DIRECTORATE OF DISTANCE EDUCATION

UNIVERSITY OF JAMMU

JAMMU

SELF-LEARNING MATERIAL
B.A. SEMESTER III

SUBJECT : MATHEMATICS **UNIT : I - V**

COURSE : UGM 0301 **LESSON NO. : 1 - 18**

DR. HINA S. ABROL

COURSE CO-ORDINATOR

http:/www.distanceeducationju.in

Printed and Published on behalf of the Directorate of Distance Education, University of Jammu, Jammu by the Director, DDE, University of Jammu, Jammu.

STATISTICS, NUMERICAL, ANALYSIS AND MATRICES

Course Contributors

- **Dr. Sunil Gupta**
- **Dr. Kamaljeet Kour**

Proof Reading

- **Dr. Narinder Sharma**

Printed at : AJANTA ART PRESS/ 20 / 50 Nos.

Semester III

STATISTICS, NUMERICAL, ANALYSIS AND MATRICES

Credit : 04 Semester Examination : 80 marks
C. No. : UGM 0301 Sessional Assessment : 20 marks

STATISTICS

UNIT 1

Probability Distribution : Defination of random variable, distribution of random variable. Distribution of cumulative distribution function (d.f.) and its properties. Two types of distribution (i) disdiscrete (ii) continuous

The idea of probability function (p.f.) and probability distribution function (p.d.f.). Construction of probability distribution. Bernoulli s distribution Conditions to be satisfied by a certain function in order that if may be p.f. of a p.d.f. of a certain distribution

Expectations and Moments : Expectation of a random variable and its interpretation. Expectation of $(X-a)^k$ and $(x-E(X))^k$ (for k= 1,2,3) whwre a is any arbitrary point and vice-versa. Moment generating function (m.g.f.) and its properties. Expectation of the sum and product of points whwn n dice are thrown. Expectation of the sum of n Bernoulli variates each with probability od sucess being. Exercises based on these concepts.

UNIT 2

Binomial Distribution : Definition, mean, variance and moment generating function, mode of the binomial distribution measure of skewness and Kurtosis. IF X is b(n.p). find the distribution of $Z = \frac{X - np}{\sqrt{npq}}$

Poisson Distribution : Defination mean variance and m.g.f. Poisson Distribution as a limiting case of Binomial Distribution. Measure of skewness and Kurtosis. If X is Poisson variate with parameter X then find the distribution of $Z = \frac{X - X}{\sqrt{X}}$

Simple problems based on these distributions.

UNIT 3

Normal Distributions : Defination, mean, variance and m.f.g. Properties of the normal curve. Normal Distribution as a limiting case of Binomial and poisson Distribtion. If X ~N $(\mu.\sigma^2)$, then $Z = \frac{X - \mu}{\sigma}$ ~ N (0.1).

Gamma Distribution : Defination, mean, variance and m.f.g.

Chi-Square Distribution : Defination, mean, variance and m.f.g.

Simple problems based on these distribution

UNIT 4

NUMERICALANALYSIS

Finite difference operator E and D. Interpolation, Inverse interpolation and linear interpolation. Newton's forward and backward difference interpolation formula. Lagrange's formula for interpolation and inverse interpolation. Simple problems based on these concepts.

UNIT 5

MATRICES

Matrices and their types : symmetric, skew-symmetric. Hermitian, skew Hermitian, orthogonal and unitary matrices with examples. Rank a matrix. Cayley-Hamilton Theorem eigenvalues and eigenvectors of matrices. Examples and exercise based on these concepts

Text Books :-

1. Mood, A.M. Graybill, F.A. and Boes D.C. Introduction to the Theory of Statistics, Mc Gray-Hill.

2. Hogg. P. G. Introduction to Mathematical statistics. John Wiley and Sons.

3. Kapur, J. N. and sazena, H. C. Mathematical Stastistics, S. Chand and Co.

4. Gupta. S. C. Fundamental of Stastistics. S. Chand and Co.

5. Bajaj, C.P. and Gupta, P.N. Elements of Stastistics. R.Chand and Co.

6. Sastry, S. S. Introductory Methods of Numerical Analysis, Prentice-Hall of India.

7. Ralston, A., A First Course in Numerical Analysis, Prentice Hall of India.

8. Saxena, H. C,. Calculus of Finite Difference. S. Chand and Co.

9. Sharma, H. S., Sharma, F.C., and Choudhary, S. S., A Text Book of Numerical Analysis, Rattan Prakasham Mandir, Agra.

10. Biswas S., A Text Book of Matrix Algebra, New Age International.

Note :-

1. Each lecture will be of one hour duration.
2. The question paper shall consist of ten question, two question from each unit. The candidate will be required to do five question, selecting exactly one question from each unit.

TABLE OF CONTENTS

TABLE OF CONTENTS

STATISTICS

PROBABILITY DISTRIBUTION

By: Dr. Sunil Gupta

1.1 INTRODUCTION : In this lesson the concept of probability, random variable, types of random variables such as discrete, continuous with suitable examples wherever required are given It is assumed that the concept of probability had already been studied by the students opting this course in class 12th. There the idea was introduced in its simplest form with the help of examples pertaining to different experiments e.g., the coin tossing experiment, a cardexperiment, one die experiment and a two dice experiment.

1.2 OBJECTIVES: Objective of studying this lesson is to give awareness among the students about the concept of Probability.This lesson will certainly helps students to understand the concepts of random variables which enables them to collect data from given data in scattered form.

1.3 PROBABILITY:

1.3.1.History The term probability- a measure of chance was defined in its classical form. Three importance consequences of the classical definition of probability were also mentioned there. In the present course our endeavor is to develop the 'axiomatic definition' of probability starting from 'classical definition'. To do this, we need the following: -

1.3.2.Experiment: The word 'experiment' is used in probability and statistics in a much broader sense than in everyday conversation. For example, tossing a coin or rolling a die is considered an experiment Many more serious experiments are associated with every medical and scientific research campaign, such as the search for a polio vaccine, the study of the cause and cure of cancer. The only difference between the above two types of experiments is that in the former case we can list all possible outcomes and it is expected that these outcomes would be 'equally likely' whereas in the latter case it is difficult to imagine all possible outcomes and it is not to be expected that these outcomes would be equally likely. We focus our attention on experiments that have only a finite number of possible outcomes or as many possible outcomes as there are positive integers.

1.3.3.Sample Space& Sample point. A sample space S of an experiment E is the set of allpossible outcomes. An element in a sample space is called a sample point.

Example: We give below the idea how to construct sample spaces for the following experiments:-

(i) A coin is tossed twice and the number of heads observed,

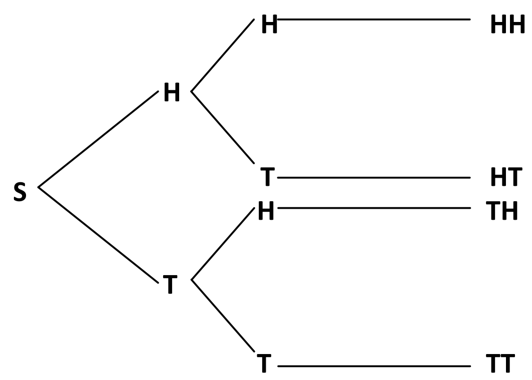
(ii) A family having three children,

(iii) A die is rolled and a note is made of the number shown on top.

(iv) From an urn containing only five red and three black balls, a ball is drawn.

(v) A coin is tossed until a head appears.

