** A + (-A) = O = (-A) + A, where (-A) is obtained by changing the sign of every element of A, which is the additive inverse of the matrix.

Commutative property of addition: $A + B = B + A$

This property states that you can add two matrices in any order and get the same result.

This parallels the commutative property of addition for real numbers. For example, $3 + 5 = 5 + 3$.

The following example illustrates this matrix property.

$$\begin{bmatrix} 3 & 7 \\ 2 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 2 \\ 8 & 1 \end{bmatrix} = \begin{bmatrix} 3+5 & 7+2 \\ 2+8 & 4+1 \end{bmatrix}$$

$$= \begin{bmatrix} 5+3 & 2+7 \\ 8+2 & 1+4 \end{bmatrix} \quad \text{(Real \# addition is commutative.)}$$

$$= \begin{bmatrix} 5 & 2 \\ 8 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 7 \\ 2 & 4 \end{bmatrix}$$

Associative property of addition: $(A+B)+C = A+(B+C)$

This property states that you can change the grouping in matrix addition and get the same result. For example, you can add matrix A to B first, and then add matrix C, **or**, you can add matrix B to C, and then add this result to A.

This property parallels the associative property of addition for real numbers. For example, $(2+3)+5 = 2+(3+5)$.

Let's justify this matrix property by looking at an example.

$(A+B)+C$	$A+(B+C)$
$\left(\begin{bmatrix}5 & 2\\2 & 8\end{bmatrix}+\begin{bmatrix}3 & 4\\9 & 6\end{bmatrix}\right)+\begin{bmatrix}1 & 0\\3 & 7\end{bmatrix}$	$\begin{bmatrix}5 & 2\\2 & 8\end{bmatrix}+\left(\begin{bmatrix}3 & 4\\9 & 6\end{bmatrix}+\begin{bmatrix}1 & 0\\3 & 7\end{bmatrix}\right)$
$\downarrow$	$\downarrow$
$\begin{bmatrix}5+3 & 2+4\\2+9 & 8+6\end{bmatrix}+\begin{bmatrix}1 & 0\\3 & 7\end{bmatrix}$	$\begin{bmatrix}5 & 2\\2 & 8\end{bmatrix}+\begin{bmatrix}3+1 & 4+0\\9+3 & 6+7\end{bmatrix}$
$\downarrow$	$\downarrow$
$\begin{bmatrix}(5+3)+1 & (2+4)+0\\(2+9)+3 & (8+6)+7\end{bmatrix}$	$\begin{bmatrix}5+(3+1) & 2+(4+0)\\2+(9+3) & 8+(6+7)\end{bmatrix}$

$=$

Additive inverse property: $A+(-A)=O$

The **opposite** of a matrix A is the matrix $-A$, where each element in this matrix is the *opposite* of the corresponding element in matrix A.

For example, if $A = \begin{bmatrix}-2 & 8\\-3 & 1\end{bmatrix}$, then $-A = \begin{bmatrix}2 & -8\\3 & -1\end{bmatrix}$.

If we add A to $-A$ we get a zero matrix, which illustrates the additive inverse property.

$$A+(-A) = \begin{bmatrix}-2 & 8\\-3 & 1\end{bmatrix}+\begin{bmatrix}2 & -8\\3 & -1\end{bmatrix}$$

$$= \begin{bmatrix}-2+2 & 8+(-8)\\-3+3 & 1+(-1)\end{bmatrix}$$

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

Subtraction of Matrix:

$$A - B = \begin{bmatrix} -2 & 3 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 8 & 1 \\ 5 & 4 \end{bmatrix}$$
$$= \begin{bmatrix} -2-8 & 3-1 \\ 0-5 & 1-4 \end{bmatrix}$$
$$= \begin{bmatrix} -10 & 2 \\ -5 & -3 \end{bmatrix}$$

Properties of Subtration Matrix:

There are various properties associated with matrix Subtraction that are, for matrices A, B, and C of the same order, then

1) The number of rows and columns should be the same for the matrix subtraction.
2) The subtraction of matrices is not commutative, that is, A - B ≠ B – A
3) The subtraction of matrices is not associative, that is, (A - B) - C ≠ A - (B - C)
4) The subtraction of a matrix from itself results in a null matrix, that is, A - A = O.
5) Subtraction of matrices is the addition of the negative of a matrix to another matrix, that is, A - B = A + (-B).

Scalar Multiplication of a Matrix

For any matrix A = $[a_{ij}]_{m\times n}$ if we multiply the matrix A with any scaler (say k) then the scaler is multiplied by each element of the matrix and this is called the scalar multiplication of matrices.

For any matrix A = $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ if it is multiplied by any scaler k then,

kA = $\begin{bmatrix} ka & kb \\ kc & kd \end{bmatrix}$

Example 4: If

$$A = \begin{bmatrix} 1 & 0 & -3 \\ -2 & 4 & 1 \end{bmatrix}$$

then the scalar multiple 2 A is obtained by multiplying every entry of A by 2:

$$2A = \begin{bmatrix} 2 & 0 & -6 \\ -4 & 8 & 2 \end{bmatrix}$$

Properties of Scalar Multiplication of a Matrix

Property	Example
Associative property of multiplication	$(cd)A = c(dA)$
Distributive properties	$c(A+B) = cA + cB$
	$(c+d)A = cA + dA$
Multiplicative identity property	$1A = A$
Multiplicative properties of zero	$0 \cdot A = O$
	$c \cdot O = O$
Closure property of multiplication	cA is a matrix of the same dimensions as A.

Associative property of multiplication: (cd)A = c(dA)

The following example illustrates this property for $c = 2$, $d = 3$, and
$A = \begin{bmatrix} 5 & 4 \\ 8 & 1 \end{bmatrix}$.

$(c \cdot d)A$	$c(dA)$
$(2 \cdot 3)\begin{bmatrix} 5 & 4 \\ 8 & 1 \end{bmatrix}$	$2\left(3\begin{bmatrix} 5 & 4 \\ 8 & 1 \end{bmatrix}\right)$
$\downarrow$	$\downarrow$
	$2\begin{bmatrix} (3 \cdot 5) & (3 \cdot 4) \\ (3 \cdot 8) & (3 \cdot 1) \end{bmatrix}$
	$\downarrow$
$\begin{bmatrix} (2 \cdot 3) \cdot 5 & (2 \cdot 3) \cdot 4 \\ (2 \cdot 3) \cdot 8 & (2 \cdot 3) \cdot 1 \end{bmatrix}$	$\begin{bmatrix} 2 \cdot (3 \cdot 5) & 2 \cdot (3 \cdot 4) \\ 2 \cdot (3 \cdot 8) & 2 \cdot (3 \cdot 1) \end{bmatrix}$

$=$

Distributive property of Scalar multiplication: c(A+B) = cA+cB

Here's an example where $c = 2$, $A = \begin{bmatrix} 5 & 2 \\ 3 & 1 \end{bmatrix}$, and $B = \begin{bmatrix} 3 & 4 \\ 2 & 6 \end{bmatrix}$:

$c(A+B)$	$cA+cB$
$2\left(\begin{bmatrix} 5 & 2 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 4 \\ 2 & 6 \end{bmatrix}\right)$ $\downarrow$ $2\begin{bmatrix} 5+3 & 2+4 \\ 3+2 & 1+6 \end{bmatrix}$ $\downarrow$ $\begin{bmatrix} 2(5+3) & 2(2+4) \\ 2(3+2) & 2(1+6) \end{bmatrix}$	$2\begin{bmatrix} 5 & 2 \\ 3 & 1 \end{bmatrix} + 2\begin{bmatrix} 3 & 4 \\ 2 & 6 \end{bmatrix}$ $\downarrow$ $\begin{bmatrix} 2\cdot 5 & 2\cdot 2 \\ 2\cdot 3 & 2\cdot 1 \end{bmatrix} + \begin{bmatrix} 2\cdot 3 & 2\cdot 4 \\ 2\cdot 2 & 2\cdot 6 \end{bmatrix}$ $\downarrow$ $\begin{bmatrix} 2\cdot 5+2\cdot 3 & 2\cdot 2+2\cdot 4 \\ 2\cdot 3+2\cdot 2 & 2\cdot 1+2\cdot 6 \end{bmatrix}$

$=$

Multiplicative identity property: $1A = A$

This property says that when you multiply any matrix A by the scalar 1, the result is simply the original matrix A.

So, for example, if $A = \begin{bmatrix} 2 & 5 \\ 1 & 7 \end{bmatrix}$, then we have:

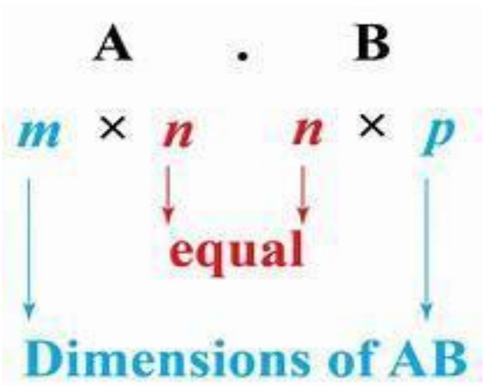
$$1\begin{bmatrix} 2 & 5 \\ 1 & 7 \end{bmatrix} = \begin{bmatrix} 1\cdot 2 & 1\cdot 5 \\ 1\cdot 1 & 1\cdot 7 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 5 \\ 1 & 7 \end{bmatrix}$$

Matrix Multiplication

- Matrix Multiplication is a binary operation performed on two matrices to get a new matrix called the product matrix.
- For matrix multiplication , the number of columns in the first matrix must be equal to the number of rows in the second matrix.

- The resulting matrix, known as the **matrix product**, has the number of rows of the first and the number of columns of the second matrix. The product of matrices **A** and **B** is denoted as **AB**.



$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \text{x} \begin{bmatrix} 10 & 11 \\ 20 & 21 \\ 30 & 31 \end{bmatrix}$$

$$= \begin{bmatrix} 1\text{x}10 + 2\text{x}20 + 3\text{x}30 & 1\text{x}11 + 2\text{x}21 + 3\text{x}31 \\ 4\text{x}10 + 5\text{x}20 + 6\text{x}30 & 4\text{x}11 + 5\text{x}21 + 6\text{x}31 \end{bmatrix}$$

$$= \begin{bmatrix} 10+40+90 & 11+42+93 \\ 40+100+180 & 44+105+186 \end{bmatrix} = \begin{bmatrix} 140 & 146 \\ 320 & 335 \end{bmatrix}$$

Properties of matrix multiplication

In this table, A, B, and C are $n \times n$ matrices, I is the $n \times n$ identity matrix, and O is the $n \times n$ zero matrix

Property	Example
The commutative property of multiplication **does not hold!**	$AB \neq BA$
Associative property of multiplication	$(AB)C = A(BC)$
Distributive properties	$A(B+C) = AB + AC$
	$(B+C)A = BA + CA$
Multiplicative identity property	$IA = A$ and $AI = A$
Multiplicative property of zero	$OA = O$ and $AO = O$
Dimension property	The product of an $m \times n$ matrix and an $n \times k$ matrix is an $m \times k$ matrix.

Matrix multiplication is not commutative

One of the biggest differences between real number multiplication and matrix multiplication is that matrix multiplication is **not commutative**.

In other words, in matrix multiplication, the order in which two matrices are multiplied matters!

See for yourselves!

Let's take a look at a concrete example with the following matrices.

$$A = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 6 & 2 \\ 3 & 2 \end{bmatrix}$$

1) Find AB and BA.

$AB =$ $\qquad$ $BA =$

Associative property of multiplication: $(AB)C = A(BC)$

This property states that you can change the grouping surrounding matrix multiplication.

For example, you can multiply matrix A by matrix B, and then multiply the result by matrix C, **or** you can multiply matrix B by matrix C, and then multiply the result by matrix A.

When using this property, be sure to pay attention to the order in which the matrices are multiplied, since we know that the commutative property does not hold for matrix multiplication!

Transpose of a Matrix:

A Matrix which is formed by turning all the rows of a given matrix into columns and vice-versa."

Let's Work Out-

Example- Find the transpose of the given matrix

$$M = \begin{bmatrix} 2 & -9 & 3 \\ 13 & 11 & -17 \\ 3 & 6 & 15 \\ 4 & 13 & 1 \end{bmatrix}$$

Solution- Given a matrix of the order 4×3.

The transpose of a matrix is given by interchanging rows and columns.

$$M^T = \begin{bmatrix} 2 & 13 & 3 & 4 \\ -9 & 11 & 6 & 13 \\ 3 & -17 & 15 & 1 \end{bmatrix}$$

Properties of Transpose Matrix

1) A square matrix "A" of order "n × n" is said to be an orthogonal matrix, if $AA^T = A^TA = I$, where "I" is an identity matrix of order "n × n."
2) A square matrix "A" of order "n × n" is said to be a symmetric matrix if its transpose is the same as the original matrix, i.e., $A^T = A$.
3) A square matrix "A" of order "n × n" is said to be a skew-symmetric matrix if its transpose is equal to the negative of the original matrix, i.e., $A^T = –A$.
4) Transpose of the transpose matrix is the original matrix itself i.e., $(A^T)^T = A$
5) A Transpose of Product of Matrices: This property says that $(AB)^T = A^T B^T$
6) Multiplication by Constant, i.e $(KA)^T = KA^T$.
7) Transpose of Addition of Matrices: This property says that. $(A+B)^T = A^T$

Trace of Matrix

- The sum of diagonal elements of a matrix is known as the trace of a matrix.
- The Trace of a matrix A can be represented as tr(A).
- The Trace of a matrix can be calculated for a square matrix only.

$$A = \begin{bmatrix} 15 & 12 & 9 \\ 4 & 6 & 11 \\ 5 & 9 & 0 \end{bmatrix}_{3\times 3}$$

$$tr(A) = 15 + 6 + 0 = 21$$

Properties of Trace Matrix

Property 1: **Trace of the sum of two matrices is equal to the sum of a trace of an individual matrix**

Mathematically it can be written as $tr(A+B) = tr(A) + tr(B)$

$$A = \begin{bmatrix} 15 & 12 & 9 \\ 4 & 6 & 11 \\ 5 & 9 & 0 \end{bmatrix}_{3\times 3}$$

$$tr(A) = 15 + 6 + 0 = 21$$

$$B = \begin{bmatrix} 4 & 3 & 7 \\ 8 & 1 & 2 \\ 5 & 6 & 1 \end{bmatrix}_{3\times 3}$$

$$tr(B) = 4 + 1 + 1 = 6$$

Now, $tr(A)+tr(B) = 21+6 = 27$

$$A + B = \begin{bmatrix} 19 & 15 & 16 \\ 12 & 7 & 13 \\ 10 & 15 & 1 \end{bmatrix}_{3\times 3}$$

$$tr(A + B) = 19 + 7 + 1 = 27$$

Property 2: **Trace of a matrix that is multiplied by some scalar is equal to the multiplication of the trace of the matrix and scalar.**

Mathematically it can be represented as ***tr(kA) = k tr(A)***

$$A = 2 \times \begin{bmatrix} 1 & 4 & 3 \\ 5 & 7 & 2 \\ 1 & 3 & 8 \end{bmatrix}_{3\times3}$$

tr(2 × A) *= 2 + 14 + 16 = 32*

$$2 \times A = \begin{bmatrix} 2 & 8 & 6 \\ 10 & 14 & 4 \\ 2 & 6 & 16 \end{bmatrix}_{3\times3}$$

2 × tr(A) *= 2 * (1 + 7 + 8)*

= 32

You can see ***tr(2 × A) = 2 × tr(A)***

Cofactor of the Matrix

- A cofactor matrix is a matrix that comprises the cofactors of each element in a matrix.
- A cofactor is a number obtained when the minor M_{ij} of the element a_{ij} is multiplied by the $(-1)^{i+j}$. i and j represent the row and column of the particular element whose cofactor is being determined.
- Minor of an element is obtained by eliminating the row and column of that particular element i.e. eliminating row **i** and column **j** and then taking the remaining part of the matrix.
- Then we calculate the determinant of the remaining part which gives us the value of the minor of that particular element.