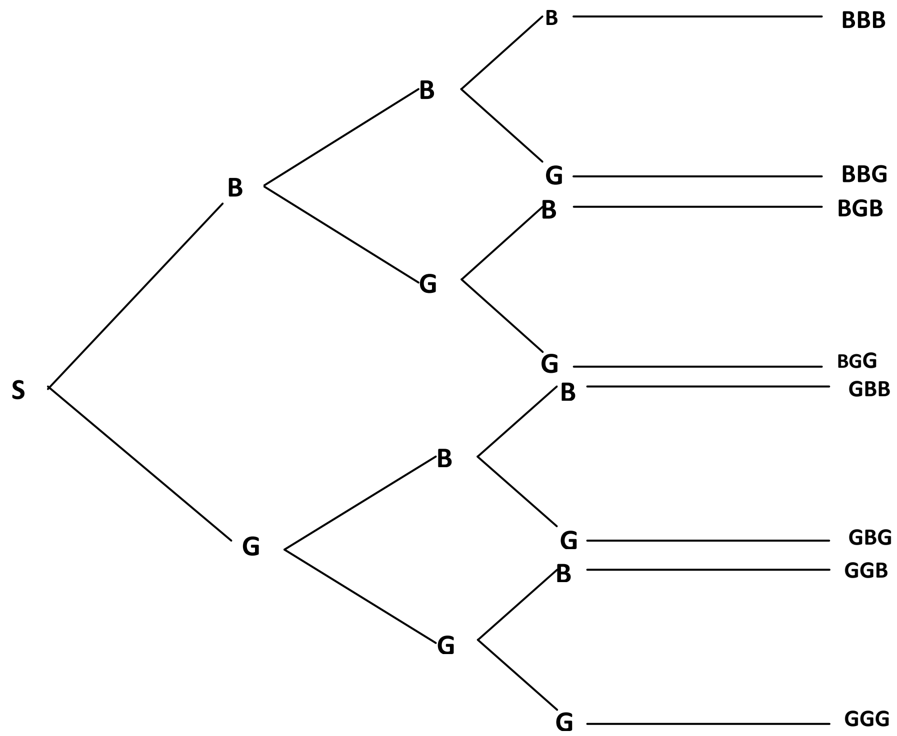
Solution (i): The sample space S for (i) is S= {HH,HT,T ,TT},



where HH means head on the first toss, head on the second toss and so on outcomes are -equally likely i.e. they occur with equal frequency.

Note: The students are required in check that if a coin is tossed n times, then it's possible -outcomes is 2^n. These outcomes are also equally likely.

ii) The sample space S for (ii) is



S= {BBB,BBG,BGB,BGG,GBB,GBG,GGB,GGG}

where BBB means that the first offspring is male, the second off spring is male and the third off spring is also male.

The sample space S for (ii) is

S = {x:x= 1, 2, 3, 4, 5, 6}

where 1,2, 3, 4, 5, 6, denote the number of dots shown on the top face of the die. The sample space S for (iv) is

$S = \{R_1, R_2, R_3, R_4, R_5, B_1, B_2, B_3\}$

The possible outcomes of the experiment (v) are H, TH, TTH, TTTH, TTTTH, TTTTTH,……………. Clearly, the number of possible outcomes is not finite We could have an indefinitely long sequences of tails before getting a head. 'However, we can make a list of enumerate the set of all possible outcomes. The sample space of this experiment is countablyinfinite. The countably infinite space has as many sample points as there are positive integers.

1.3.4.Event. An event is a subset of a sample space S of an experiment.

Since an event is a subset of the sample space S, this means that S itself is sure event and the empty set ϕ is the impossible event which can never occur.

1.4.PROBABILITY OF AN EVENT

The theory of probability has always been associated with games of chair, dice, cards etc. Directly or indirectly, probability plays an important role in science, business, and everyday life, where decisions and understanding involve uncertainty or risk. Thus, it is unfortunate that the term 'probability' itself, is difficult to define

1.4.1. Definition : If an experiment has n different equally likely outcomes and if exactly m of these outcomes are favorable to the event A, then the probability of event A, denoted by P(A), is

$$P(A) = \frac{\text{Number of favourable outcomes}}{\text{Number of possible outcomes}} = \frac{m}{n}$$

NOTE. Clearly, $\frac{m}{n}$ is a positive number not greater than unity so that $0 \leq \frac{m}{n} \leq 1$

1.4.2.Axioms of Probability

We shall now formulate the axioms of probability and, shall write the probability of event A as P (A), the probability of event B as P (B), and so on. We shall also continue to denote the set of all possible outcomes, namely the sample space, by the letter S. As we shall formulate them here, the axioms of probability apply only when the sample space S is discrete(finite or countably infinite); the modifications that are required for the continuous case will be discussed in later lessons.

Axiom 1: The probability of any event is a non-negative real number, that is, $P(A) > 0$, for any subset A of S.

Axiom 2: $P(S)=1$.

Axiom 3: If A_1 and A_2 are any two disjoint subsets of S (or A_1 and A_2 are any two mutually exclusive events), then

$P(A_1 U A_2) = P(A_1) + P(A_2)$.

1.5. RANDOM VARIABLE. A variable whose value is a number determined by theoutcome of an experiment is called a random variable.

Or

Consider an experiment for which the sample space is S. A real valued function that is defined on the space S is called a random variable. In other words, in a particular experiment,a random variable X would be some function that assigns a real number X(S) to each possible outcomes $s \in S$. For instance in above example

X(HHT)= 2, where X stands for the numberof heads when a coin is tossed thrice.

1.5.1. (Discrete and continuous)

Discrete Random Variable. A random variable which can assume only a discrete setof real values and for which the value which the variable depends on chance, is calleddiscrete random variable or a stochastic random variable.

Thus if we throw a perfect die, X, the number of points on the die is a discrete random variable since (i) X takes only a set of discrete values viz., 1, 2, 3, 4 ,5, 6 (ii) the value which it actually takes in an experiment depends on chance. In fact X can take values 1, 2, 3, 4, 5,6,each with probability 1/6.

Also note that 'twice the number of points on a die' is also a discrete random variable which takes the values 2, 4, 6, 8, 10, 12 with probability 1/6 in each case.

Other Examples of discrete random variable:

(i) If X denotes the sum of points in a throws of two perfect dice, then X is random variable. Here X takes the values 2, 3, 4, 5, 6, 7, 8, 9, 10, .11, 12 with probabilities 1/36, 2/36, 3/36, 4/36, 5/36, 6/36, 5/36, 4/36, 3/36, 2/36, 1/36.

(ii) Three coins are tossed. How many falls as "head"?

Solution. The answer is a number determined by the outcome of the experi ment. The number may be 0, 1, 2 or 3. Although we cannot predict the out come exactly, we can say what the possibilities and probabilities are. A sample space for the experiment is shown in the first column (Table 1). The second column shows the number of heads for same sample point, and the third column shows the probabilities of the sample points.

Table 1. Three coins

Sample points	Number of heads	probability
HHH	3	1/8
HHT	2	1/8
HTH	2	1/8
THH	2	1/8
THT	1	1/8
HTT	1	1/8
TTH	1	1/8
TTT	0	1/8

The information about the possible number of heads, and their probabilities, is collected in Table 2 The probability of getting exactly 2 heads is found by adding the probabilities of all two heads HHT, HTH, THH, (because the three events are mutually exclusive).

$$1x\frac{1}{8}\,3\times\frac{1}{8},3\times\frac{1}{8},1\times\frac{1}{8}$$

Table 2. Three coins. probability function for number of heads

X=x	0	1	2	3
P[X=x]	1/8	3/8	3/8	1/8

(iii) If we let the variable X to denote the number of heads, the Table 2 shows the possible values that X can have, and the probability of each value. This set of ordered pairs, each of form (number of heads, probability of that number) is the probability function of X. Since value of X is a number determined by the outcome of an experiment, X is called a random variable.

(iv) 1.5.2. REMARK.In this example, the values of f(x) are

(v) The coefficients 1, 3, 3, 1 are the binomial coefficients for

(vi) $\binom{3}{x}$, x=0,1,2,3.Consequently all values of f(x) are given by the following formula

(vii) f(x)= P(X=x) = $\binom{3}{x}\left(\frac{1}{2}\right)^3$, x=0,1,2,3

Example. : We select one of the integers 1 through 10 factors. What is the probability that it has exactly 2 divisors? More than 4?

(viii) Table 3. Number of divisors of integers 1 through 10

Integer	1	2	3	4	5	6	7	8	9	10
No of divisors	1	2	2	3	2	4	2	4	3	4

Solution. The first row of Table 3 is a sample space integer from I through 10 at random. Let X be the number of divisors of the selected integer. The second row shows the value of the random variable of each sample point. Each integer has probability 0.1 of being drawn because the expression 'at random' means that the integer 1 through 10 are equally likely.

(ix) Next, we combine case according to the number of divisions for each sample point, thus obtaining table 4.

(x) Table 4. Numbers of divisors and their probabilities (for integer selected at random from 1 through 10).

No of divisors, X = x	1	2	3	4
P [X=x]= f(x)	0.1	0.4	0.2	0.3

1. 6. EXAMINATION ORIENTED QUESTIONS/LESSON END EXERCISE

Q.1.: A sample space S serves as the universal set for all questions related to an experiment An event A with respect to a particular sample space S is simply a set of possible outcomes favorable to the event A. For example, in the experiment of tossing a coin twice, the set

$$S=\{HH, HT, TH, TT\}$$

is the sample space for the experiment.For each outcome of the experiment we can determine whether a given event does or does not occur. We may be interested in the event "both tosses show head" we find that this event occurs if the experiment results in an outcome corresponding to an element of the set

$$A=\{HH\}$$

We recognize A as a subset of the sample space S.

The following are some examples of events:

Q.2. In the experiment of tossing a coin twice, Write event of

a)"getting two heads"

b) "getting two tails.

Q.3. In the experiment of rolling a die, "an even number occurs" is an event "An odd number occurs" is another event. Write it in tabular form

1. Event "A or B" . if A and B are events (sets).$A\cup B$ is the event which occurs iff A or B or both occur.

2. Event "A and B" . if A and B are events (sets), $A\cap B$ is the event which occurs if A and B occur.

3. Event "not A", if A is an event (set), A ? is the event which occurs iff A does not occur.

Q.4.: A die is rolled. What is the probability that the even number occurs?

Solution: The sample space S for this experiment is the set

S = {1, 2, 3, 4, 5, 6}, of all possible outcomes. Let A be the event "even number occurs". Then. A = {2, 4, 6}. The number

of possible outcomes is 6 and the number of outcomes favorable to the event A is 3.

Therefore, $P(A)=\frac{3}{6}=\frac{1}{2}$.

Q.5.: A card is drawn from a well shuffled deck of 52 cards. What is the probability that the card is

'a) the ace of hearts ? (b) a heart (c) an ace?

Solution. We define the following events :

A : the cards is the ace of hearts,.

B : the card is a heart.

C : the card is an ace.

(a) Since there are 52 cards, and there is only one ace of hearts,

$$P(A) = \frac{1}{52}$$

(b) There are 13 cards of heart. Therefore,

$$P(B) = \frac{13}{52} = \frac{1}{4}$$

(c) There are four aces amongst the 52 cards. Therefore,

$$P(C) = \frac{4}{52} = \frac{1}{13}$$

1.7. SUGGESTIONS:

After readingthis lesson student should also read the problems from the following references.

1.8. REFERENCES:

1).Mathematical Statistics by J.N. Kapur& H.C. Saxena, S.Chand& Co.

2). Introduction to mathematical Statistics by P.G. Hoel, John Wiley & sons

3). Numerical Methods, Probability & statistics by Sandip Banerjee, Books & Allied

Publications Ltd. Calcutta.

4). A text Book of Statistics, Numerical Analysis & Matrices by Sunil Gupta,Anupama Gupta, Narinder Kumar, Malhotra Brothers PaccaDanga, Jammu.

1.9.MODEL TEST PAPER

Q.1. Define Probability & Event. Give one example in each Case.

Q.2. Define Random Variable & illustrate an example.

Q.3. Define Continuous & Discrete Random variables. Give one example in each Case.

Q.4. Write probability of getting tails in tossing three coins experiment.

Q.5. Find the probability of getting

a) black card

b) red ace card

c) jack card

from a deck of 52 cards.

STATISTICS

THE IDEA OF PROBABILITY FUNCTION (P.F.)

By: Prof. Sunil Gupta

2.1 **INTRODUCTION:** in this lesson the idea of probability& distribution functions of different random variables are reported with help of suitable examples.

2.2 **OBJECTIVES:** The objective of this lesson is to explain the difference between probability & distribution functions.

2.3 **PROBABILITY FUNCTION (P.F.).**

2.3.1. **DEFINITION:** Let X be a random variable with possible values $x_1, x_2, \ldots\ldots x_n$ and associated probabilities $f(x_1)$, $f(x_2)$, .. f(xn).then the set f whose elements are the ordered pairs $(x_i, f(x_i)$, i= 1, 2,.........n is called the probability function of X if it satisfies two conditions.

i) $f(x) \geq 0, \forall\ x$ ii) $\sum_{i=1}^{n} f(x_i) = 1$

Alternate Definition of Probability Function: A probability function is a function which assigns a probability f(x)=P(X=x)to each real number x within the range of a discrete random variable X and which

2.3.2. Distribution Function. Let the random variable X takes values x_1, x_2......., x_n, with probabilities p_1, p_2. ...... p_n, and let $x_1 < x_2 < x_3 < \ldots\ldots < x_n$, then it is easily seen that

$P[X < x_{i-1}] = 0$, $P[X \le x_1] = p_1$, $P[X < x_2] = p_1$,

$P[X \le x_2] = P[X < x_2] + P[X = x_2] = p_1 + p_2$,

$P[X \le x_i] \le p_1 + p_2 + \dots\dots\dots + p_{i-1}$

$P[X \le x_i] \le p_1 + p_2 + \dots\dots\dots + p_i$

$P[X \le x_n] = p_1 + p_2 + \dots\dots\dots + p_n = 1$

We define the **distribution function** F(x) of X as :

$$F(x) = P[X \le x]$$

2.3.3. Properties of Probability or Distribution Functions

(i) $0 \le F(x) < 1$

(ii) F(x) is a monotonic increasing function of x

i.e $a \le b \Rightarrow F(a) \le F(b)$

(iii) F(x) has n points of discontinuity at x_1, x_1 .... x_n. At each of function is continuous on the right, but is discontinuous on the left.

(iv) At the point x_i

$F(x_i + 0) = \quad F(x_i) = F(x_i - 0) = p_i$

where F(x) has a jump of p_i at x_i

(v) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$,

$F(\infty) = \lim_{x \to \infty} F(x) = 1$

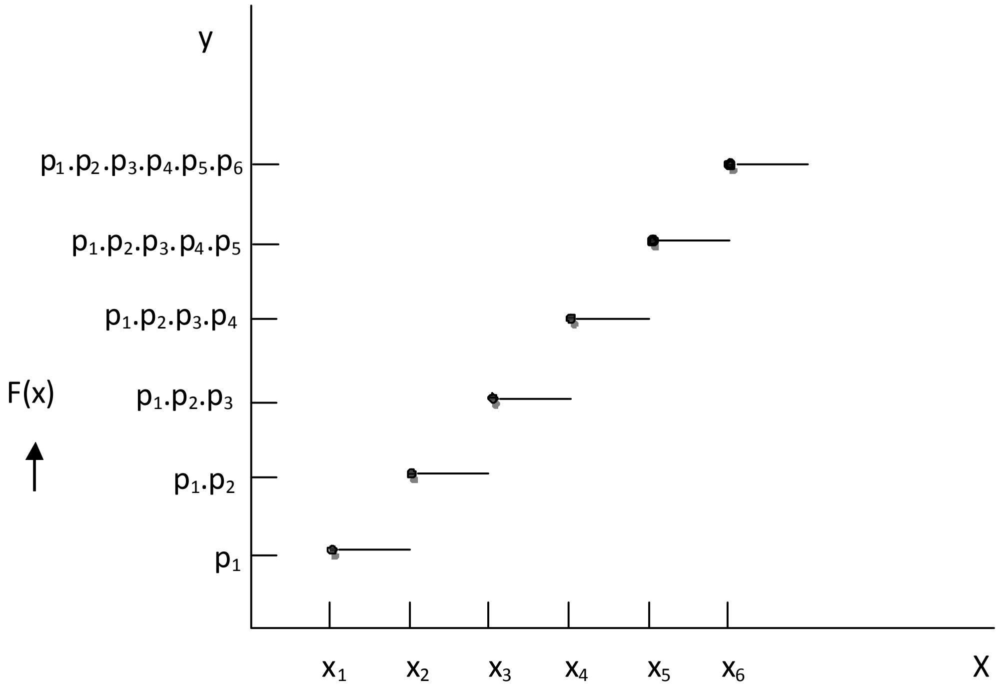


Figure 2. Distribution function of a discrete random variable.

Example Refer to the two dice experiment in lesson 01. Construct (i) the probability function (p.f) and (ii) the distribution function of the random variable X, where X denote the total obtained with the pair of dice.

Solution. (i) Refer to the sample space of the two dice experiment, (i) The possible values of the random variable X are 2, 3....., 12 and associated probabilities.

These are exhibited in the following table:

X=x	2	3	4	5	6	7	8	9	10	11	12
P[X=x]=f(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Note that the probabilities obtained in above can be obtained by the formula

$f(x) = \frac{6-|x-7|}{36}$, for x=2,3……………………12.

where the absolute value |x - 7|equals x - 7 or 7- x, whichever is positive or zero.

(ii) The distribution function F(x) is exhibited infollowing Table .

Table . Distribution function of the total obtained with the pair of dice.