- If we denote the Cofactor using C_{ij}, then the cofactor of any element for $\mathbf{C_{ij} = M_{ij} \times (-1)^{i+j}}$

  Where,

  i is the number of rows for the element under consideration,

  j is the number of columns for the element under consideration, and

  $\mathbf{M_{ij}}$ is the minor of the element in the i^{th} row and j^{th} column.

What is Minor?

A minor of a particular element is obtained by eliminating the row and column of the matrix to which that particular element belongs and then finding the determinant of the remaining part. The matrix formed by combining all the minors is called the minor matrix. For example minor of the element a11 matrix $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ is calculated as:

$$M_{11} = det \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}$$
$$= 45 - 48$$
$$= -3$$

How to Find Cofactor of a Matrix?

In order to find a cofactor matrix we need to perform the following steps:

Step 1: Find the minor of each element of the matrix and make a minor matrix.

Step 2: Multiply each element in the minor matrix by $(-1)^{i+j}$.

Thus, we obtain the cofactor matrix.

Question 2:

ɔfactor of the element a_{11} of the matrix $A = \begin{bmatrix} 2 & -3 & 5 \\ 6 & 0 & p \\ 1 & 5 & -7 \end{bmatrix}$ is -20, then find the value of p.

Solution: Given matrix is:

$$A = \begin{bmatrix} 2 & -3 & 5 \\ 6 & 0 & p \\ 1 & 5 & -7 \end{bmatrix}$$

Using the formula of cofactor of an element,

$C_{ij} = (-1)^{i+j} \det(M_{ij})$

Cofactor of a_{11} is:

$C_{11} = (-1)^{1+1} \det(M_{11})$

$$-20 = \begin{vmatrix} 0 & p \\ 5 & -7 \end{vmatrix}$$

-20 = 0 – 5p

-20 = -5p

⇒ 5p = 20

⇒ p = 20/5

⇒ p = 4

Hence, the value of p is 4.

Question 1: Find the cofactor matrix of the matrix:

$$A = \begin{bmatrix} 1 & 9 & 3 \\ 2 & 5 & 4 \\ 3 & 7 & 8 \end{bmatrix}$$

Solution:

Given matrix is:

$$A = \begin{bmatrix} 1 & 9 & 3 \\ 2 & 5 & 4 \\ 3 & 7 & 8 \end{bmatrix}$$

Let M_{ij} be the minor of elements of the ith row and jth column.

Minor of the elements of matrix A are:

$$M_{11} = \begin{vmatrix} 5 & 4 \\ 7 & 8 \end{vmatrix} = 40 - 28 = 12$$

$$M_{12} = \begin{vmatrix} 2 & 4 \\ 3 & 8 \end{vmatrix} = 16 - 12 = 4$$

$$M_{13} = \begin{vmatrix} 2 & 5 \\ 3 & 7 \end{vmatrix} = 14 - 15 = -1$$

$$M_{21} = \begin{vmatrix} 9 & 3 \\ 7 & 8 \end{vmatrix} = 72 - 21 = 51$$

$$M_{22} = \begin{vmatrix} 1 & 3 \\ 3 & 8 \end{vmatrix} = 8 - 9 = -1$$

$$M_{23} = \begin{vmatrix} 1 & 9 \\ 3 & 7 \end{vmatrix} = 7 - 27 = -20$$

$$M_{31} = \begin{vmatrix} 9 & 3 \\ 5 & 4 \end{vmatrix} = 36 - 15 = 21$$

$$M_{32} = \begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} = 4 - 6 = -2$$

$$M_{33} = \begin{vmatrix} 1 & 9 \\ 2 & 5 \end{vmatrix} = 5 - 18 = -13$$

Matrix of cofactors of A is:

$$= \begin{bmatrix} +12 & -4 & +(-1) \\ -51 & +(-1) & -(-20) \\ +21 & -(-2) & +(-13) \end{bmatrix}$$

$$= \begin{bmatrix} 12 & -4 & -1 \\ -51 & -1 & 20 \\ 21 & 2 & -13 \end{bmatrix}$$

For a 2×2 Matrix

For a 2×2 matrix (2 rows and 2 columns):

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The determinant is:

$$|A| = ad - bc$$

"The determinant of A equals a times d minus b times c"

It is easy to remember when you think of a cross:

- Blue is positive (+ad),
- Red is negative (−bc)

Example: find the determinant of

$$C = \begin{bmatrix} 4 & 6 \\ 3 & 8 \end{bmatrix}$$

Answer:

$$|C| = 4 \times 8 - 6 \times 3$$
$$= 32 - 18$$
$$= 14$$

For a 3×3 Matrix

For a 3×3 matrix (3 rows and 3 columns):

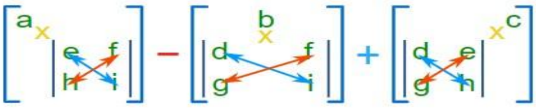
$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

The determinant is:

$$|A| = a(ei - fh) - b(di - fg) + c(dh - eg)$$

"The determinant of A equals ... etc"

It may look complicated, but **there is a pattern**:



To work out the determinant of a **3×3** matrix:

- Multiply **a** by the **determinant of the 2×2 matrix** that is **not in a**'s row or column.
- Likewise for **b**, and for **c**
- Sum them up, but remember the minus in front of the **b**

As a formula *(remember the vertical bars || mean "determinant of")*:

$$|A| = a \cdot \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \cdot \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \cdot \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

Example:

$$D = \begin{bmatrix} 6 & 1 & 1 \\ 4 & -2 & 5 \\ 2 & 8 & 7 \end{bmatrix}$$

$$|D| = 6\times(-2\times7 - 5\times8) - 1\times(4\times7 - 5\times2) + 1\times(4\times8 - (-2\times2))$$

$$= 6\times(-54) - 1\times(18) + 1\times(36)$$

$$= -306$$

Adjoint of a Matrix

- It is the transpose of the cofactor of the given matrix.
- To calculate the adjoint of a matrix, you need to find the cofactor matrix of the given matrix and then transpose it.
- *The key application or use of the adjoint of a matrix is to find the inverse of invertible matrices.*

Example: Find the Adjoint of $A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$.

the matrix.

s elements.

omes from after Step 2.

Solution:

Given matrix is $A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$

Step 1: *Find the Cofactor of each element.*

Cofactor of element at A[1,1]: 5

Cofactor of element at A[1,2]: -4

Cofactor of element at A[2,1]: -3

Cofactor of element at A[2,2]: 2

Step 2: *Create matrix from Cofactors*

i.e., $\begin{bmatrix} 5 & -4 \\ -3 & 2 \end{bmatrix}$

Step 3: *Transpose of Cofactor matrix,*

$\mathbf{Adj(A)} = \begin{bmatrix} 5 & -3 \\ -4 & 2 \end{bmatrix}$

Example 1: Find the Adjoint of the given matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 7 & 4 & 5 \\ 6 & 8 & 9 \end{bmatrix}$.

Solution:

Step 1: To find the cofactor of each element

To find the cofactor of each element, we have to delete the row and column of each element one by one and take the present elements after deleting.

Cofactor of elements at A[0,0] = 1 : $+\begin{bmatrix} 4 & 5 \\ 8 & 9 \end{bmatrix}$ *= +(4×9 – 8×5) = -4*

Cofactor of elements at A[0,1] = 2 : $-\begin{bmatrix} 7 & 5 \\ 6 & 9 \end{bmatrix}$ *= -(7×9 – 6×5) = -33*

Cofactor of elements at A[0,2] = 3 : $+\begin{bmatrix} 7 & 4 \\ 6 & 8 \end{bmatrix}$ *= +(7×8 – 6×4) = 32*

Cofactor of elements at A[2,0] = 7 : $-\begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix}$ *= -(2×9 – 8×3) = 6*

Cofactor of elements at A[2,1] = 4 : $+\begin{bmatrix} 1 & 3 \\ 6 & 9 \end{bmatrix}$ *= +(1×9 – 6×3) = -9*

Cofactor of elements at A[2,2] = 5 : $-\begin{bmatrix} 1 & 2 \\ 6 & 8 \end{bmatrix}$ *= -(1×8 – 6×2) = 4*

Cofactor of elements at A[3,0] = 6 : $+\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$ *= +(2×5 – 4×3) = -2*

Cofactor of elements at A[3,1] = 8 : $-\begin{bmatrix} 1 & 3 \\ 7 & 5 \end{bmatrix}$ *= -(1×5 – 7×3) = 16*

Cofactor of elements at A[3,2] = 9 : $+\begin{bmatrix} 1 & 2 \\ 7 & 4 \end{bmatrix}$ *= +(1×4 – 7×2) = -10*

The matrix looks like with the cofactors:

$$A = \begin{bmatrix} +\begin{bmatrix} 4 & 5 \\ 8 & 9 \end{bmatrix} & -\begin{bmatrix} 7 & 5 \\ 6 & 9 \end{bmatrix} & +\begin{bmatrix} 7 & 4 \\ 6 & 8 \end{bmatrix} \\ -\begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix} & +\begin{bmatrix} 1 & 3 \\ 6 & 9 \end{bmatrix} & -\begin{bmatrix} 1 & 2 \\ 6 & 8 \end{bmatrix} \\ +\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} & -\begin{bmatrix} 1 & 3 \\ 7 & 5 \end{bmatrix} & +\begin{bmatrix} 1 & 2 \\ 7 & 4 \end{bmatrix} \end{bmatrix}$$

The final cofactor matrix:

$$A = \begin{bmatrix} -4 & -33 & 32 \\ 6 & -9 & 4 \\ -2 & 16 & -10 \end{bmatrix}$$

Step 2: Find the transpose of the matrix obtained in step 1

$$adj(A) = \begin{bmatrix} -4 & 6 & -2 \\ -33 & -9 & 16 \\ 32 & 4 & -10 \end{bmatrix}$$

This is the ***Adjoint of the matrix.***

Example: Find the Adjoint of $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

Solution:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Step 1: *Find the Cofactor of each element.*

$$C_{12} = (-1)^{1+2} \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} = -(36 - 42) = 6$$
$$C_{13} = (-1)^{1+3} \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = 3 - 28 = -25$$
$$C_{21} = (-1)^{2+1} \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} = -(18 - 24) = 6$$
$$C_{22} = (-1)^{2+2} \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} = 9 - 21 = -12$$
$$C_{23} = (-1)^{2+3} \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} = -(8 - 14) = 6$$
$$C_{31} = (-1)^{3+1} \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = 12 - 15 = -3$$
$$C_{32} = (-1)^{3+2} \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = -(6 - 12) = 6$$
$$C_{33} = (-1)^{3+3} \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = 5 - 8 = -3$$

Step 2: *Create matrix from Cofactors*

$$C = \begin{bmatrix} -3 & 6 & -25 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{bmatrix}$$

Step 3: *Transpose of Matrix C to adjoint of given matrix.*

$$\text{adj}(A) = C^T = \begin{bmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -25 & 6 & -3 \end{bmatrix}$$

Which is adjoint of given matrix A.

Inverse of a Matrix

Finding the inverse is one of the important applications of the Adjoint of the Matrix. To find the inverse of a

Matrix using Adjoint we can use the following steps:

Step 1: Find the determinant of the matrix.

Step 2: If the determinant is zero, then the matrix is not invertible, and there is no inverse.

Step 3: If the determinant is non-zero, then find the adjoint of the matrix.

Step 4: **Divide the adjoint of the matrix by the determinant of a matrix.**

Step 5: The result of Step 4 is the Inverse of given Matrix.

The inverse of a matrix is obtained by dividing its adjoint by its determinant. That is, if A is a square matrix and det(A) is non-zero, then

$$A^{-1} = adj(A)/det(A)$$

Question 2: Find out the inverse of

$$\begin{bmatrix} 1 & -1 & 2 \\ 4 & 0 & 6 \\ 0 & 1 & -1 \end{bmatrix}$$

?

Solution:

Let A =

$$\begin{bmatrix} 1 & -1 & 2 \\ 4 & 0 & 6 \\ 0 & 1 & -1 \end{bmatrix}$$

be the given matrix.

$$A^{-1} = \frac{adj(A)}{|A|}$$

To find out the adj(A), first we have to find out cofactor(A).

a_{11} = -6, a_{12} = 4, a_{13} = 4

a_{21} = 1, a_{22} = -1, a_{23} = -1

a_{13} = -6, a_{32} = 2, a_{33} = 4So, cofactor(A) =

$$\begin{bmatrix} -6 & 4 & 4 \\ 1 & -1 & -1 \\ -6 & 2 & 4 \end{bmatrix}$$

adj(A) =

$[cofactor(A)]^T$

adj(A) =

$$[cofactor(A)]^T = \begin{bmatrix} -6 & 4 & 4 \\ 1 & -1 & -1 \\ -6 & 2 & 4 \end{bmatrix}^T$$

adj(A) =

$$\begin{bmatrix} -6 & 1 & -6 \\ 4 & -1 & 2 \\ 4 & -1 & 4 \end{bmatrix}$$

Then, | A | = 1(0−6)+1(−4−0)+2(4−0) = −6−4+8 = −2

$$A^{-1} = \frac{adj(A)}{|A|} = \frac{\begin{bmatrix} -6 & 1 & -6 \\ 4 & -1 & 2 \\ 4 & -1 & 4 \end{bmatrix}}{-2}$$

$$A^{-1} = \begin{bmatrix} 3 & -\frac{1}{2} & 3 \\ -2 & \frac{1}{2} & -1 \\ -2 & \frac{1}{2} & -2 \end{bmatrix}$$

Matrix Norm

The Norm of a matrix is a real number which is a measure of the magnitude of the Matrix.

The norm of a square matrix A is a non-negative real number denoted $\|A\|$. There are several different ways of defining a matrix norm, but they all share the following properties:

1. $\|A\| \geq 0$ for any square matrix A.
2. $\|A\| = 0$ *if and only if* the matrix $A = 0$.
3. $\|kA\| = |k| \ \|A\|$, for any scalar k.
4. $\|A + B\| \leq \|A\| + \|B\|$.
5. $\|AB\| \leq \|A\| \ \ \|B\|$.

The 1-norm

$$\|A\|_1 = \max_{1 \leq j \leq n} \left(\sum_{i=1}^{n} |a_{ij}| \right)$$

(the maximum absolute column sum). Put simply we sum the absolute values down each column and then take the biggest answer.

Example Calculate the 1-norm of $A = \begin{bmatrix} 1 & -7 \\ -2 & -3 \end{bmatrix}$.

Solution

The absolute column sums of A are $1 + |-2| = 3$ and $|-7| + |-3| = 10$. The larger of these is 10 and therefore $\|A\|_1 = 10$.

Infinity -Norm (The Infinity-Norm of a Square Matrix is the maximum of the absolute Row Sums)

The inifinity-norm

$$\|A\|_\infty = \max_{1\le i\le n}\left(\sum_{j=1}^{n}|a_{ij}|\right)$$

(the maximum absolute row sum). Put simply we sum the absolute values along each row and then take the biggest answer.

Example Calculate the infinity-norm of $A = \begin{bmatrix} 1 & -7 \\ -2 & -3 \end{bmatrix}$.

Solution

The absolute row sums of A are $1 + |-7| = 8$ and $|-2| + |-3| = 5$. The larger of these is 8 and therefore $\|A\|_\infty = 8$.

Euclidean Norm (The Euclidean Norm of a Square Matrix is the square root of the sum of all the squares of the elements

The Euclidean norm

$$\|A\|_E = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij})^2}$$

(the square root of the sum of all the squares). This is similar to ordinary "Pythagorean" length where the size of a vector is found by taking the square root of the sum of the squares of all the elements.

Example Calculate the Euclidean norm of $A = \begin{bmatrix} 1 & -7 \\ -2 & -3 \end{bmatrix}$.