X=x	2	3	4	5	6	7	8	9	10	11	12
P[X≤x]=f(x)	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}$

The above table can be written in the following form :

$$F(x) = \begin{cases} 0 & \text{for } x < 2 \\ \frac{1}{36} & \text{for } 2 \le x < 3 \\ \frac{3}{36} & \text{for } 3 \le x < 4 \\ \frac{6}{36} & \text{for } 4 \le x < 5 \\ \frac{10}{36} & \text{for } 5 \le x < 6 \\ \frac{15}{36} & \text{for } 6 \le x < 7 \\ \frac{21}{36} & \text{for } 7 \le x < 8 \\ \frac{26}{36} & \text{for } 8 \le x < 9 \\ \frac{30}{36} & \text{for } 9 \le x < 10 \\ \frac{33}{36} & \text{for } 10 \le x < 11 \\ \frac{35}{36} & \text{for } 11 \le x < 12 \\ 1 & \text{for } 12 \le x \end{cases}$$

2.3.4. Let us prove following properties:

Property 1. The function F(x) is non decreasing as x increases i.e if $x_1 < x_2$, then $F(x_1) \leq F(x_2)$.

Property 2. $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to -\infty} F(x) = 1$

Method for finding probability function from the distribution function

If the d.f of a random variable X is known, then the probability that X will lie in any specified interval of the real line can be determined from the d.f

2.3.5. Theorem For any given value x,

$P(X > x) = 1 — F(X)$.

2.3.6. Theorem For any given values x_1 and x_2 such that $x_1 < x_2$

$P(x_1 < X < x_2) = F(x_2) — F(x_1)$

Example : Suppose that a random variable X has a discrete distribution with the following p.f

$$f(x) = \begin{cases} cx, for x = 1,2,3,4,5 \\ 0, \quad elsewhere \end{cases}$$

Determine the value of the constant c.

Solution. Since f (x) as defined above is given to be thep.f, it must satisfy the following:

(i) f (x) > 0 (ii) Σ f(x) = 1

The first of the two conditions is being satisfied for x=1,2,3,4,5.

The second condition will also be satisfied jf

$$\sum_{x=1}^{5} f(x) = 1$$

or $c+2c+3c+4c+5c=1$

or $15c = 1$ or $c = \frac{1}{15}$

Example: Suppose that the p.f. of a random variable X is specified as follows :

$$f(x) = \begin{cases} 0 \;, & \text{for } x < 0 \\ 1 - e^{-x}, & \text{for } x \geq 0 \end{cases}$$

Find each of the following probabilities

(i)	$P(X \leq -2)$	(ii)	$P(X > -1)$	(iii)	$P(x = 0)$
(iv)	$P(X > 0)$	(v)	$P(X \leq 1)$	(vi)	P(-1< xd"1)

Solution (i) $P(X \leq -2) = f(-2) = 0$

(ii) $P(X > -1) = 1 — P(X \leq 1)$

$= 1 - f(-1) = 1 - 0 = 1$

(iii) $P(X = 0) = P(X \leq 0) - P(X < 0)$

$= f(0) - 0 = 1 - 0 = 0$

(iv) $P(X > 0) = 1 — P(X \leq 0)$

$= 1 - f(0) = 1 - 0 = 1$

(v) $P(X \leq 1) = f(1)$

$= 1 - e^{-1}$

(vi) $P(-1 < X_1 \leq 1) = P(X \leq 1) — P(X \leq -1)$

$= f(1) - f(-1)$

$= (1 - e^{-1}) - 0 = 1 — e^{-1}$

2.4. Examination Oriented Questions/Lesson End Exercise

Q.1. Verify that $f(x) = \frac{2x}{k(k+1)}$, x=1, 2.......k, zero elsewhere can serve as the probability function (p.f) of a random variable X.

Q.2. Let $f(x) = \frac{x}{15}$, x = 1,2,3,4,5 zero elsewhere, be the probability function (p.f) of X. Find P (X = 1 or 2), $P\{\frac{1}{2}<X<\frac{5}{2}\}$ and $P(1 \leq X \leq 2)$.

Q.3. (a) Is the function defined as follows a probability density function (p.d.f)?

$$f(x) = \begin{cases} 0, for x < 2 \\ \frac{1}{18}(3+2x), \ 2 < x \leq 4 \\ 0, for x < 4 \end{cases}$$

(b) Find the probability that a variate having this probability density function (p.d.f) fall in the interval $2 \leq x \leq 3$.

Q.4. For each of the following probability density functions of X compute P(|X|<1) and $P(X^2 < 9)$.

(a) $f(x) = \begin{cases} \frac{x^2}{18}, -3 < x < 3 \\ 0, \quad elsewhere \end{cases}$

(b) $f(x) = \begin{cases} \frac{x+2}{18}, -2 < x < 4 \\ 0, \quad elsewhere \end{cases}$

Q.5. Let f(x) be the p.f of a random variable X. Find the distribution function X and sketch its graph if

$$f(x) = \begin{cases} \frac{x}{6}, x = 1,2,3 \\ 0, \quad elsewhere \end{cases}$$

Q.6. Let f(x) be the p.d.f of a random variable X. Find the distribution function X and sketch its graph if

Q.7. Let the distribution function be given by

$$f(x)=\begin{cases} \frac{2}{x^3}, 1 < x < \infty \\ 0, \quad elsewhere \end{cases}$$

Find the probability P(-3 < X 1/2) and P(X =0).

Q.8. Define probability function (p.f.) of a discrete random variable X.A coin is tossed thrice, find the probability function of the number of heads obtained and draw its graph.

Q.9. Define the distribution function of a random variable X. List its important properties and prove any one of them.

Q.10. Suppose that the distribution function of a random X is as follows :

$$F(x) = \begin{cases} 0, x \leq 0 \\ \frac{1}{9}x^2, 0 < x \leq 3 \\ 1, \ for x > 3 \end{cases}$$

Find the probability density function of X.

Q.11. Give example of continuous and discrete random variables. Explain the terms probability density function (p.d.f) and distribution function (d.f).

ANSWERS

3. (a) yes (b)$\frac{4}{9}$

4.(a) $\frac{1}{27}$,1 (b)$\frac{2}{9}, \frac{25}{36}$

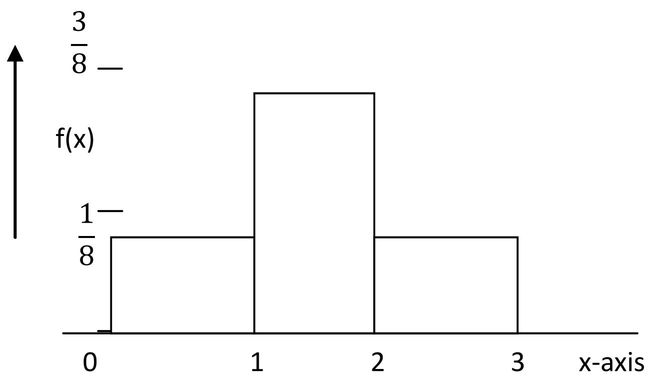
6. F(x) = $\begin{cases} 0, & x < 1 \\ \frac{1}{6}, & 1 \leq x < 2 \\ 1, & 3 \leq 1 \end{cases}$

7.$\frac{3}{4}, \frac{1}{2}$

8.

X=x	0	1	2	3
F(x)-P(X=x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

10. f(x) = $\frac{d}{dx}$F(x) = $\begin{cases} \frac{2}{9}x & for\ 0 < x \leq 3 \\ 0 & elsewhere \end{cases}$



2.5. SUGGESTIONS: After readingthis lesson student should also read the problems from the following references.

2.6. REFERENCES:

1). Mathematical Statistics by J.N. Kapur& H.C. Saxena, S.Chand& Co.

2). Introduction to mathematical Statistics by P.G. Hoel, John Wiley & sons

3). Numerical Methods, Probability & statistics by Sandip Banerjee, Books & Allied Publications Ltd. Calcutta.

4). A text Book of Statistics, Numerical Analysis & Matrices by Sunil Gupta,Anupama Gupta, Narinder Kumar, Malhotra Brothers PaccaDanga, Jammu.

2.7. MODEL TEST PAPER

Q.1. Define Probability & Distribution functions. Give one example in each Case.

Q.2. Find probability function of getting no. of divisors from integers 1,2, …..10.

Q.3. Write important properties of probability functions.

Q.4. Define probability function (p.f.) of a discrete random variable X.A coin is tossed thrice, find the probability function of the number of heads obtained and draw its graph.

Q.5. (a) Is the function defined as follows a probability density function (p.d.f) ?

$$f(x) = \begin{cases} 0, \; for\, x < 2 \\ \frac{1}{18}(3+2x), \; 2 \le x \le 4 \\ 0, \; for\, x < 4 \end{cases}$$

(b) Find the probability that a variate having this probability density function (p.d.f) fall in the interval 2d” xd”3.

Q.6. Suppose that a random variable X has a discrete distribution with the following p.f

$$f(x) = \begin{cases} cx, for x = 1,2,3,4,5 \\ 0, \quad elsewhere \end{cases}$$

Determine the value of the constant c.