Solution

$$\begin{aligned} \|A\|_E &= \sqrt{1^2 + (-7)^2 + (-2)^2 + (-3)^2} \\ &= \sqrt{1 + 49 + 4 + 9} \\ &= \sqrt{63} \approx 7.937. \end{aligned}$$

Matrix Norm

Calculate the norms indicated of these matrices

$$A = \begin{bmatrix} 2 & -8 \\ 3 & 1 \end{bmatrix} \quad \text{(1-norm)}, \qquad B = \begin{bmatrix} 3 & 6 & -1 \\ 3 & 1 & 0 \\ 2 & 4 & -7 \end{bmatrix} \quad \text{(infinity-norm)},$$

$$C = \begin{bmatrix} 1 & 7 & 3 \\ 4 & -2 & -2 \\ -2 & -1 & 1 \end{bmatrix} \quad \text{(Euclidean-norm)}.$$

Vector

- **Scalar** will have only **magnitude** .

  Eg : An Airplane moves 300 miles per hourn

- Whereas **Vector** has **both Magnitude and direction .**

  Eg : An Airplane Moves 300 miles per hour in the direction 45 degree North of east

- **The vector value changes when a**

Eigen Values and Eigen Vectors

n eigenvector is a non-zero vector that remains unchanged in direction when a linear transformation is applied to it, though it may be scaled by a scalar factor, known as its corresponding eigenvalue.

- ❖ We know that the vectors change its magnitude and direction when some linear transformation is applied to it.
- ❖ But some vectors do not change much (or in other words they change at most by its scale factor) even after the application of transformations on them.
- ❖ Such vectors are called eigenvectors.
- ❖ We have to find eigenvalues always before finding the eigenvectors.

How to Find Eigenvectors of a 2 × 2 matrix?

To find the eigenvectors, we first have to compute the eigenvalues using the above-mentioned steps. Let us understand the process by an example.

Example: Find the eigenvalues and eigenvectors of matrix $A = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$.

Solution:

Let λ represent the eigenvalue(s) and **v** = $\begin{bmatrix} x \\ y \end{bmatrix}$ represent the eigenvector(s).

$)\vec{x} = \vec{0} \quad or \quad (\lambda I - A)\vec{x} = \vec{0}$

Then the characteristic equation is:

$|A - \lambda I| = 0$

$$\begin{vmatrix} 5-\lambda & 4 \\ 1 & 2-\lambda \end{vmatrix} = 0$$

$(5 - \lambda)(2 - \lambda) - (4)(1) = 0$

$10 - 5\lambda - 2\lambda + \lambda^2 - 4 = 0$

$\lambda^2 - 7\lambda + 6 = 0$

$(\lambda - 6)(\lambda - 1) = 0$

$\lambda = 6, \lambda = 1.$

Thus, the eigenvalues are 1 and 6. Let us find the corresponding eigenvector to each eigenvalue in each case.

When $\lambda = 1$:

Substitute $\lambda = 1$ in the equation:

$(A - \lambda I)\,\mathbf{v} = O$

$$\begin{bmatrix} 5-1 & 4 \\ 1 & 2-1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

To solve this system, we apply the elementary row transformations (Alternatively, we can use Cramer's rule as well) on the coefficient matrix:

Apply $R_2 \to 4R_2 - R_1$,

$$\begin{bmatrix} 4 & 4 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Expanding as equations,

$4x + 4y = 0$

We have one equation in two variables. So assume $y = t$. Then

$4x + 4t = 0 \Rightarrow 4x = -4t \Rightarrow x = -t$

So the solution is, $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -t \\ t \end{bmatrix} = t\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Thus, the eigenvector is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

When $\lambda = 6$:

Substitute $\lambda = 1$ in the equation:

$(A - \lambda I)\, \mathbf{v} = O$

$$\begin{bmatrix} 5-6 & 4 \\ 1 & 2-6 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

To solve this system, we apply the elementary row transformations (Alternatively, we can find Cramer's rule as well) on the coefficient matrix:

Apply $R_2 \to R_2 + R_1$,

$$\begin{bmatrix} -1 & 4 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Expanding as equations,

$-x + 4y = 0$

Let $y = t$. Then $-x + 4t = 0 \Rightarrow x = 4t$.

So the solution is, $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4t \\ t \end{bmatrix} = t\begin{bmatrix} 4 \\ 1 \end{bmatrix}$

Thus, the eigenvector is $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$.

Thus, the eigenvectors of the given 2×2 matrix are $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$.

How to Find Eigenvectors of a 3 × 3 matrix?

We follow the same steps as above for the 3×3 matrix as well. Here is an example.

Example: Find the eigenvectors of 3×3 matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

Solution:

Let λ be the eigenvalue and $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ be the eigenvector of A.

Finding eigenvalues:

The characteristic equation is:

$|A - \lambda I| = 0$

$$\begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} = 0$$

Calculating the determinant, we get

$-\lambda^3 + 3\lambda^2 = 0$

$\lambda^2 (-\lambda+3) = 0$

$\lambda = 0, \lambda = 3.$

Thus, the eigenvalues are 0 and 3. Let us find the corresponding eigenvectors.

When $\lambda = 0$:

$(A - \lambda I)\,\mathbf{v} = 0$

$$\begin{bmatrix} 1-0 & 1 & 1 \\ 1 & 1-0 & 1 \\ 1 & 1 & 1-0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Apply $R_2 \to R_2 - R_1$ and $R_3 \to R_3 - R_1$,

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Expanding the above matrix, we get

$x + y + z = 0$

We have only one equation with two unknowns. So let us assume two of the variables to be $y = t_1$ and $z = t_2$. Then the above equation becomes:

$x + t_1 + t_2 = 0 \Rightarrow x = - t_1 - t_2$.

Thus, the eigenvector is:

$\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ = \(\left[\begin{array}{l}

$- t_1 - t_2$ \\

t_1 \\t_2

\end{array}\right]\) $= t_1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

Thus, the eigenvectors that correspond to $\lambda = 0$ are $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

When $\lambda = 3$:

$(A - \lambda I)\,\mathbf{v} = 0$

$$\begin{bmatrix} 1-3 & 1 & 1 \\ 1 & 1-3 & 1 \\ 1 & 1 & 1-3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Apply $R_2 \to 2R_2 + R_1$ and $R_3 \to 2R_3 + R_1$,

$$\begin{bmatrix} -2 & 1 & 1 \\ 0 & -3 & 3 \\ 0 & 3 & -3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Now, apply $R_3 \to R_3 + R_2$,

$$\begin{bmatrix} -2 & 1 & 1 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Expanding the above matrix, we get

$-2x + y + z = 0$

$-3y + 3z = 0$

Let $z = t$. Then

$-3y + 3z = 0 \Rightarrow z = y = t$.

$-2x + y + z = 0 \Rightarrow -2x + t + t = 0 \Rightarrow x = t$.

Thus, the eigenvector is:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} t \\ t \\ t \end{bmatrix} = t\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Thus, the eigenvector that correspond to $\lambda = 3$ is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Thus, the eigenvectors of the given 3×3 matrix are $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Singular Value Decomposition

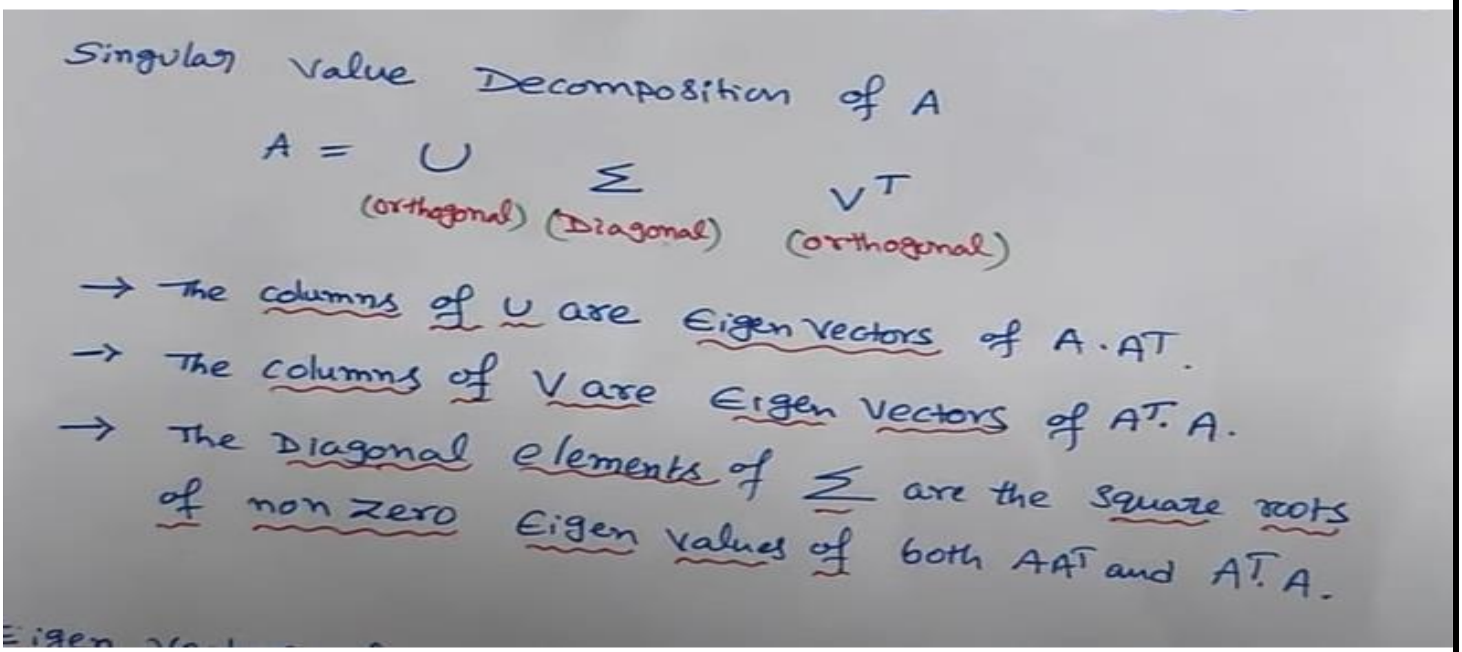
Singular Value Decomposition (SVD) is a fundamental concept in linear algebra and numerical analysis. It decomposes a matrix into three simpler matrices, providing insights into its structure and allowing for various computations and analyses.

Given a matrix A, its Singular Value Decomposition is represented as:

$A = U\Sigma V^T$

Where:

- U is an orthogonal matrix containing the left singular vectors.
- Σ is a diagonal matrix containing the singular values.
- V^T is the transpose of an orthogonal matrix containing the right singular vectors.



SVD has numerous applications, including:

1. **Dimensionality Reduction**: SVD can be used to reduce the dimensionality of data while preserving important information.

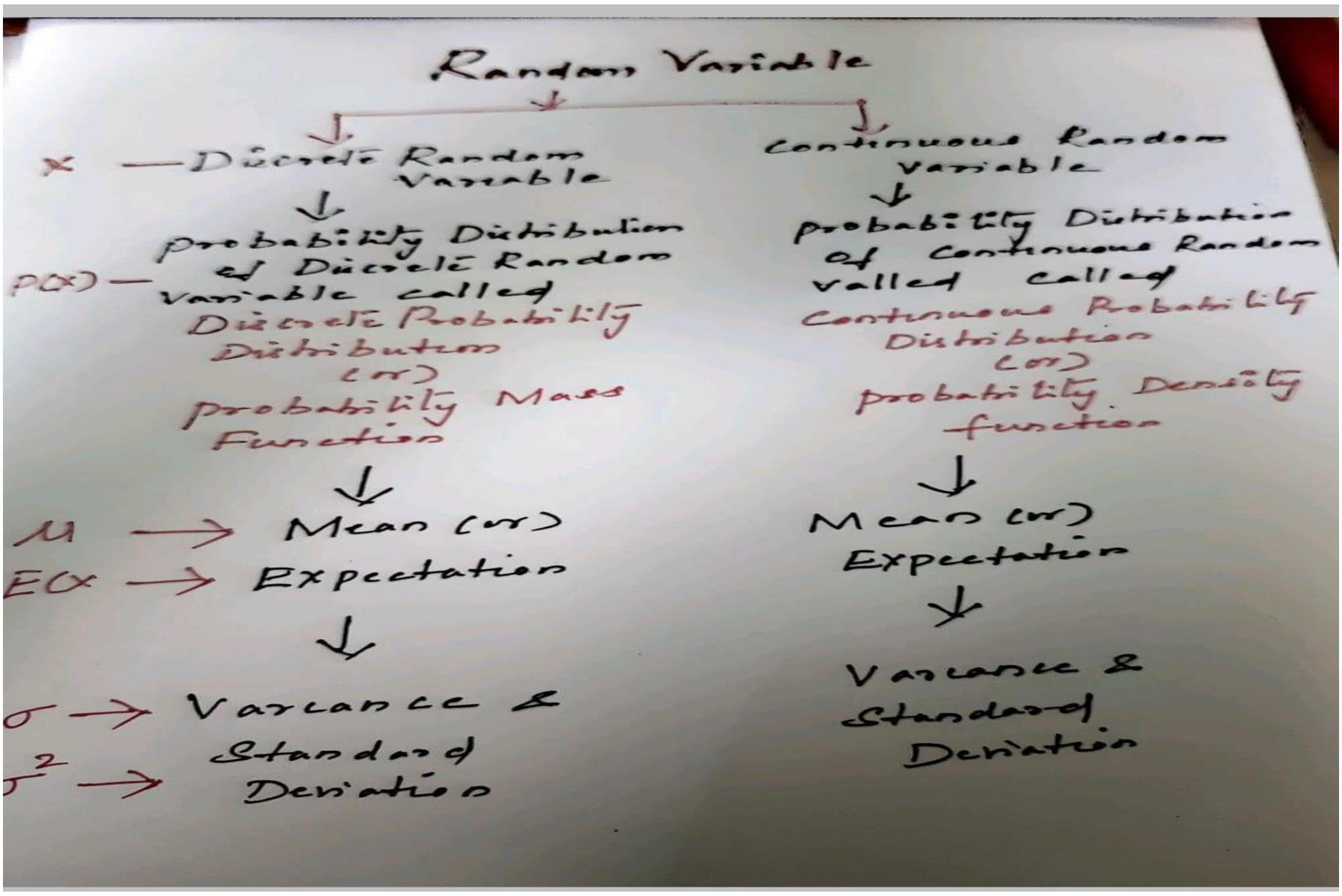
2. **Matrix Approximation**: By truncating the singular values and corresponding singular vectors, you can approximate the original matrix.

3. **Pseudoinverse**: SVD can be used to compute the pseudoinverse of a matrix, which is useful in solving linear least squares problems.

4. **Principal Component Analysis (PCA)**: SVD is closely related to PCA, where the singular vectors represent the principal components of the data.

5. **Image Compression**: SVD can be used for compressing images by representing them in terms of their singular vectors and values.

STATISTICS



Introduction to Random Variable

Experiment: Is any Physical action or process that is observed and the result is noted

Experiment: An Experiment in which all possible outcomes are known and the exact outcomes cannot be predicted in advance is called "Random Experiment".

- Eg1 : Tossing a coin is an example of Random experiment because in this experiment all possible outcomes {H,T} are known but exact outcome cannot be predicted.
- Eg2: Boiling Water – Outcome is Predicted i.e it gets Evaporated . So its not an Random experiment

Sample Space

- Set of all Possible outcomes of an Experiment is called "Sample Space".
- Examples :

✓ Sample Space of "Tossing a Coin experiment is **S = {H,T}**

✓ Sample Space of "Tossing 2 Coins Experiment is **S = {HH,HT,TH,TT}**

✓ Sample Space of " Rolling a Dice Experiment is **S = {1,2,3,4,5,6}**

Random Variable

✓ Sample Space of "Tossing a Coin experiment is **S = {H,T}**

✓ A Radom Variable is a function that assigns a numerical value to each sample point in the sample space.

✓ Random Variables are denoted by capital Letters. Eg : X, Y etc.