B.A. III Semester	Unit-I
Course No. : UGM 0301	Lesson 3

STATISTICS

BERNOULLI S DISTRIBUTION

By: Prof. Sunil Gupta

3.1 INTRODUCTION: In this lesson the problems for finding arbitrary constants from probability & distribution functions are discussed.

3.2 OBJECTIVES: The objective of this lesson is to explain the idea to find constants from probability functions. Also how to construct the distribution function from probability function & vice versa

3.3 BERNOULLI DISTRIBUTION

This probability distribution applies when an experiment has two possible outcomes(often referred to an 'failure and 'success'). If the probabilities of failure and success are,respectively,1-θ and θ, and we code these two outcomes as 0(zero success) and 1 (one success), the probability function of the Bernoulli distribution can be written as

$f(x;\theta) = \theta^x(1-\theta)^{1-x}$, for x=0,1 ....(1)

(1) Can be represented in the form as shown in table 7.

Table. Probability function of Bernoulli distribution

X=x	0	1
P(X=x)=f(x)	1-θ	θ

3.4 : Continuous Distributions - A random variable X is said to be a continuous distribution if there exists a non-negative function f, defined on the real line, such that for any interval A,

P(X∈A)=$\int_A f(x)$dx

The functionf is called the probability density function (p.d.f.) of X. Thus, if a random variable X has a continuous distribution, the probability that X will belong to any subset of thereal line can be found by integrating the p.d.f of X over that subset. Every p.d.f must satisfy following two requirements:

(1) f(x)>0and

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

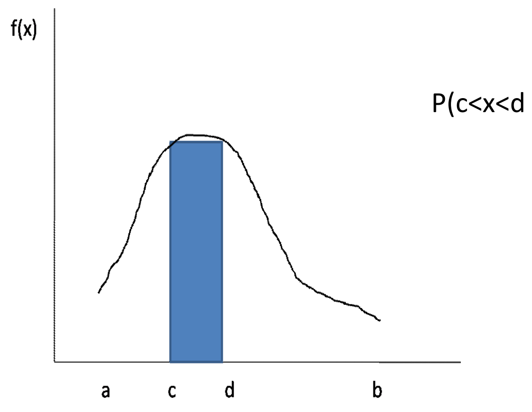


Figure 4.

Furthermore, if f(x) satisfies (i) and (ii), then the probability that variable X lies between c and d (a < c <d < b), denoted by P(c<X<d),is given by P (c <X < d) = Area under the curve f(x) bounded by the x-axis and the two ordinates x=c and x=d

Again if a random variable X has a continuous distribution, then P(X=x)=0,for every individual value x, and the d.f of X is a continuous function over the entire line. Furthermore, at any point x at which the p.d.f is continuous, the d.f. is differentiable and

$$\frac{dF(x)}{dx} = f(x)$$

3.5 EXAMPLES

Example 1.: Suppose that the p.d.f of a certain random variable X has the following form

$$f(x) = \begin{cases} cx, & for 0 < x < 4 \\ 0, & elsewhere \end{cases}$$

where c is a given constant. Determine: the value of c and then determine the values of P (1≤ X ≤ 2) and P (X >2). $=\int_c^d f(x)dx$

Solution. For every p.d.f it must be true that

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

Therefore, in this example

$c\int_0^4 x dx=1$ or $8c=1$

here $c=\frac{1}{8}$

Also P (1≤ X ≤2)=$\int_1^2 f(x)dx = \frac{1}{8}\int_1^2 x dx = \frac{3}{16}$

P(x>2)=$\int_2^4 f(x)dx = \frac{1}{8}\int_2^4 x dx = \frac{3}{4}$

f(x)= $\begin{cases} cx^2, for\ 1 \le x \le 2 \\ 0\ , \quad elsewhere \end{cases}$

(a) Find the value of the constant c

(b) Find the value of P {X>$\frac{3}{2}$}

Solution (a) For every p.d.f it must be true that

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

Therefore, in this example

c$\int_1^2 x^2$dx=1

$\frac{7c}{3}$ =1 or c=$\frac{3}{7}$

(b) P[X>$\frac{3}{2}$]=$\int_{3/2}^2 f(x)$dx=$\frac{3}{7}\int_{3/2}^2 x^2$dx

=$\frac{3}{7}.\frac{1}{3}\{8-\frac{27}{8}\}= \frac{1}{7}.\frac{37}{8}=\frac{37}{56}$

Example 3. : Suppose that the d.f. of a random variable X is as follows:

F(x) = $\begin{cases} 0, forx \le 0 \\ \frac{1}{9}x^2, for 0 < x \le 3 \\ 1\ , forx > 3 \end{cases}$

Find the p.d.f. of X.

Solution.

f(x)= $\frac{dF(x)}{dx}$ = $\begin{cases} 0\ , forx \le 0 \\ \frac{2}{9}x, for 0 < x \le 3 \\ 0, forx > 3 \end{cases}$

The above can be re-written as

$$f(x)=\begin{cases} \frac{2}{9}x, for 0 < x \le 3 \\ 0, \quad elsewhere \end{cases}$$

3.6 EXAMINATION ORIENTED EXERCISES/LESSON END EXERCISE

Q.1. For each of the following, find the constant c so that f(x) satisfies the conditions of being a probability function (p.f) of one random variable X.

(a) $f(x) = cx^2$, x=1,2 .... k, zero elsewhere.

(b) $f(x) = c\left(\frac{2}{3}\right)^x$, x=l,2,3 .... zero elsewhere.

Q.2. Find a constant c so that f(x) satisfies the conditions of being a p.d.f of one randomvariable X.

$$f(x)=\begin{cases} cxe^{-x}, 0 < x < \infty \\ 1, \quad elsewhere \end{cases}$$

Q.3. A random variable X has the following probability function:

X	:	—2	—1	0	1	2	3
P(x)	:	0.1	k	0.2	2k	0.3	3k

(i) Find k

(ii) Evaluate P(X<2), P(X>2), P(- 2 < X <2)

(iii) Determine the distribution function F(x) of X.

Solution. (i)Since $\sum_{x=-2}^{3} P(x)$=1 gives 0.1 + k+ 0.2 + 2k+0.3 +3k= 1

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

(ii)P(X<2) =P(X=- 2)+P(X= - 1)+P(X=0)+P(x=1)

= 0.1 + k + 0.2 + 2k

= 3k + 0.3=0.5

$P(X\geq 2) = P(x=2)+P(x=3)$

$=0.3 +3k =0.3+0.2=0.5$

$P(X\geq 2) =1 — P (X < 2) = 1 \quad -0.5= 0.5$

$P(-2<X<2)=P(X= -1)+P(X=0)+P(X=1)$

$=k+0.2+2k=3k+0.2$

$= 0.1 + 0.2 = 0.3$

The distribution function F(x) is exhibited in the table given below:

X	-2	-1	0	1	2	3
P(x)	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{2}{10}$	$\frac{2}{15}$	$\frac{3}{10}$	$\frac{3}{15}$
$F(x)=P(X\leq x)$	$\frac{1}{10}$	$\frac{5}{30}$	$\frac{11}{30}$	$\frac{15}{30}$	$\frac{24}{30}$	$\frac{30}{30}$

3.7. SUGGESTED READING The students are advised to go through following books for references

3.8. REFERENCES:

1).Mathematical Statistics by J.N. Kapur& H.C. Saxena, S.Chand& Co.

2). Introduction to mathematical Statistics by P.G. Hoel, John Wiley & sons

3). Numerical Methods, Probability & statistics by Sandip Banerjee, Books & Allied Publications Ltd. Calcutta.

4). A text Book of Statistics, Numerical Analysis & Matrices by Sunil Gupta,Anupama Gupta, Narinder Kumar, Malhotra Brothers PaccaDanga, Jammu.

3.9. MODEL TEST PAPER

Q. 1. Find a constant c so that f(x) satisfies the conditions of being a p.d.f of one random variable X.

$$f(x)=\begin{cases} cxe^{-x}, 0 < x < \infty \\ 1, \quad elsewhere \end{cases}$$

Q.2. Define Bernoulli distribution & write its Probability function.

Q.3. . A random variable X has the following probability function:

X	:	—2	—1	0	1	2	3
P(x)	:	0.1	k	0.2	2k	0.3	3k

(i) Find k

(ii) Evaluate P(X<2), P(X>2), P(- 2 < X <2)

(iii) Determine the distribution function F(x) of X.

Q.4. Suppose that the p d f of a random variable X is as follows:

$$f(x)=\begin{cases} cx^2, for\ 1 \le x \le 2 \\ 0\ , \quad elsewhere \end{cases}$$

(a) Find the value of the constant c

(b) Find the value of P {X>$\frac{3}{2}$}

STATISTICS

EXPECTATIONS AND MOMENTS

By: Dr. Sunil Gupta

4.1. INTRODUCTION: In this lesson the concept of mean, variance, moments, moment generating functions etc. are discussed. The concept is explained in a simpler way.

4.2. OBJECTIVES: Objective of studying this lesson is to give the idea of finding mean & variance of different types of functions both discrete & continuous functions. awareness among the students about the concept of.

4.3. MEAN OF FUNCTION :

4.3.1. Definition→Let x be random variable with p.f. f(x). Then expectation of function, g(x), denoted and defined by

$$E(g(x))=\begin{cases} \sum_{x=-\infty}^{\infty} g(x)f(x), if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} g(x)f(x)dx, if\ x\ is\ continuous \end{cases}$$

4.3.2.Remark : if we take g(x)=X , we call it s expectation of X or mean of X, denoted by μ and is defined as

$$\mu = E(x) = \begin{cases} \sum_{x=-\infty}^{\infty} xf(x), if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} xf(x)dx, if\ x\ is\ continuous \end{cases}$$

4.4. r^{th} moment

4.4.1.Definition:→ Let X be random variable with p.f. f(x), then the rth moment about mean= (μ), denoted and defined by

$$\mu_r = E[(X-\mu)^r] = \begin{cases} \sum_{x=-\infty}^{\infty}(x-\mu)^r f(x), if\ x\ is\ discrete \\ \int_{-\infty}^{\infty}(x-\mu)^r f(x)dx, if\ x\ is\ continuous \end{cases}$$

NOTE (i) If we take $\mu = 0$ in above definition, we get r^{th} moment about 0 (or origin)

$$\mu'_r = E(x^r) = \begin{cases} \sum_{x=-\infty}^{\infty} x^r f(x), if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} x^r f(x)dx, if\ x\ is\ continuous \end{cases}$$

ii) if we take r=2 in above definition, we get

$$\mu_2 = E[(X-\mu)^2] = \begin{cases} \sum_{x=-\infty}^{\infty}(x-\mu)^2 f(x), if\ x\ is\ discrete \\ \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx, if\ x\ is\ continuous \end{cases}$$

Is called **variance of X**and is written as var(x)

i.evar(X) = μ_2 (ii^{nd} moment about mean)

Example: Find mean and variance of number of tails appearing in tossing three coins experiment.

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

Let X, denote no.of tails

∴ X(HHH)=0

X(HHT)=1=X(HTH)=X(THH)

X(HTT)=2=X(THT)=X(TTH)

X(TTT)=3

∴ X takes values 0,1,2,3

prob. functionf(x)of f(x)

X=x	0	1	2	3
f (x)	1/8	3/8	3/8	1/8

Hence X is discrete random variable

Now mean = (μ)= E(x)=$\sum_{x=-\infty}^{\infty} xf(x)$

=$\sum_{x=0}^{3} xf(x)$

=0f(0) + 1f(1)+2f(2)+3f(3)

= $0(\frac{1}{8}) + 1(\frac{3}{8}) + 2(\frac{3}{8}) + 3(\frac{1}{8}) = \frac{12}{8} = \frac{3}{2}$

Variance (x)= μ_2= E[(x-μ)^2]

= $\sum_{x=-\infty}^{\infty}(x-\mu)^2 f(x)$

But $\mu = \frac{3}{2}$

= $\sum_{x=0}^{\infty}(x-\frac{3}{2})^2 f(x)$

=$(0-\frac{3}{2})^2 f(0) + (1-\frac{3}{2})^2 f(1) + (2-\frac{3}{2})^2 f(2) + (3-\frac{3}{2})^2 f(3)$

=$\frac{9}{4}.\frac{1}{8} + \frac{1}{4}.\frac{3}{8} + \frac{1}{4}.\frac{3}{8} + \frac{9}{4}.\frac{1}{8}$

=$\frac{24}{32} = \frac{3}{4}$

Example: Find mean and variance of total number of divisors appearing in 1,2,3,4...........10.

Solution : Let S={1,2,3,4.....................10 } and

let X= total no. of divisors

X(1)=1

X(2)=2= x(3)= x(5)=x(7)

X(4) =3= x(9)

X(6)=4= x(8)=x(10)

∴ X takes values 1,2,3,4

Its p.f.

X=x	1	2	3	4
f (x)	1/10	4/10	2/10	3/10

Here x is discrete random variable

Now mean = μ = E(X)=$\sum_{X=-\infty}^{\infty} x\, f(x)$

=$\sum_{X=1}^{4} x\, f(x)$

=1f(1)+2f(2)+3f(3)+4f(4)

=$1(\frac{1}{10})$ +$2(\frac{4}{10})$ +$3(\frac{2}{10})$ +$4(\frac{3}{10})$

=$\frac{1}{10}+\frac{8}{10}+\frac{6}{10}+\frac{12}{10} = \frac{27}{10}$

μ =2.7

Variance (X) = μ_2 = [E(x- μ)^2]

= $\sum_{x=-\infty}^{\infty}[(x-\mu)^2$ f(x), here μ = 2.7

$=\sum_{X=1}^{4}(x-2.7)^2$f(x)

$= (1-2.7)^2$f(1) $+(2-2.7)^2$ f(2)$+(3-2.7)^2$ f(3) $+(4-2.7)^2$f(4)

$=(-1.7)^2(\frac{1}{10})+(-.7)^2(\frac{4}{10}) + (.3)^2(\frac{2}{10}) + (1.3)^2(\frac{3}{10})$

$=\frac{1010}{10}=101$

Example: Find mean and variance of X for following, if exists.

i) $f(x) = \begin{cases} \frac{1}{x} \;, & \text{if } 1 < x < \infty \\ 0 \,, & otherwise \end{cases}$

ii) $F(x) = \begin{cases} 1, x \le 0 \\ \frac{x+1}{2}, & 0 < x < 2 \\ 0 \,, x \ge 2 \end{cases}$

Solution **(i)** Here f(x) is p,f for continuous r.v ,so

μ = mean(x)= E(x) = $\int_{-\infty}^{\infty} xf(x)$dx

$= \int_{1}^{\infty} x.\frac{1}{x}$ dx

$= [x]_1^{\infty}$

$=\infty-1=\infty$

$\therefore$ mean of X doesnot exist

Since

var(X)= μ_2 =E[(x-μ)^2]

&$\mu = \infty$, i.e μ doesnot exist, so var(x) does not exists

(ii) here x is cont r,v &

$f(x) = \frac{d}{dx}F(x)$

$$= \begin{cases} 0, x \le 0 \\ \frac{1}{2}, 0 < x < 2 \\ 0, x \ge 2 \end{cases}$$

$$= \begin{cases} \frac{1}{2}, 0 < x < 2 \\ 0, \quad otherwise \end{cases}$$

μ (mean x) $= E(x) = \int_{-\infty}^{\infty} x\, f(x)dx$

$= \int_0^2 x.\frac{1}{2}dx = \frac{1}{2}\int_0^2 x.dx$

$= \frac{1}{2} \left| \frac{X^2}{2} \right|_0^2 \qquad = \frac{1}{4}\ (2^2 - 0^2) = 1$

var (x) $= \mu_2 = E[(x-\mu)^2]$

$= \int_{-\infty}^{\infty}(x - \mu)^2 f(x)\, dx$

$= \int_0^{\infty}(x - \mu)^2 \frac{1}{2}\, dx$

But $\mu = 1$

$= \frac{1}{2}\int_0^2 (x - 1)^2 dx$

$= \frac{1}{2} \left| \frac{(x-1)^3}{3} \right|_0^2$

$= \frac{1}{6}[(2-1)^3 - (0-1)^3]$

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

4.5. Moment Generating Function (m.g.f)

4.5.1.Definition Let X be a r.v with p.f. f(x), then m.g.f of X w.r.t 't' denoted and defined by m.g.f of X= $M_x(t)$

$$= E[e^{tx}] = \begin{cases} \sum_{x=-\infty}^{\infty} e^{tx} f(x), if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} e^{tx} f(x)dx, if\ x\ is\ continuous \end{cases}$$

Example :-Find expansion of m.g.f and hence find moments.