✓ Example :

✓ **Experiment :** "**Tossing 2 Coins** "

- **Sample Space :** **S = {HH,HT,TH,TT}**
- **Random Variable** : **X { no. of Heads }**
- **So, X = {2,1,1,0} (Numerical Values assigned to HH ,HT, TH and TT)**

Types of Random Variable

Random variable are classified into 2 Types

1) Discrete Random variable

2) Continuous Random variable

1) Discrete Random Variable :

✓ If a Random variable takes only a finite or a countable number of values , it is called Discrete Random Variable.

✓ Example : when 2 coins are tossed , the sample space S = {HH,HT,TH,TT} and if the function is No.of Heads, then the variable can take finite or countable values like 2,1,0 and such variable is called Discrete Random variable.

2) Continuous Random Variable :

- ✓ A Random Variable X which can take any Value between certain Interval is called Continuous Random Variable.
- ✓ Example : The Height of students in a particular class lies between 4 to 6 feet. We write this as

$$X = \{x | 4 \leq x \leq 6\}$$

Probability Distribution of a Random Variable

1) Probability Distribution

- ✓ Distribution of Probabilities of each Outcome of the random variable is called Probability Distribution
- ✓ Example : Experiment : 2 coins are tossed , Sample Space S ={HH,HT,TH,TT} function : No.of Heads, then the Discrete Random variable X = {2,1,0}

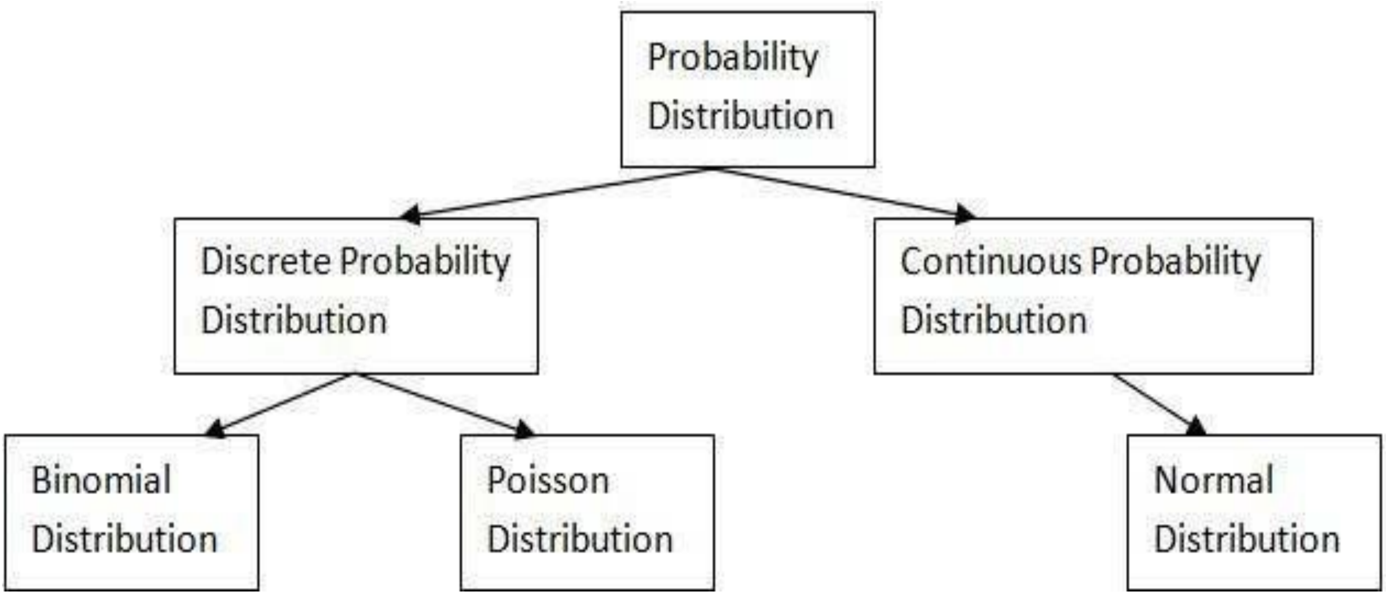
Outcome of the Variable Corresponding Probabilities

Propabilitity Distribution of the Discrete Probability Mass Function

X=x	0	1	2
f(x) = P(X=x)	1/4	2/4	1/4

Probability Distribution of the Discrete Random variable and f(X) is called Probability Mass Function

Types of Probability Distribution



Discrete Probability Distribution Function or Probability Mass Function:

- Probability Distribution of a Discrete Random Variable is called **Discrete Probability Distribution**.
- The Discrete Probability Distribution Function **P(X = x) or Pi** is called the **Probability Mass Function (PMF)**.
- The Properties of Probability Mass Function are

i. $P_i \geq 0$ *for all i, i.e* P_i'*s are all non − negative.*

ii. $\sum P_i = P_1 + P_2 + \cdots \ldots\ldots\ldots + P_n = 1$ *i.e the total probability is one*

1) TO check the given function is Probability Mass Function or Not.

<u>Solution</u> : Experiment = Tossing 2 coins

Sample space = {HH, HT, TH, TT}

X = {No. of Heads}

X = {2, 1, 1, 0}

↓

Discrete Random Variable because It takes finite (or) countable values.

The Probability Distribution of the outcomes of the Random variable are

x	0	1	2
$P(X=x)$ (or) $f(x)$ or P_i	$1/2$ ↓ Prob of getting 0 heads out of 4 outcomes	$2/4$	$1/4$

(is a PMF → $f(x)$)

TO check given function is PMF or not, the $f(x)$ should satisfy 2 conditions

① Each $P_i > 0$

② Sum of all probabilities should be Equal to 1

In our Example, all three P_i are greater than zero and

$1/2 + 2/4 + 1/4 = 1$. So $f(x)$ is a PMF

Mean (or) Expectation, Variance and Standard Deviation of a Discrete Random Variable.

Mean: The Mean of a Discrete Random Variable is the central Value or Average of its Corresponding PMF.
It is also called as Expected value.
It is computed using the formula

$$\mu = E(x) = \sum x \cdot P(x)$$ where

x - outcome
P(x) - probability of the outcome

Variance: It is a Measure used to find how the data is spread w.r.t the Mean.
Variance is denoted by σ^2 and Standard Deviation by σ

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

$$\sigma = \sqrt{E(x^2) - (E(x))^2}$$

1) Find Mean, Variance and SD of the Random Variable X that represents No. of heads in tossing 3 coins.

Sol : Exp : Tossing 3 coins

S : {HHH, HHT, HTT, HTH, TTT, TTH, THH, THT}

X : No. of heads

: {0, 1, 2, 3}

probability Distribution of X

X	0	1	2	
P(x)	1/8	3/8	3/8	1/8

To find Mean

X	P(X)	E(X) = X.P(X)
0	$\frac{1}{8}$	0
1	$\frac{3}{8}$	3/8
2	$\frac{3}{8}$	3/4
3	$\frac{1}{8}$	3/8

$\Sigma x.p(x) = 1.5$ or E(X)

To find Variance

X^2	P(X)	$E(X^2) = X^2.P(X)$
0	1/8	0
1	3/8	3/8
4	3/8	3/2
9	1/8	9/8

$\Sigma X^2.P(X)$ or $E(X^2)$ = 3

$$\text{Variance } \sigma^2 = E(X^2) - [E(X)]^2$$
$$= 3 - (1.5)^2$$
$$= 3 - 2.25$$
$$= 0.75$$

$SD = \sigma = \sqrt{0.75} = 0.86$

Problem 2:

The Probabilities that a customer will buy 1,2,3,4 or 5 items in a grocery store are 3/10,1/10,1/10,2/10,3/10 respectively. What is the average number of items that a customer will buy

Solution :

The Probability Distribution Table for the Above problem is

X	P(x)	x.Px)
1	3/10	3/10
2	1/10	2/10
3	1/10	3/10
4	2/10	8/10
5	3/10	15/10
E x.P(X) (or) E(x)		**31/10 = 3.1**

The Average Number of Items the customer can buy is 3.1

Problem 3:

The Following Probability Distribution tells us the Probability that a certain Soccer Team Scores certain number of goals in a given time. **Find Mean, Variance and SD**

Goals x	Probabilities P(x)
0	0.18
1	0.34
2	0.34
3	0.11
4	0.02

Solution :

The Probability Distribution Table for the Above problem is

Goals x	Probabilities P(x)	x.P(x)	X^2	$X^2.P(X)$
0	0.18	0	0	0
1	0.34	0.34	1	0.34
2	0.34	0.68	4	1.36
3	0.11	0.33	9	0.99
4	0.02	0.08	16	0.32
$E(X) = \sum x.P(X) = 1.43$			$E(X^2) = \sum X^2.P(X) = 3.01$	

Mean $\mu = E(X) = 1.43$

Variance $\sigma^2 = E(X^2) - (E(X))^2$

$= 3.01 - (1.43)^2$

$= 0.96$

Problem 4:

If x: No. of Colleges to which you apply

x = {0,1,2,3,4}

f(x) = {
- Kx, x=0 or 1
- 2Kx, x=2
- K(5-x), x=3 or 4
- 0, x>4

}

1) Find K
2) Probability of getting Admission in a College
3) Atmost 2 Colleges
4) Atleast 2 Colleges

Solution :

No.of Colleges to which you can apply X	Probability of Getting Admission P(x)
0	Kx
1	Kx
2	2Kx
3	K(5-x)
4	K(5-x)
>4	0

1) Find K

P(X=0)+P(X=1)+ P(X=2)+ P(X=3)+ P(X=4)+ P(X>4) = 1

K(0) + K(1) + 2K(2) + K(5-3) + k(5-4)+0 = 1

0 + K + 4K + 2K + k =1

8k = 1,

K = 1/8

2) Probability of Getting Admission in a college - **Means Find P(x=1)**

P(x=1)

K(1)

1/8

3) Atmost 2 Colleges (Max. 2 Colleges) - **Means Find P(x≤ 2)**

P(x=0)+P(x=1)+P(x=2)

K0) + K(1) + 2k(2)

0 + K + 4K

5K , 5*1/8 = 5/8

4) Atleast 2 Colleges (Min. 2 Colleges) - **Means Find P(x ≥ 2)**

P(x=2) + P(x=3) + P(x=4) + p(x>4)

2K(2) + K(5-3) + k(5-4) + 0

4K + 2K + K

7K , 7*1/8 = 7/8

Binomial Experiments

A **binomial experiment** is a probability experiment that satisfies the following conditions.

1. The experiment is repeated for a fixed number of trials, where each trial is independent of other trials.
2. There are only two possible outcomes of interest for each trial. The outcomes can be classified as a success (S) or as a failure (F).
3. The probability of a success $P(S)$ is the same for each trial.
4. The random variable x counts the number of successful trials.

Notation for Binomial Experiments

Symbol	*Description*
n	The number of times a trial is repeated.
$p = P(S)$	The probability of success in a single trial.
$q = P(F)$	The probability of failure in a single trial. ($q = 1 - p$)
x	The random variable represents a count of the number of successes in n trials: $x = 0, 1, 2, 3, \ldots, n$.

Binomial Experiments

Example:
Decide whether the experiment is a binomial experiment. If it is, specify the values of n, p, and q, and list the possible values of the random variable x. If it is not a binomial experiment, explain why.

- You randomly select a card from a deck of cards, and note if the card is an Ace. You then put the card back and repeat this process 8 times.

This is a binomial experiment. Each of the 8 selections represent an independent trial because the card is replaced before the next one is drawn. There are only two possible outcomes: either the card is an Ace or not.

$n = 8 \quad p = \frac{4}{52} = \frac{1}{13} \quad q = 1 - \frac{1}{13} = \frac{12}{13} \quad x = 0, 1, 2, 3, 4, 5, 6, 7, 8$

Binomial Experiments

Example:
Decide whether the experiment is a binomial experiment. If it is, specify the values of n, p, and q, and list the possible values of the random variable x. If it is not a binomial experiment, explain why.

- You roll a die 10 times and note the number the die lands on.

This is not a binomial experiment. While each trial (roll) is independent, there are more than two possible outcomes: 1, 2, 3, 4, 5, and 6.

Binomial Probability Formula

In a binomial experiment, the probability of exactly x successes in n trials is

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

Example:

A bag contains 10 chips. 3 of the chips are red, 5 of the chips are white, and 2 of the chips are blue. Three chips are selected, with replacement. Find the probability that you select exactly one red chip.

$p =$ the probability of selecting a red chip $= \frac{3}{10} = 0.3$

$q = 1 - p = 0.7$

$n = 3$

$x = 1$

$$P(1) = {}_3C_1(0.3)^1(0.7)^2$$
$$= 3(0.3)(0.49)$$
$$= 0.441$$

Binomial Probability Distribution

Example:

A bag contains 10 chips. 3 of the chips are red, 5 of the chips are white, and 2 of the chips are blue. Four chips are selected, with replacement. Create a probability distribution for the number of red chips selected.

$p =$ the probability of selecting a red chip $= \frac{3}{10} = 0.3$

$q = 1 - p = 0.7$

$n = 4$

$x = 0, 1, 2, 3, 4$

x	$P(x)$
0	0.240
1	0.412
2	0.265
3	0.076
4	0.008

The binomial probability formula is used to find each probability.

Mean, Variance and Standard Deviation

Population Parameters of a Binomial Distribution

Mean: $\mu = np$

Variance: $\sigma^2 = npq$

Standard deviation: $\sigma = \sqrt{npq}$

Example:

One out of 5 students at a local college say that they skip breakfast in the morning. Find the mean, variance and standard deviation if 10 students are randomly selected.

$n = 10$

$p = \frac{1}{5} = 0.2$

$q = 0.8$

$\mu = np = 10(0.2) = 2$

$\sigma^2 = npq = (10)(0.2)(0.8) = 1.6$

$\sigma = \sqrt{npq} = \sqrt{1.6} \approx 1.3$

Poisson Distribution

The **Poisson distribution** is a discrete probability distribution of a random variable x that satisfies the following conditions.

1. The experiment consists of counting the number of times an event, x, occurs in a given interval. The interval can be an interval of time, area, or volume.
2. The probability of the event occurring is the same for each interval.
3. The number of occurrences in one interval is independent of the number of occurrences in other intervals.

The probability of exactly x occurrences in an interval is

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

where $e \approx 2.71818$ and μ is the mean number of occurrences.

Poisson Distribution

Example:

The mean number of power outages in the city of Brunswick is 4 per year. Find the probability that in a given year,

a.) there are exactly 3 outages,

b.) there are more than 3 outages.

a.) $\mu = 4, \quad x = 3$

$$P(3) = \frac{4^3(2.71828)^{-4}}{3!}$$

$$\approx 0.195$$

b.) $P(\text{more than } 3)$

$$= 1 - P(x \leq 3)$$

$$= 1 - [P(3) + P(2) + P(1) + P(0)]$$

$$= 1 - (0.195 + 0.147 + 0.073 + 0.018)$$

$$\approx 0.567$$

Mean and variance of a Poisson distribution

The Poisson distribution has only one parameter, called λ.

- The mean of a Poisson distribution is λ.
- The variance of a Poisson distribution is also λ.

In most distributions, the mean is represented by μ (mu) and the variance is represented by σ^2 (sigma squared). Because these two parameters are the same in a Poisson distribution, we use the λ symbol to represent both.

Poisson distribution formula

The probability mass function of the Poisson distribution is:

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

Where:

- X is a random variable following a Poisson distribution
- k is the number of times an event occurs
- $P(X = k)$ is the probability that an event will occur *k* times
- e is Euler's constant (approximately 2.718)
- λ is the average number of times an event occurs
- ! is the factorial function

Example: Applying the Poisson distribution formula

An average of 0.61 soldiers died by horse kicks per year in each Prussian army corps. You want to calculate the probability that exactly two soldiers died in the VII Army Corps in 1898, assuming that the number of horse kick deaths per year follows a Poisson distribution.