Solution : Let X be r.v with p.f. f(x)

Case I when X is discrete

Then m.g.f. of X = $M_x(t)$

$=E[e^{tx}]=\sum_{x=-\infty}^{\infty} e^{tx}\ f(x)$

$$e^{\theta} = 1+\frac{\theta}{1!}+\frac{\theta}{2!}+\frac{\theta}{3!}\text{------------------}$$

$=\sum_{x=-\infty}^{\infty}[1 + \frac{t(x)}{1!} + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \ldots\ldots\ldots\ldots]f(x)$

$=\sum_{x=-\infty}^{\infty}[f(x) + txf(x) + \frac{t^2x^2f(x)}{2!} + \frac{t^3x^3f(x)}{3!} + \ldots\ldots\ldots]$

$=\sum_{x=-\infty}^{\infty} f(x) + t\sum_{-\infty}^{\infty} xf(x) + \frac{t^2}{2!x^2}\sum_{-\infty}^{\infty} x^2 f(x)+\frac{t^2}{3!}\sum_{-\infty}^{\infty} x^3 f(x)+\ldots\ldots\ldots$

$=1+tE(x)+ \frac{t^2}{2!}E(x^2) + \frac{t^3}{3!}E(x^3)+\ldots\ldots\ldots\ldots$

$=1+t\mu_1' + \frac{t^2}{2!}\mu_2' + \frac{t^3}{3!}\mu_3'+\ldots\ldots\ldots(*)$

Is required expansion of m.g.f.

Case II : when X is continuous, we get same result.

For moments,

Diff(*) b/s w.r.t 't'

$M_x'(t) = \mu_1' + t\,\mu_2' + \frac{t^2}{2!}\mu_3' + \ldots\ldots\ldots\ldots(1)$

Put t=0

= first moment about origin

$M_x'(t)\,|_{t=0} = \mu_1'$

Differentiate (1) b/s w.r.t 't'

$M_x''(t)\,| = \mu_2' + \frac{2t\mu_3'}{2} + \ldots\ldots\ldots$

Put t=0

$M_x''(t)\,|_{t=0} = \mu_2'$ = II[nd] moment about origin

& so on $M_x^r(t)\,|_{t=0} = \mu_r'$ =r[th] moment about origin.

4.5.2. THEOREM Let X and Y be r.v & a,b are constants. Then

i) $M_{x+y}(t) = [M_x(t)][M_y(t)]$

ii) $M_{x+a}(t) = e^{at}M_x(t)$

(iii) $M_{xb}(t) = M_x(bt)$

(iv) $M_{\frac{x+a}{b}}(t) = e^{\frac{a}{b}t}\, M_x\left(\frac{t}{b}\right)$

<u>**Proof :**</u> (ii) $M_{x+a}(t) = E[e^{t(x+a)}]$

$=E[e^{tx+at}]$

$= E[e^{tx}. e^{at}]$

$= e^{at} E(e^{tx})$

$= e^{at} M_x(t)$

$= e^{at} M_x(t)$

(iii) $M_{xb}(t) = E[e^{t(xb)}]$

$= E[e^{(tb)x}]$

$= M_x(bt)$

(iv) $M_{\frac{x+a}{b}}(t) = E[e^{\frac{t(x+a)}{b}}]$

$= E[e^{\frac{t(x+a)}{b}}]$

$= E[e^{\frac{tx}{b}}.e^{\frac{ta}{b}}]$

$= e^{\frac{at}{b}} E[e^{\frac{tx}{b}}]$

$= e^{\frac{at}{b}} (E\, e^{\left(\frac{t}{b}\right)x})$

$= e^{\frac{at}{b}} M_x(\frac{t}{b})$

(i) $M_{x+y}(t) = E[e^{t(x+y)}]$

$= E[e^{tx}.\, e^{ty}]$

$= E(e^{tx}.\, e^{ty}]$

$= [M_x(t)]\, [M_y(t)]$

4.6. Examination Oriented Exercise/ Lesson End Exercise

Q.1. Define Mean, variance of a function.

Q.2. Find expansion of m.g.f

Q.3. Find mean and variance of X for following, if exists.

$$f(x) = \begin{cases} \frac{1}{x}, & \text{if } 1 < x < \infty \\ 0, & otherwise \end{cases}$$

Q.4. Find mean and variance of number of tails appearing in tossing three coins experiment.

4.7.SUGGESTED READING

The students are advised to go through following references for details

4.8. REFERENCES

1). Mathematical Statistics by J.N. Kapur & H.C. Saxena, S.Chand & Co.

2). Introduction to mathematical Statistics by P.G. Hoel, John Wiley & sons

3). Numerical Methods, Probability & statistics by Sandip Banerjee, Books & Allied Publications Ltd. Calcutta.

4). A text Book of Statistics, Numerical Analysis & Matrices by Sunil Gupta,Anupama Gupta, Narinder Kumar, Malhotra Brothers Pacca Danga, Jammu.

4.9. MODEL TEST PAPER

Q.1. Find variance of no. of dots appearing on top of six dice when rolled.

Q.2. A gambler is playing with a dice. He is promised

(i) the sum

(ii) the product of number thrown

Q.3. Find mean and variance of number of tails appearing in tossing three coins experiment.

Q.4. If 'x' has p.d.f. given as

$$f(x)=\begin{cases}\frac{x+1}{2}, -1 < x < 1\\ 0, elsewhere\end{cases}$$

then find mean of x.

STATISTICS

EXPECTATION OF THE SUM AND PRODUCT OF POINTS WHEN 'N' DICE ARE THROWN.

By: Dr. Sunil Gupta

5.1. INTRODUCTION: In this lesson properties of mean, variance, moments such as sum, product etc. are discussed. The idea is very useful for students.

5.2. OBJECTIVES: Objective of studying this lesson is to aware students about the concepts of addition, multiplication of finding mean, variance, moments of the functions when more than one experiment are performed.

5.3. PROPERTIES OF MEAN

5.3.1.Theorem:- let x,y be r.v with p.f. f(x). Then

(i) E(1) =1

(ii) E(a)=a,a is any real number.

(iii) E(x+y) =E(x) +E(y)

Solution : Let x,y be r.v with p.f. f(x). Then

(i) E(1) = $\begin{cases} \sum_{x=-\infty}^{\infty} 1.f(x), if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} f(x)dx, if\ x\ is\ continuous \end{cases}$

$= \begin{cases} \sum_{-\infty}^{\infty} f(x), if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} f(x)dx, if\ x\ is\ continuous \end{cases}$

=1 $\because$ f(x) is p.f

(ii) E(a) = $\begin{cases} \sum_{x=-\infty}^{\infty} af(x), if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} af(x)dx, if\ x\ is\ continuous \end{cases}$

= a$\begin{cases} \sum_{x=-\infty}^{\infty} f(x), if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} f(x)dx, if\ x\ is\ continuous \end{cases}$

=a.1 $\because$ f(x) is p.f

=a

(iii) E[(x+y)] = $\begin{cases} \sum_{x,y=-\infty}^{\infty}(x+y)f(x), if\ x,y\ is\ discrete \\ \int_{-\infty}^{\infty}(x+y)f(x)dx, if\ x,y\ is\ continuous \end{cases}$

=$\begin{cases} \sum_{x=-\infty}^{\infty} xf(x) + \sum_{y=-\infty}^{\infty} yf(x)\ , if\ x,y\ is\ discrete \\ \int_{-\infty}^{\infty} xf(x)dx + \int_{y=-\infty}^{\infty} yf(x)\ \ dx, if\ x,y\ is\ continuous \end{cases}$

=$\begin{cases} \sum_{x,y=-\infty}^{\infty} xf(x), if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} xf(x)dx, if\ x\ is\ continuous \end{cases}$ + $\begin{cases} \sum_{x,y=-\infty}^{\infty} yf(x), if\ y\ is\ discrete \\ \int_{-\infty}^{\infty} yf(x)dx, if\ y\ is\ continuous \end{cases}$

=E(x) +E(y)

Example: let x be r.v with p.f. f(x) and 'a' is constant. Then E(ax)=aE(x)

Solution E(ax)= $\begin{cases} \sum_{x=-\infty}^{\infty} af(x)\ , if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} af(x)\ , if\ x\ is\ continuous \end{cases}$

= a$\begin{cases} \sum_{x=-\infty}^{\infty} f(x)\ , if\ x\ is\ discrete \\ \int_{-\infty}^{\infty} f(x)\ , if\ x\ is\ continuous \end{cases}$

= aE(x)

Example : Find mean of product of getting number of dots in throwing n die simultaneously

Solution: let $X_1, X_{2,}$.................... X_n be n dice throwing simultaneously

We have to find mean of $(x_1,x_2.......x_n)$

i.e $E[x_1,x_2,........x_n]$

we know

$E[x_1,x_2.......x_n] = E(x_1)\, E(x_2)+.................. E(x_n)$.......(1)