Calculation

The specific army corps (VII Army Corps) and year (1898) don't matter because the probability is constant.

k = 2 deaths by horse kick

λ = 0.61 deaths by horse kick per year

e = 2.718

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

$$P(X = 2) = \frac{(2.718^{-0.61})(0.61^2)}{2!}$$

$$P(X = 2) = \frac{(0.54339)(0.3721)}{2}$$

$$P(X = 2) = 0.101$$

The probability that exactly two soldiers died in the VII Army Corps in 1898 is 0.101.

Practice questions

1 → At a small walk-in clinic, an average of five patients arrive at the clinic per hour during opening hours. What is the probability that exactly three patients will arrive in the next hour?

k = 3 patients
λ = 5 patients/hour
e = 2.718

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

$$P(X = 3) = \frac{(2.718^{-5})(5^{3})}{3!}$$

$$P(X = 3) = \frac{(0.00674)(125)}{6}$$

$$P(X = 3) = 0.14$$

2 → If you receive an average of two emails per week from your statistics professor, what is the probability that you will receive exactly one email from your statistics professor on Monday?

k = 1 email
λ = 2 emails per week = 2/7 emails per day = 0.286 emails per day
(The question asks for the probability in units of days, so λ needs to be converted from units of weeks to days.)
e = 2.718

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

$$P(X = 1) = \frac{(2.718^{-0.286})(0.286^{1})}{1!}$$

$$P(X = 1) = \frac{(0.75128)(0.286)}{1}$$

$$P(X = 1) = 0.215$$

3 → Over the last 300 years, there were 87 floods in Statsville. Assuming that the number of floods per year follows a Poisson distribution, what is the probability that there will be no floods in Statsville next year?

k = 0 floods
λ = 87 floods per 300 years = 87/300 floods per year = 0.29 floods per year
e = 2.718

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

$$P(X = 0) = \frac{(2.718^{-0.29})(0.29^0)}{0!}$$

$$P(X = 0) = \frac{(0.74829)(1)}{1}$$

$$P(X = 0) = 0.748$$

Discrete Random Variable	Continuous Random Variable
① if a Random variable takes only a finite or Countable number of values, then the variable is called DRV.	if a Random variable takes any value b/w certain interval, then the variable is called CRV.
② Probability Distribution of a Discrete Random variable is called Discrete Probability Distribution or Probability Mass Function	Probability Distribution of a Continuous Random variable is called continuous Probability Distribution or Probability Density Function.
③ Properties of PMF are ① $P_i \geq 0$ for all i, i.e P_i's are all non-Negative ② $\sum P_i = 1$ i.e sum of Probabilities must be Equal to 1	Properties of PDF is ① $f(x) \geq 0$ ② $\int_{-\infty}^{\infty} f(x) \cdot d(x) = 1$
Discrete Probability Distribution → Binomial Distribution, Poisson Distribution	Continuous Probability Distribution ↓ Normal Distribution

Eg: check whether the following function can serve as PMF

$$f(x) = (x-2)/2 \text{ for } x = 1,2,3,4$$

Solution: Condition 1: $f(x) \geq 0$ for Every x but

$f(1) = (1-2)/2 = -1/2$ is a negative.

So, the function does not satisfy first condition so the $f(x) = (x-2)/2$ is not a PMF.

Eg: Check whether the following function can serve as PDF

$$f(x) = \frac{1}{3}x^2 \text{ for } x \geq 0$$

Solution: cond 1: $f(x) \geq 0$

cond 2: $\int_{-\infty}^{\infty} f(x) \cdot d(x) = 1$

The above function Satisfies Condition 1 since x^2 is always non-negative and $1/3$ is a positive value.

To check second Condition

$$\int_0^{\infty} f(x) \cdot dx = 1 \Rightarrow \int_0^{\infty} \frac{1}{3} \cdot x^2 \cdot dx = 1$$

$$\Rightarrow \frac{1}{3}\int_0^{\infty} x^2 \cdot dx = 1 \Rightarrow \frac{1}{3} \cdot \left[\frac{x^3}{3}\right]_0^{\infty} = 1$$

$$\Rightarrow \frac{1}{3}\left[\frac{\infty^3}{3} - \frac{0^3}{3}\right] = \frac{1}{3} \cdot \frac{\infty^3}{3} = \frac{\infty^3}{9} = \infty \neq 1$$

Since the function doesnot satisfy $\int_{-\infty}^{\infty} f(x) \cdot d(x) = 1$, it doesnot serve PDF

Mean $= \mu = \sum x \cdot p(x)$

Find Mean for the Discrete Random Variable X

$$p(x) = \begin{cases} \frac{1}{6}; & x \in \{1,2,3,4,5,6\} \\ 0, & \text{otherwise} \end{cases}$$

Solution: $\sum_{x=1}^{6} x \cdot p(x)$

$= 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right)$

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

$= \frac{1}{6}(21) = 3.5$

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

Find Mean for the Continuous Random variable X

$$f(x) = \begin{cases} \frac{1}{20}x^2, & 2 \le x \le 4 \\ 0, & \text{otherwise.} \end{cases}$$

Solution:

$E(X) = \mu = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx$

$= \int_{2}^{4} x \cdot \left(\frac{1}{20}x^2\right) \cdot dx$

$= \frac{1}{20}\int_{2}^{4} x^3 \cdot dx = \frac{1}{20}\left[\frac{x^4}{4}\right]_2^4$

$\Rightarrow \frac{1}{20} \cdot \left[\frac{4^4}{4} - \frac{2^4}{4}\right] = \frac{1}{20}\left[\frac{240}{4}\right] = \frac{240}{80}$

$= 3$ //

Variance $= E(X^2) - (E(X))^2$

$E(X^2) = 1 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6}$

$= \frac{1}{6}[1+4+9+16+25+36]$

$= \frac{1}{6}[91] = 15.167$

$E(X) = 3.5 \Rightarrow (E(X))^2 = (3.5)^2 = 12.25$

So, the variance is

$\Rightarrow E(X^2) - (E(X))^2$

$\Rightarrow 15.167 - 12.25$

$\Rightarrow 2.917$

Variance $= E(X^2) - (E(X))^2$

$E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x) \cdot dx$

$= \int_{2}^{4} x^2 \cdot \left(\frac{1}{20}x^2\right) \cdot dx$

$= \frac{1}{20}\int_{2}^{4} x^4 \cdot dx$

$= \frac{1}{20}\left[\frac{x^5}{5}\right]_2^4$

$= \frac{1}{20}\left[\frac{4^5}{5} - \frac{2^5}{5}\right] = \frac{1}{20}\left[\frac{992}{5}\right]$

$\Rightarrow \frac{992}{100} = 9.92$

$E(X) = 3 \Rightarrow (E(X))^2 = 3^2 = 9$

So, the variance is

$E(X^2) - (E(X))^2$

$\Rightarrow 9.92 - 9$

problem : A Random Variable X has PDF as

$$f(x) = \begin{cases} kx^2, & -3 < x < 3 \\ 0, & \text{otherwise} \end{cases}$$

(i) find k, (ii) find $P(X \le 2)$

solution : one of the property of PDF is

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

so

$$\int_{-3}^{3} k \cdot x^2 \cdot dx = 1$$

$$\Rightarrow k \int_{-3}^{3} x^2 \cdot dx = 1$$

$$\Rightarrow k \cdot \left[\frac{x^3}{3} \right]_{-3}^{3} = 1$$

$$\Rightarrow k \cdot \left[\frac{3^3}{3} - \frac{(-3)^3}{3} \right] = 1$$

$$\Rightarrow k \cdot \left[\frac{27}{3} + \frac{27}{3} \right] = 1 \Rightarrow k \left[\frac{54}{3} \right] = 1$$

$$\Rightarrow k \cdot 18 = 1 \Rightarrow \boxed{k = 1/18}$$

(ii) find $P(X \le 2)$

$$= \int_{-3}^{2} f(x) \cdot dx = \int_{-3}^{2} kx^2 \cdot dx = k \int_{-3}^{2} x^2 \cdot dx$$

$$= \frac{1}{18} \int_{-3}^{2} x^2 \cdot dx = \frac{1}{18} \left[\frac{x^3}{3} \right]_{-3}^{2}$$

$$= \frac{1}{18} \left[\frac{2^3}{3} - \frac{(-3)^3}{3} \right]$$

$$= \frac{1}{18} \left[\frac{8}{3} + \frac{27}{3} \right]$$

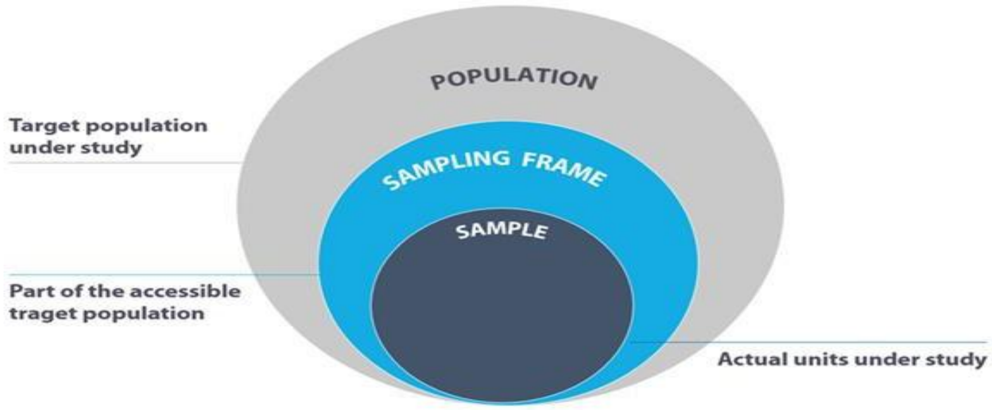
$$= \frac{1}{18} \left[\frac{35}{3} \right] = \frac{35}{54} = \boxed{0.64}$$

$$\boxed{P(X \le 2) = 0.64}$$

SAMPLING:

1.Population

- In statistics, a **population** refers to the entire group that you want to draw conclusions about.
- For instance, if you're conducting a survey about favorite ice cream flavors among all adults in a city, the **population** would be all adults in that city.

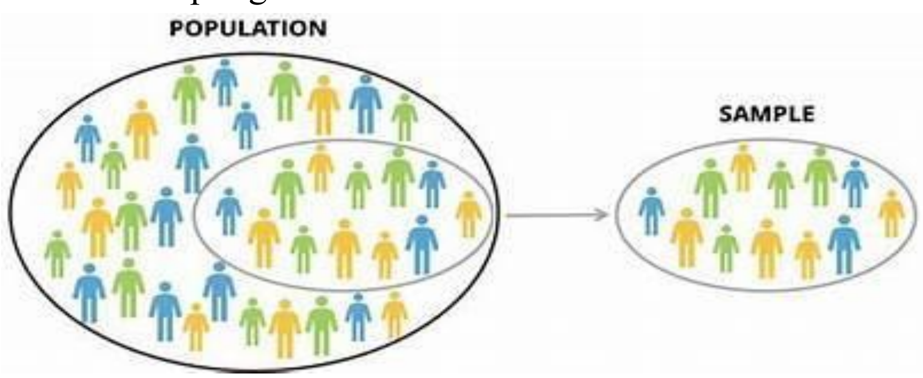


2. Sampling Frame

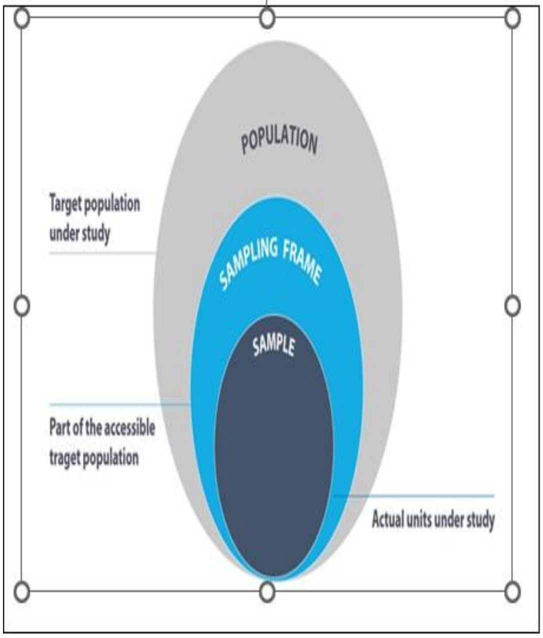
- The **sampling frame** is a list or an operational definition of the target population from which a sample is drawn.
- It is essentially a subset of the population from which you will actually select your sample.
- In the ice cream example, the sampling frame might be a list of all adult residents in the city obtained from a census or voter registration database.

3. Sample

- A **sample** is a subset of the population that is selected for study.
- Ideally, a sample should be chosen in such a way that it is **representative of the population**, meaning that the characteristics of the sample closely match those of the population.
- In the ice cream survey, the sample might consist of a randomly selected group of **500 adults** from the sampling frame.



Suppose you're a researcher interested in studying the **average income of households in a particular city.**



Population

Population :people or items with characteristics one wishes to analyse
Eg : all households in that city. So, if there are 100,000 households in the city, the population would consist of all 100.000 households.

Sampling Frame

Sampling Frame a list or an operational definition of the target population from which a sample is drawn.
Eg : city's housing registry, which lists all residential addresses. This registry serves as your sampling frame.

Sample

Sample :people or items with characteristics one wishes to analyse
Eg : randomly select 500 addresses from the housing registry. These 500 households would constitute your sample.

Sampling Design Process

Define Population

↓

Determine Sampling Frame

↓

Determine Sampling Procedure

Probability Sampling
Type of Procedure
Simple Random Sampling
Stratified Sampling
Cluster Sampling

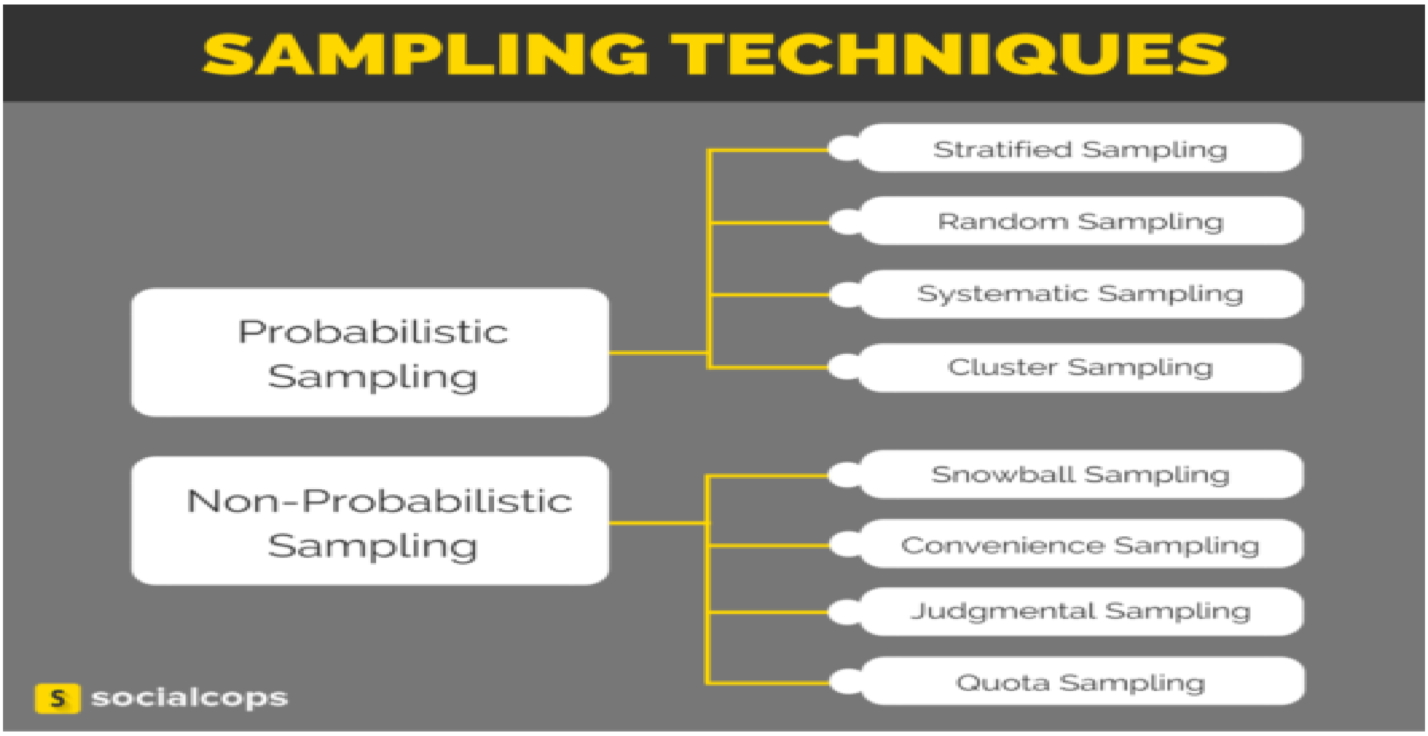
Non-Probability Sampling
Type of Procedure
Convenience
Judgmental
Quota

↓

Determine Appropriate Sample Size

↓

Execute Sampling Design



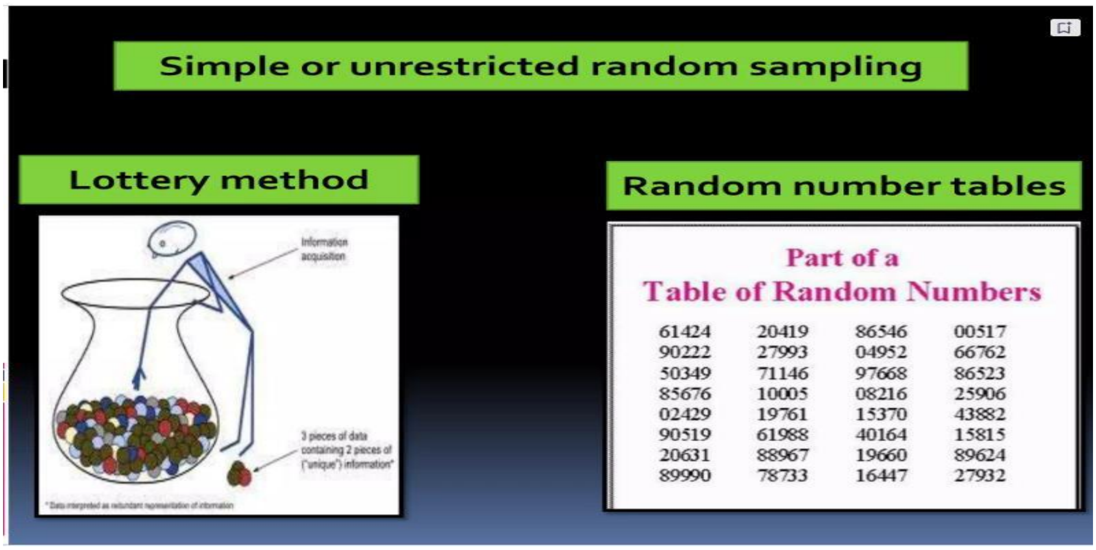
PROBABILISTIC SAMPLING TECHNIQUES

Simple Random Sampling

- In a **simple random sample**, every member of the population has an **equal chance of being selected**.
- It is the most basic and unbiased sampling method.

Example: Simple Random Sampling

- You want to select a simple random sample of **100 employees** from a social media marketing company.
- Assign a number to every employee in the company database (e.g., 1 to 1000).
- Use a **random number generator** to select 100 numbers.
- The employees corresponding to those numbers form the sample.

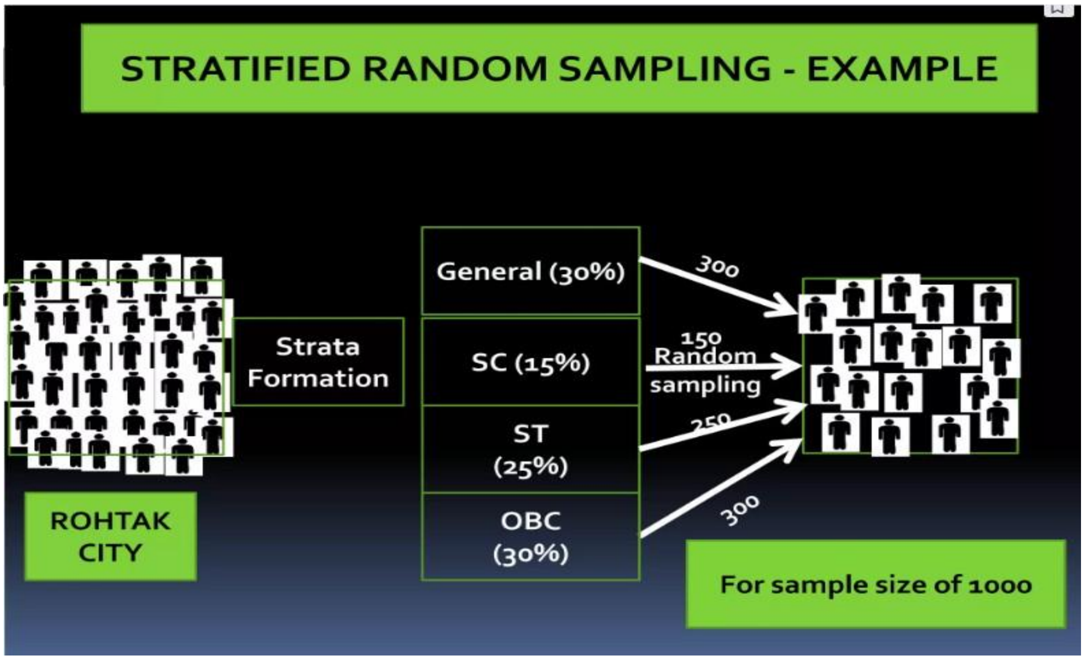


Stratified Sampling

- In **stratified sampling**, the population is divided into distinct subgroups called **strata** based on characteristics such as age, gender, income level, etc.
- Then, samples are randomly selected from each stratum **in proportion to their size** in the population.
- This method ensures **representation from all subgroups**.

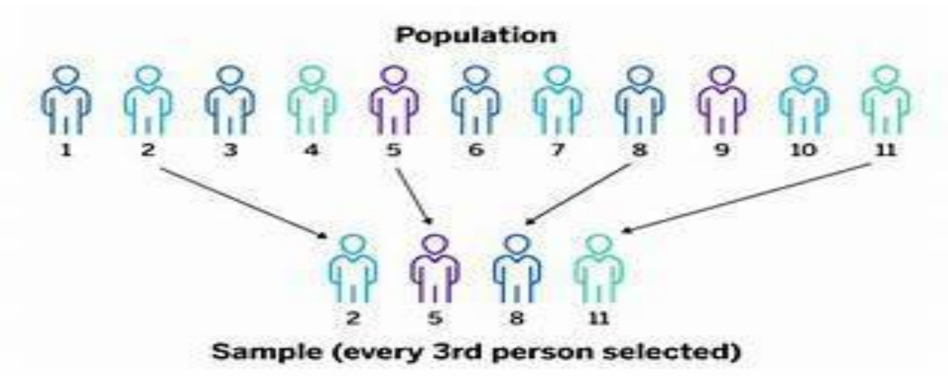
Example: Stratified Sampling

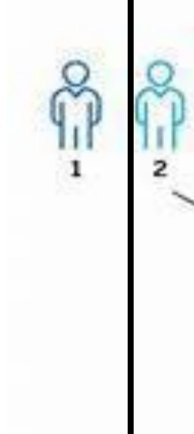
- A company has **800 female employees** and **200 male employees**.
- To maintain gender balance, the population is divided into two strata based on gender.
- Random sampling is applied to each group:
    - Select **80 females**
    - Select **20 males**
- This results in a **representative sample of 100 employees**.



Systematic Sampling:

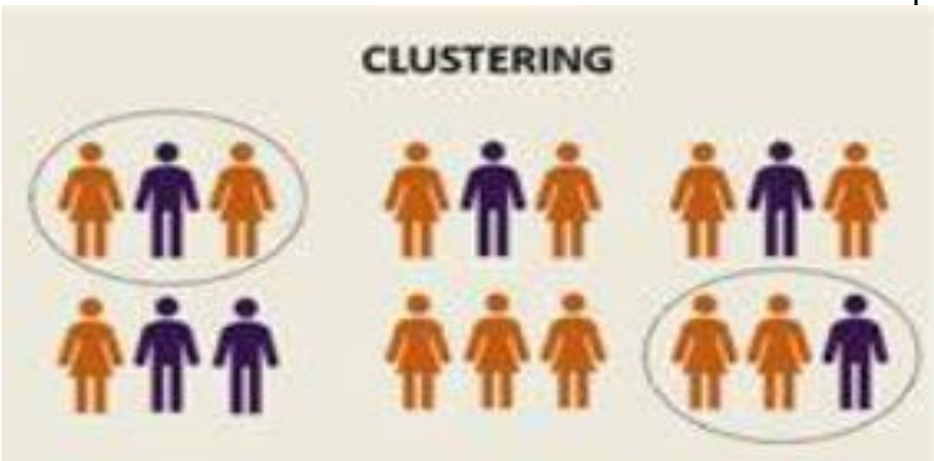
- Systematic sampling involves selecting every nth individual from the population after a random start.
- For example, if you have a population of 1000 and you want a sample of 100, you could select every 10th individual after choosing a random starting point between 1 and 10.

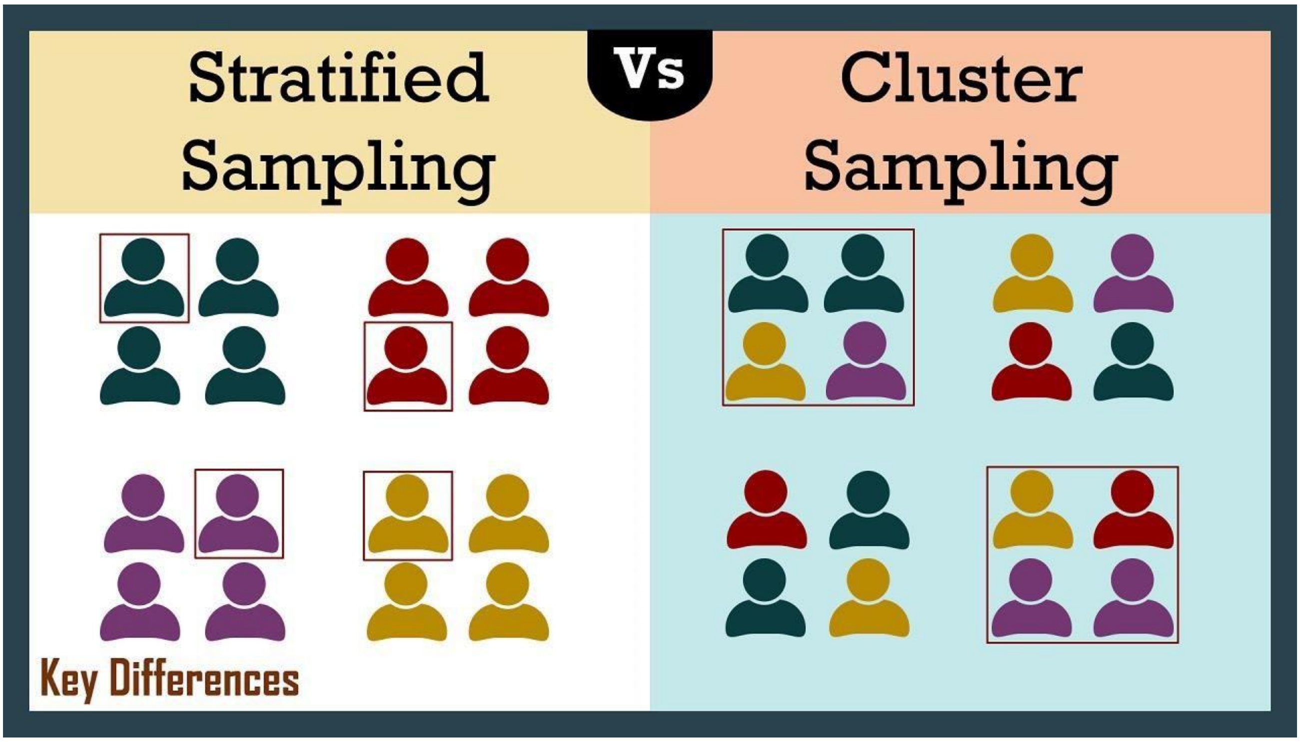




Cluster Sampling:

- Cluster sampling involves dividing the population into clusters or groups (e.g., neighborhoods, schools), and then randomly selecting clusters to include in the sample.
- All individuals within the selected clusters are then included in the sample





PROBABILISTIC SAMPLING TECHNIQUES:

1.Simple Random Sampling:

- In simple random sampling, every individual in the population has an equal chance of being selected, and each selection is independent of the others.
- This can be done by randomly selecting individuals from a sampling frame using methods like random number generators or random sampling software.

2. Systematic Sampling:

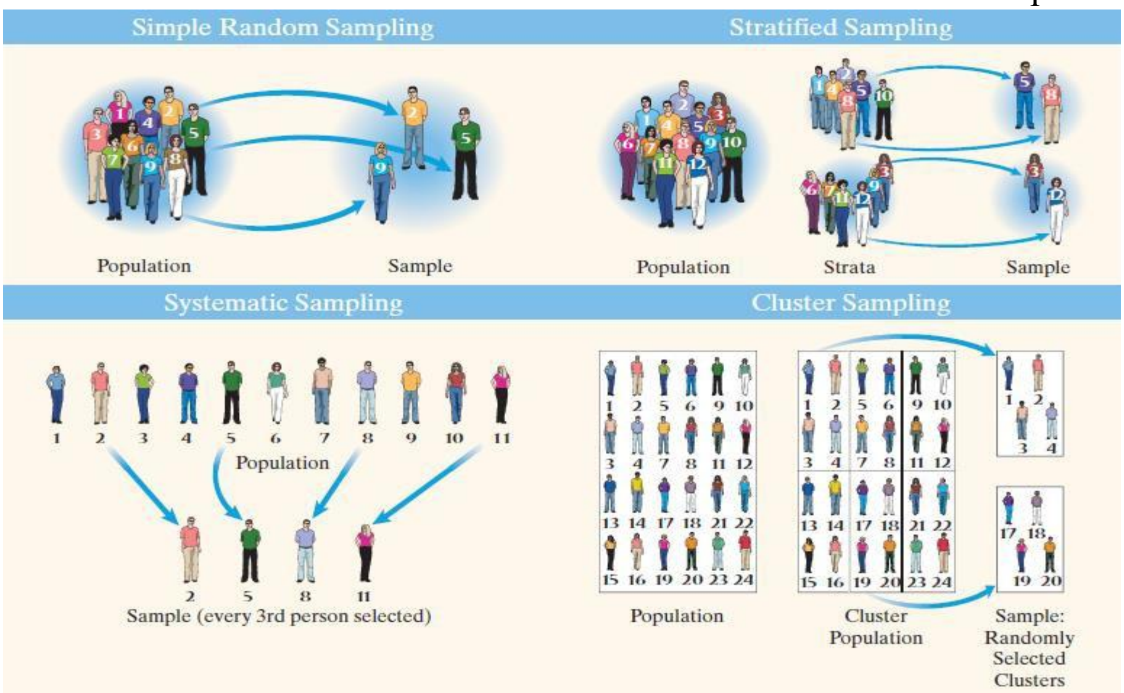
- Systematic sampling involves selecting every nth individual from the population after a random start.
- For example, if you have a population of 1000 and you want a sample of 100, you could select every 10th individual after choosing a random starting point between 1 and 10.

3. Stratified Sampling:

- In stratified sampling, the population is divided into distinct subgroups or strata based on certain characteristics (e.g., age, gender, income level), and then samples are randomly selected from each stratum in proportion to their size in the population.
- This ensures representation from all subgroups.

4. Cluster Sampling:

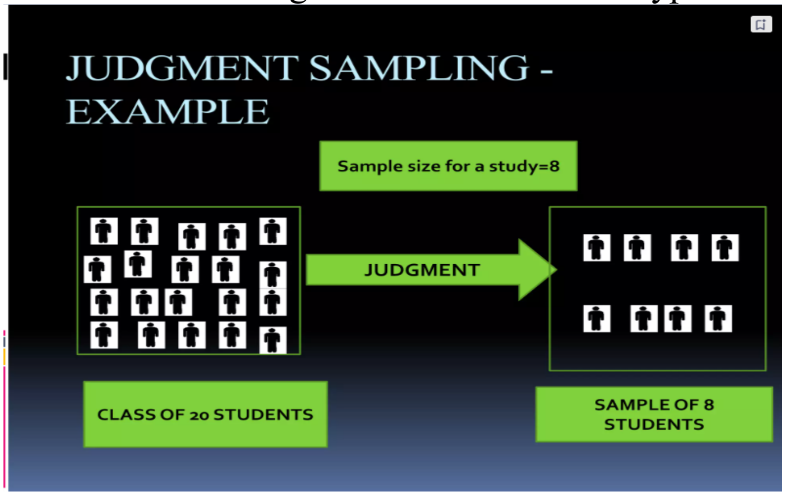
- Cluster sampling involves dividing the population into clusters or groups (e.g., neighborhoods, schools), and then randomly selecting clusters to include in the sample.
- All individuals within the selected clusters are then included in the sample.



NON PROBABILISTIC SAMPLING TECHNIQUES:

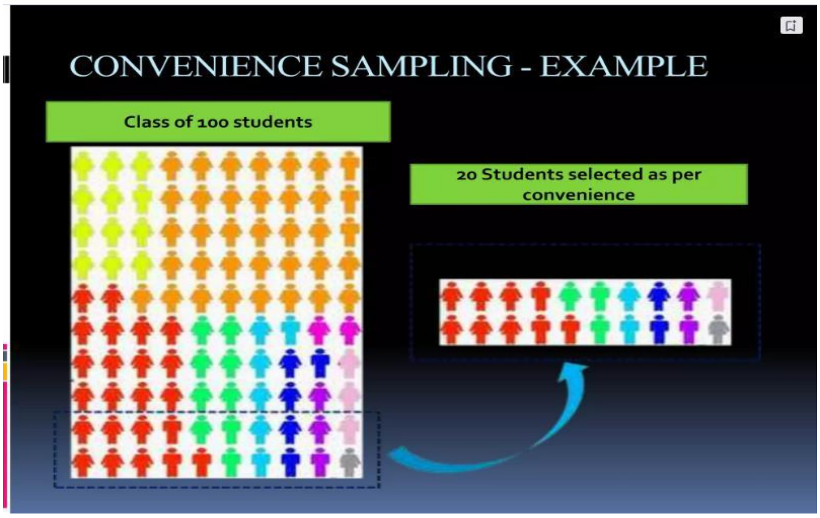
1. Judgement Sampling

- Depends exclusively on the Judgement of Investigator .
- Sample selected which investigator thinks to be most typical of the Universe.



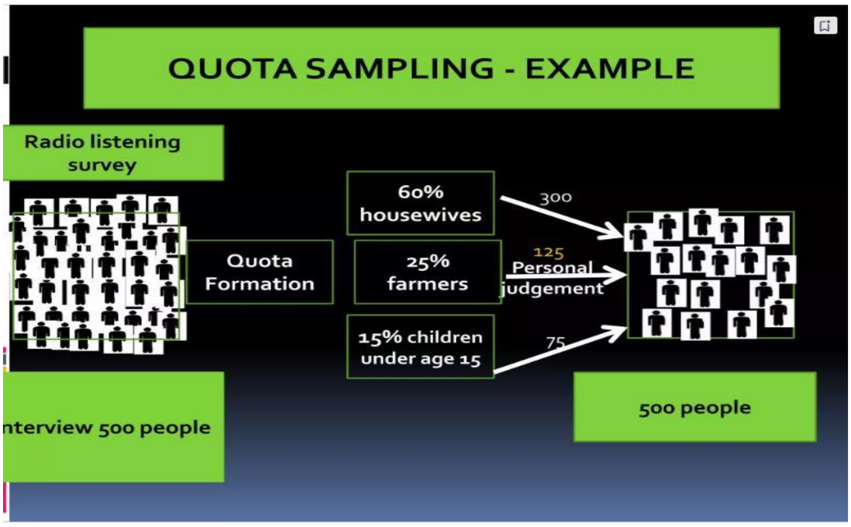
2. Convenience Sampling

- selecting individuals who are conveniently available or easily accessible to the researcher.
- This method is quick and inexpensive but may introduce bias as it doesn't ensure that the sample is representative of the population



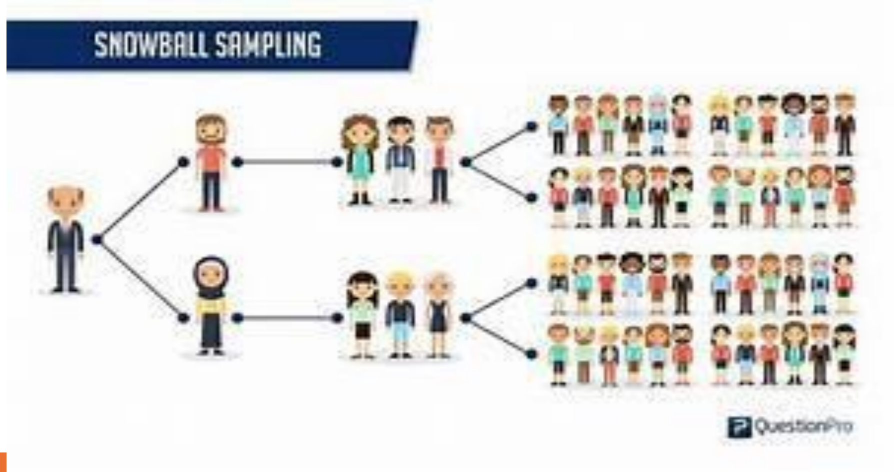
3. Quota Sampling

- This involves dividing the population into subgroups or strata based on certain characteristics (e.g., age, gender, socioeconomic status) and then setting quotas for each subgroup.
- Researchers then purposively sample individuals to meet these quotas



4. Snow Ball Sampling

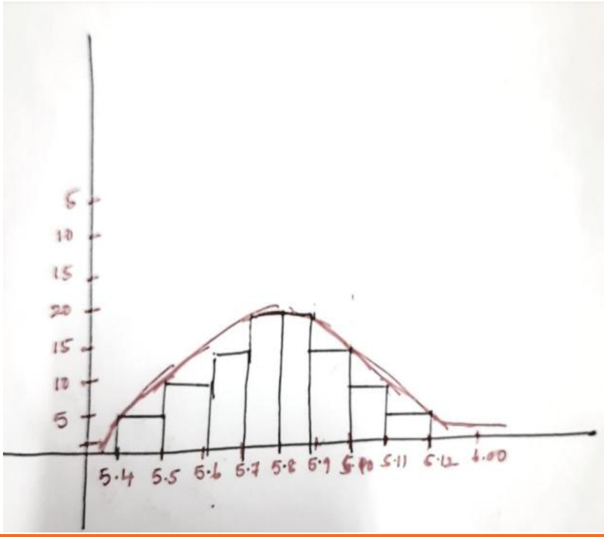
- This technique is used when members of the population are difficult to identify or locate.
- The researcher starts with a small number of initial participants who are known to belong to the population of interest and then asks them to refer other potential participants.
- This process continues iteratively, with new participants referring additional participants. Snowball sampling can be useful for studying hard-to-reach populations but may result in a biased sample if the initial participants are not representative of the population.



NORMAL DISTRIBUTION:

- In a normal distribution, data is **symmetrically** distributed with no skew.
- When plotted on a graph, the data follows a **bell shape**, with most values clustering around a central region
- Normal distributions are also called **Gaussian distributions** or bell curves because of their shape.

No.of.students	Heights
5	5.4-5.5
10	5.5-5.6
15	5.6-5.7
20	5.7-5.8
20	5.8-5.9
15	5.9-5.10
10	5.10-5.11
5	5.11-5.12



PROPERTIES OF NORMAL DISTRIBUTION

Normal distributions have key characteristics that are easy to spot in graphs:

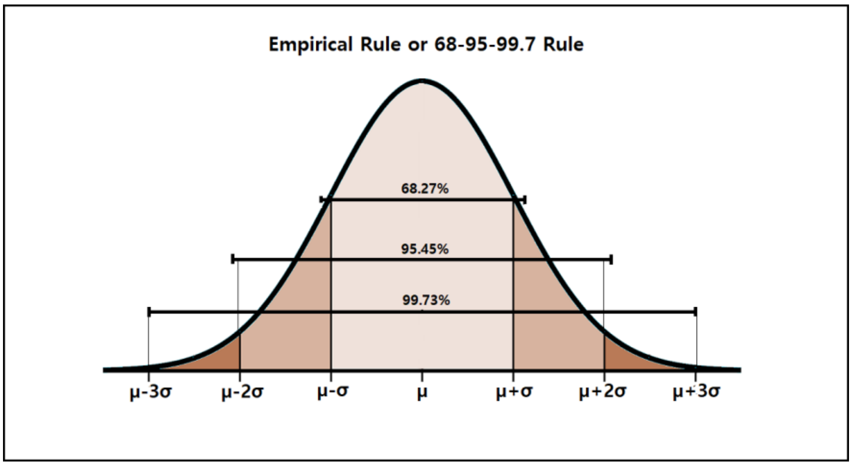
- The **mean, median** and **mode** are exactly the same.
- The **distribution is symmetric about the mean**—half the values fall below the mean and half above the mean.

The distribution can be described by two values: the **mean and the standard deviation**

EMPIRICAL RULE OF NORMAL DISTRIBUTION

Empirical rule
The empirical rule, or the 68-95-99.7 rule, tells you where most of your values lie in a normal distribution:

- Around 68% of values are within 1 standard deviation from the mean.
- Around 95% of values are within 2 standard deviations from the mean.
- Around 99.7% of values are within 3 standard deviations from the mean.



PROPERTIES OF NORMAL DISTRIBUTION

Problem 1: X is a normally distributed variable with mean $\mu = 30$ and standard deviation $\sigma = 4$. Find the probabilities

a) $P(X < 40)$

b) $P(X > 21)$

c) $P(30 < X < 35)$

Formula : $Z = (X- \mu)/ \sigma$

a) For x = 40, the z-value $z = (40 - 30) / 4 = 2.5$

Hence $P(x < 40) = P(z < 2.5)$ = [area to the left of 2.5] = 0.9938

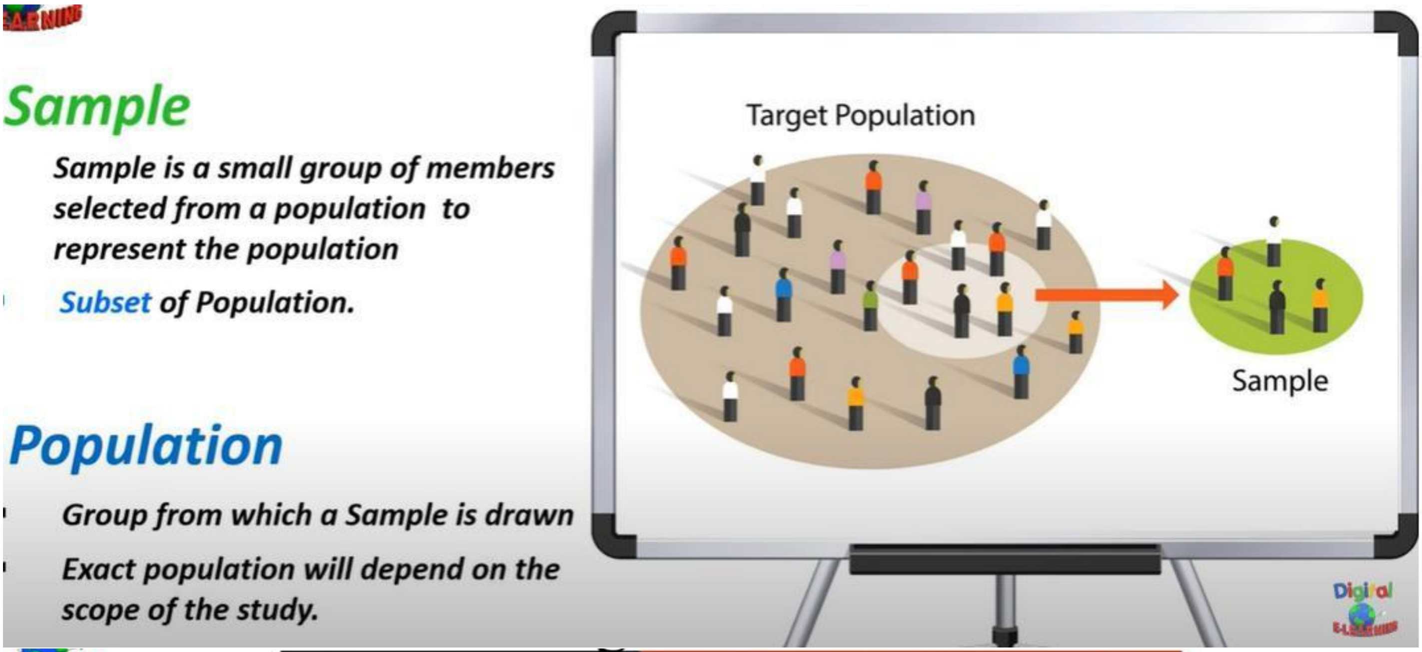
b) For x = 21, $z = (21 - 30) / 4 = -2.25$

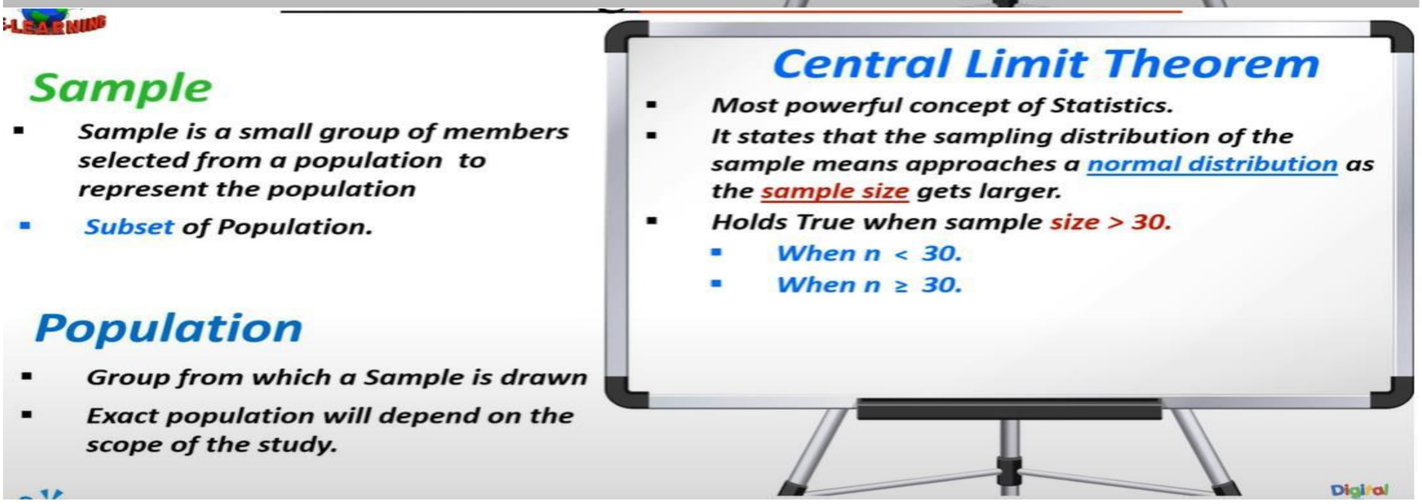
Hence $P(x > 21) = P(z > -2.25)$ = [total area] - [area to the left of -2.25] = 1 - 0.0122 = 0.9878.

c) For x = 30 , $z = (30 - 30) / 4 = 0$ and for x = 35, $z = (35 - 30) / 4 = 1.25$

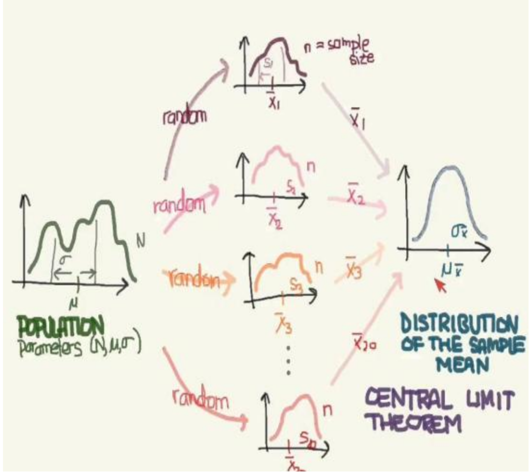
Hence $P(30 < x < 35) = P(0 < z < 1.25)$ = [area to the left of z = 1.25] - [area to the left of 0]== 0.8944 - 0.5 = 0.3944

CENTRAL LIMIT THEOREM





- The Central Limit Theorem is a powerful tool that enables researchers to draw conclusions about populations based on samples.
- The Central Limit Theorem (CLT) is a fundamental concept in statistics that states that the sampling distribution of the sample mean approaches a normal distribution as the sample size increases, regardless of the shape of the population distribution.
- This theorem is crucial because it allows us to make inferences about population parameters using sample statistics.



Question : There are **250 cats** at show, where average weight of **12 Kg**, with standard deviation of **8 Kg**. If we choose **4** samples, then what is the probability they have an average weight of greater than **10 Kg** and less than **25 Kg**?

$$Z = \frac{X-\mu}{\sigma_x} \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Sol : Mean of Population $\mu = 12$ Kg $\quad \sigma = 8$ Kg $\quad n = 4 \quad \sigma_x = 8/\sqrt{4}$

$$Z_{25} = \frac{25-12}{8/\sqrt{4}} \qquad Z_{25} = \frac{13}{4} \qquad Z_{25} = 3.25$$

Z value $3.25 = 0.994$

Z value $3.25 = 0.994 - 0.5$

Z value $3.25 = 0.494$

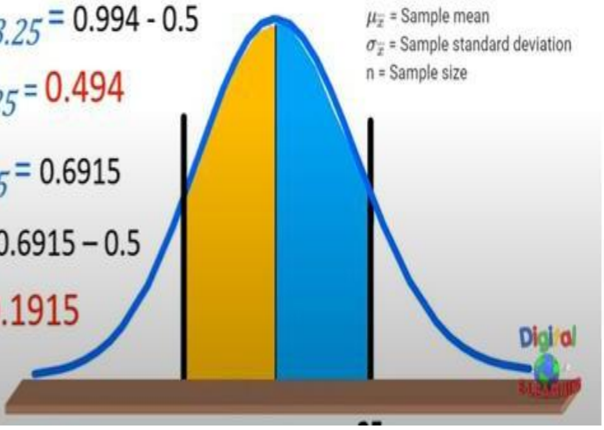
Where,
μ = Population mean
σ = Population standard deviation
$\mu_{\bar{x}}$ = Sample mean
$\sigma_{\bar{x}}$ = Sample standard deviation
n = Sample size

$$Z_{10} = \frac{10-12}{8/\sqrt{4}} \qquad Z_{10} = \frac{-2}{4} \qquad Z_{10} = -0.5$$

Z value $-0.5 = 0.6915$

Z value $-0.5 = 0.6915 - 0.5$

Z value $-0.5 = 0.1915$



Adding = 0.494 +0.1915 = .6855

Let $\bar{X}$ be the mean of a random sample of size 60 drawn from a population with mean 120 and standard deviation 45.

a. Find the probability that $\bar{X}$ assumes a value between 112 and 125.

Solve for Sample Standard Deviation

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \qquad \sigma = 45 \qquad n = 60$$

$$\sigma_{\bar{X}} = \frac{45}{\sqrt{60}}$$

$$\sigma_{\bar{X}} = 5.8095$$

Solve for $P(112 < X < 125)$

Z-Score Formula: $z = \dfrac{x - \mu_X}{\sigma_X}$

@ $x = 112$

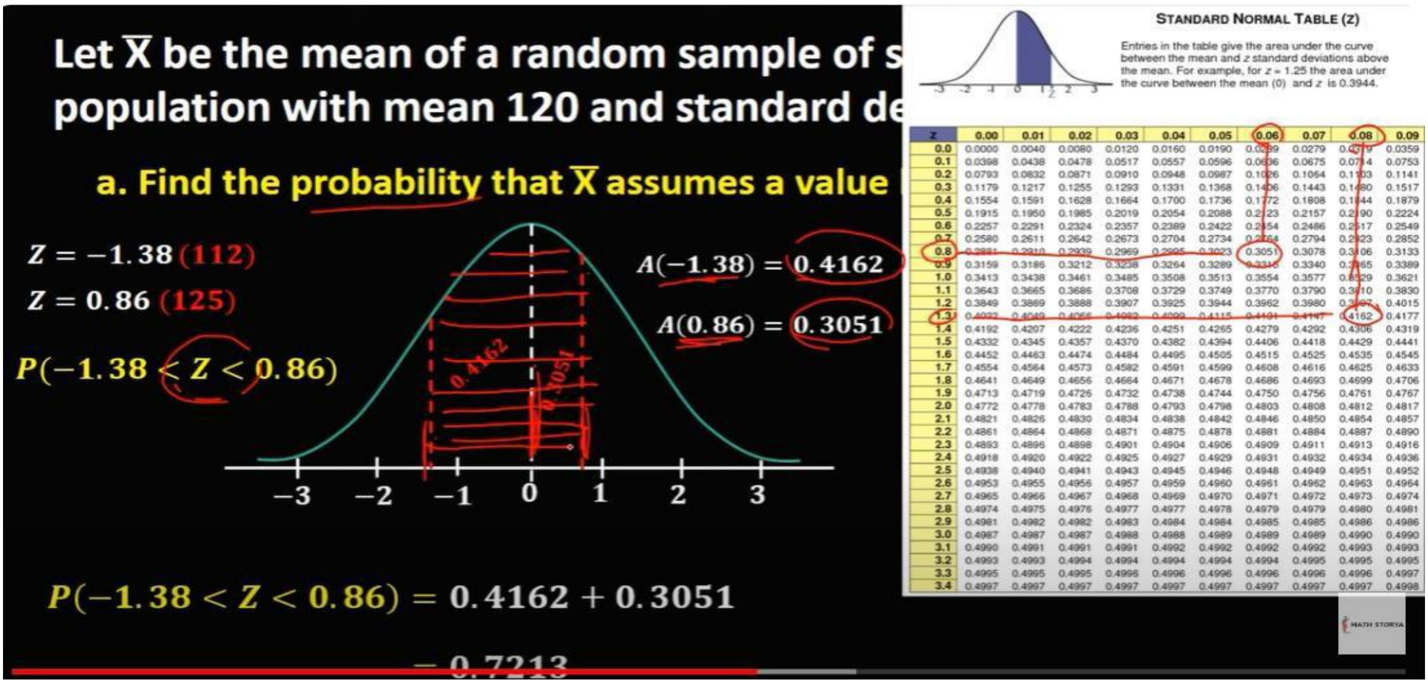
$$z = \frac{112 - 120}{5.8095}$$

$$z = -1.38$$

@ $x = 125$

$$z = \frac{125 - 120}{5.8095}$$

$$z = 0.86$$



Covariance

- **Covariance** is a measure of the relationship between two random variables and to what extent, they change together.

- Or in other words, it defines the changes between the two variables, such that change in one variable is equal to change in another variable.

- Mathematically, the covariance between two random variables X and Y, denoted as cov(X,Y).

- **Types of Covariance**

Covariance can have both positive and negative values. Based on this, it has two types:

1. Positive Covariance
2. Negative Covariance

Positive Covariance

If cov(X,Y)>0, it indicates a positive relationship, meaning that as the value of one variable increases, the other variable tends to increase as well.

Negative Covariance

If cov(X,Y)<0, it indicates a negative relationship, meaning that as the value of one variable increases, the other variable tends to decrease.

$$Cov\ (X,Y) = \frac{\sum (X_i - \bar{X})\ (Y_i - \bar{Y})}{N - 1}$$

Where:

- X_i and Y_i represent the observed values of X and Y.
- $\bar{X}$ and $\bar{Y}$ denote their respective means.
- N is the number of observations.

Suppose we want to evaluate the relationship between the number of hours studied (X) and the test scores (Y) obtained by a group of five students. The data are below.

Hours (X)	Score (Y)
3	70
5	80
2	60
7	90
4	75

	Hours (X)	Score (Y)	$(X_i - \bar{X})$ 1	$(Y_i - \bar{Y})$	Product 2
	3	70	-1.2	-5	6
	5	80	0.8	5	4
	2	60	-2.2	-15	33
	7	90	2.8	15	42
	4	75	-0.2	0	0
Average	4.2	75		Total	85 3
				N-1	4
				Covariance	21.25 4

The positive covariance (21.25) suggests a positive association between the number of hours studied and exam scores. This result implies that as the number of study hours increases, the scores tend to increase.

Dereviatives

Integration Rules:

1. $\int f(x)dx = F(x) + C \Leftrightarrow F'(x) = f(x)$
2. $\int a\, f(x)dx = a \int f(x)dx$
3. $\int - f(x)dx = - \int f(x)dx$
4. $\int [f(x) \pm g(x)]\, dx = \int f(x)dx \pm \int g(x)dx$

Differentiation Formulas:

1. $\frac{d}{dx}(x) = 1$
2. $\frac{d}{dx}(ax) = a$
3. $\frac{d}{dx}(x^n) = nx^{n-1}$
4. $\frac{d}{dx}(\cos x) = -\sin x$
5. $\frac{d}{dx}(\sin x) = \cos x$
6. $\frac{d}{dx}(\tan x) = \sec^2 x$
7. $\frac{d}{dx}(\cot x) = -\csc^2 x$
8. $\frac{d}{dx}(\sec x) = \sec x \tan x$
9. $\frac{d}{dx}(\csc x) = -\csc x(\cot x)$
10. $\frac{d}{dx}(\ln x) = \frac{1}{x}$
11. $\frac{d}{dx}(e^x) = e^x$
12. $\frac{d}{dx}(a^x) = (\ln a)a^x$
13. $\frac{d}{dx}(\sin^{-1} x) = \frac{1}{\sqrt{1-x^2}}$
14. $\frac{d}{dx}(\tan^{-1} x) = \frac{1}{1+x^2}$
15. $\frac{d}{dx}(\sec^{-1} x) = \frac{1}{|x|\sqrt{x^2-1}}$

Integration Formulas:

1. $\int 1\, dx = x + C$
2. $\int a\, dx = ax + C$
3. $\int x^n dx = \frac{x^{n+1}}{n+1} + C, \ n \neq -1$
4. $\int \sin x\, dx = -\cos x + C$
5. $\int \cos x\, dx = \sin x + C$
6. $\int \sec^2 x\, dx = \tan x + C$
7. $\int \csc^2 x\, dx = -\cot x + C$
8. $\int \sec x(\tan x)\, dx = \sec x + C$
9. $\int \csc x(\cot x)\, dx = -\csc x + C$
10. $\int \frac{1}{x} dx = \ln|x| + C$
11. $\int e^x dx = e^x + C$
12. $\int a^x dx = \frac{a^x}{\ln a} + C \quad a > 0,\ a \neq 1$
13. $\int \frac{1}{\sqrt{1-x^2}} dx = \sin^{-1} x + C$
14. $\int \frac{1}{1+x^2} dx = \tan^{-1} x + C$
15. $\int \frac{1}{|x|\sqrt{x^2-1}} dx = \sec^{-1} x + C$
16. $\int \tan x\, dx = \ln|\sec x| + C$ or $-\ln|\cos x| + C$
17. $\int \cot x\, dx = \ln|\sin x| + C$ or $-\ln|\csc x| + C$
18. $\int \sec x\, dx = \ln|\sec x + \tan x| + C$
19. $\int \csc x\, dx = \ln|\csc x - \cot x| + C$
20. $\int \ln x\, dx = x \ln|x| - x + C$
21. $\int \frac{1}{\sqrt{a^2-x^2}} dx = \sin^{-1}\left(\frac{x}{a}\right) + C$
22. $\int \frac{1}{a^2+x^2} dx = \frac{1}{a}\tan^{-1}\left(\frac{x}{a}\right) + C$
23. $\int \frac{1}{x\sqrt{x^2-a^2}} dx = \frac{1}{a}\sec^{-1}\left|\frac{x}{a}\right| + C$ or $\frac{1}{a}\cos^{-1}\left|\frac{a}{x}\right| + C$
24. $\int \sin^2 x\, dx = \frac{x}{2} - \frac{\sin(2x)}{4} + C$. Note: $\sin^2 x = \frac{1-\cos 2x}{2}$

Unit Highlights

- Linear algebra provides the language of **vectors, matrices, eigenvalues, and SVD**.
- Probability and statistics help model **randomness, data variation, and uncertainty**.
- Common distributions include **binomial, Poisson, and normal**.
- The **central limit theorem** explains why sample means often behave approximately normally.
- Calculus provides **derivatives and partial derivatives** for optimization.
- These mathematical ideas are fundamental to machine learning and A

TEXT BOOKS:

1. Amit Kumar Das, Saptarsi Goswami, Pabitra Mitra, Amlan Chakrabarti, “Deep Learning”, Pearson Paperback, First Edition,2021.

REFERENCE BOOKS:

1. Josh Patterson and Adam Gibson, “Deep learning: Apractitioner's approach ”,O' Reilly Media, First Edition,2017.
2. Nikhil Buduma, “Fundamentals of Deep Learning, Designing next generation machine intelligence algorithms”,O’Reilly,ShroffPublishers,2019.

Question Bank:

UNIT I - MATHEMATICAL TOOLS FOR DEEP LEARNING		
PART –A		
1.	Define Matrix and Vector.	L1
2.	What is the transpose of a matrix?	L2
3.	Define Tensor.	L1
4.	What is the rank of a matrix?	L2
5.	Define determinant of a matrix.	L1
6.	What is trace of a matrix?	L1
7.	Define Eigen values and Eigen vectors.	L2
8.	What is Singular Value Decomposition (SVD)?	L1
9.	Define probability.	L2
10.	What is a random variable?	L1
11.	State Binomial distribution formula.	L1
12.	Define Poisson distribution.	L2
13.	What is Normal distribution?	L1
14.	State Central Limit Theorem.	L2
15.	Define derivative.	L2
16.	State product rule of derivatives.	L2
17.	What is partial derivative?	L1
PART –A		
1.	Explain Eigen values and Eigen vectors with example and discuss their applications in Deep Learning.	L3
2.	Explain Singular Value Decomposition (SVD) with mathematical formulation and applications in dimensionality reduction.	L4
3.	Solve a system of linear equations using matrix method and explain the significance of rank.	L3
4.	Explain determinant and its properties. Discuss its role in matrix invertibility.	L4
5.	Describe Normal distribution and explain the Central Limit Theorem with suitable example.	L3
6.	Derive and explain Binomial and Poisson distributions with properties and examples.	L4
7.	Explain probability and random variables with real-time applications in Machine Learning.	L3
8.	Discuss matrix norms and tensors and explain their importance in Deep Learning models.	L4
9.	Explain derivatives and rules of differentiation with examples related to optimization.	L5
10.	Explain partial derivatives and their role in gradient-based learning algorithms.	L5