Let X_1 : no of dots in 1st dice

X=x	1	2	3	4	5	6
f(x)	1/6	1/6	1/6	1/6	1/6	1/6

Here X= 1,2.......6 is discrete

$E(x_1)= \sum_{x=1}^{6} x_1 f(x)$

=1.f(1) +2.f(2)+3.f(3)+4.f(4)+5.f(5)+6.f(6)

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

$=\frac{z1}{6} = \frac{7}{2}$ use in (1)

$E(x_1,x_2..........x_n) =(\frac{7}{2})\, (\frac{7}{2})..........(\frac{7}{2})$

$=(\frac{7}{2})^n$

5.3.2.Theorem let x,y be a r.v & a,b are constants. Then

(i) var X = $E(x)^2 - [E(x)]^2$

or

$\mu_2 = \mu_2' - \mu_1'^2$

(ii) var (ax+b) = a^2 var(x)

(iii) var (x+y) = var x+ var y

Proof: We know

$\mu_r = E[(x-\mu)^r]$ ----------------(1)

& $\mu_r' = E(X^r)$-----------------(2)

Var x = $\mu_2 = E[(x-\mu)^2]$

$=E[x^2-2\mu x+\mu^2]$

$= E[x^2]- E[2\mu x] +E[\mu]^2$

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

But $\mu = E(x)$

$\mu_2 = E(x^2)-2 E(x) E(x)+(E(x))^2$

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

Var x = $\mu_2= E(x^2)- [E(x)]^2$

=> $\boxed{\mu_2= \mu_2' - \mu_1'^2}$ using (2)

(ii) we know

$Var(x) = E(x^2) - E(x)^2$

Put x= ax+b

$Var(ax+b) = E[(ax+b)^2]- E[(ax+b)]^2$

$= E[a^2x^2+b^2 +2axb]-[aE(x)+b]^2$ (E(b) =b)

$= [(a^2E(x^2) +b^2 + 2abE(x)] - [(aE(x))^2 +b^2 + 2aE(x)b]$

$= a^2E(x^2) +b^2+ 2abE(x)-a^2(E(x)^2 - b^2 - 2abE(x)$

$=a^2E(x^2) - a^2(E(x))^2$

$= a^2E(x^2) - a^2 (E(x))^2$

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

$Var(ax+b) = a^2 var(x)$

(iii) we know

$Var (x)= E(x^2)-[E(x)]^2$ -------------------------- (*)

Put x=x+y

$Var(x+y) = E(x+y)^2 - [E(x+y)]^2$

$= E(x^2) + E(y^2) + 2E(xy) - [(E(x))^2 + (E(y))^2 + 2E(x)E(y)]$

$= E(x^2) + E(y^2) + \cancel{2E(xy)} - (E(x))^2 - (E(y))^2 - \cancel{2E(x)E(y)}$

$= (E(x^2) - (E(x))^2) + (E(y^2) - (E(y))^2)$

$\Rightarrow$ **Var(x+y) = Var x + Var y** using (*)

$\mu = E(x) = \mu_1'$

5.3.3.Article Convert first three moment

(i) about mean into origin

(ii) about origin into mean

About Mean $\mu_r = E[(x-\mu)^r]$
About origin
$\mu'_r = E(x)^r$

Proof

(i) First moment about mean

$\mu_1 = E[(x-\mu)] = E(x) - E(\mu)$

$= \mu - \mu$

$= 0$

2nd moment about mean

$\mu_2 = E[(x-\mu)^2] = \mu_2' - \mu_1'^2$ (see (i) part of above theorem)

3rd moment about mean

$\mu_3 = E[((x-\mu)^3]$

$= E[x^3 - 3\mu x^2 + 3\mu^2 x - \mu^3]$

$= E(x^3) - 3\mu E(x)^2 + 3\mu^2 E(x) - \mu^3$

$= E(x^3) - 3 E(x) E(x)^2 + 3(E(X))^2 E(x) - (E(X))^3$

$= \mu'_3 - 3\mu_1' \mu_2' + 3\mu_1'^2 \mu_1' - (\mu_1')^3$

$= \mu_3' - 3\mu_1' \mu_2' + 3\mu_1'^3 - \mu_1'^3$

$= \mu_3' - 3\mu_1' \mu_2' + 2\mu_1'^3$

(ii) First moment about origin $= \mu_1' = E(x) = \mu$

$\mu_1' = \mu$

IInd moment about origin

$\mu_2' = E(x^2) = E[(x-\mu+\mu)^2]$

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

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

$= \mu_2 + \mu^2 + 2\mu\mu_1$

$= \mu_2 + \mu^2 + 0 \because \mu_1 = 0$

$\mu_2' = \mu_2 + \mu^2$

IIIrd moment about origin

$\mu_3' = E(x^3) = E[(x-\mu+\mu)^3]$

$= E[(x-\mu)^3 + \mu^3 + 3\mu^2(x-\mu) + 3\mu(x-\mu)^2]$

$= E[(x-\mu)^3 + E(\mu^3) + 3\mu^3 E(x-\mu) + 3\mu(x-\mu)^2]$

$= \mu_3 + \mu^3 + 3\mu^2\mu_1 + 3\mu\mu_2$ But $\mu_1 = 0$

$\mu_3' = \mu_3 + \mu^3 + 3\mu\mu_2$

Example: Find m.g.f. of Bernoulli's distribution X whose p.f. is

$$f(x) = \begin{cases} \theta^x(1-\theta)^{1-x}, x = 0,1 \\ 0, otherwise \end{cases}$$

Also find mean &variance.

Solution : m.g.f of X= $M_x(t) = E(e^{tx})$

$=\sum_{x=-\infty}^{\infty} e^{tx} f(x)$

$=\sum_{x=0,1} e^{tx}\theta^x(1-\theta)^{1-x}$

$=\sum_{x=0,1}(\theta e^t)^x(1-\theta)^{1-x}$

$=\sum_{x=0,1}(\theta e^t)^0(1-\theta)^{1-0} + (\theta e^t)^1(1-\theta)^{1-1}$

$=(1-\theta) + \theta e^t$ ------------------------ (A)

Differentiate w.r.t 't'

$M_x'(t) = \theta e^t$

Put t=0

$\mu_1' = M_x'(t)|_{t=0} = \theta e^0 = \theta$

Variance (x)= $\mu_2' - \mu_2'^2$ ---------------(1)

Differentiate $M_x'(t)$ w.r.t 't'

$\mu_1' = M_x'(t)\,|_{t=0} = \theta e^0 = \theta$

Mean $\mu = \mu_1' = \theta$

Variance (x)= $\mu_2' - \mu_2'^2$ ---------------(1)

Differentiate $M_x'(t)$ w.r.t 't'

$M_x''(t) = \theta e^t$

Put t=0

$\mu_2' = M_x''(t)\,|_{t=0} = \theta$. Use in (1)

Var (x) = $\theta - \theta^2$

Example 1 Find expected values of X which takes the values $x_i = \frac{(-1)^i 2^i}{i}$ with probabilities $p_i = 2^{-i}$, i=1,2,3............

Example 2 Find mean, variance and m.g.f of distribution function $\frac{1}{k}$

$$F(x) = \begin{cases} \frac{1}{k}, & x = 1,2,\ldots\ldots\ldots..k \\ 0, & otherwise \end{cases}$$

Solution 1 : Since x is defined for discrete random variable So

Expected value of x= E(x)

$E(x) = \sum_{x=-\infty}^{\infty} x\, f(x)$

$= \sum_{i=1}^{\infty} x_i\, p_i$

$= \sum_{i=1}^{\infty} \frac{(-1)^i 2^i}{i} \times 2^{-i}$

$= \sum_{i=1}^{\infty} \frac{(-1)^i}{i}$

$= \frac{(-1)^1}{1} + \frac{(-1)^2}{2} + \frac{(-1)^3}{3} + \frac{(-1)^4}{4} + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

$= -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

Solution 2 : Here x is a discrete r.v

m.g.f of X = $M_x(t) = E(e^{tx}) = \sum_{x=-\infty}^{\infty} e^{tx} f(x)$

$= \sum_{x=1}^{k} e^{tx} . \frac{1}{k}$

$= \frac{1}{k} \sum_{x=1}^{k} e^{tx}$

$= \frac{1}{k} [e^t + e^{2t} + e^{3t} \ldots\ldots\ldots\ldots\ldots\ldots + e^{kt}]$------------------------(1)

$= \frac{1}{k} e^t (1 + e^t + e^{2t} + e^{3t} \ldots\ldots\ldots\ldots\ldots\ldots + e^{(k-1)t})$

$= \frac{e^t}{k} 1. \frac{(e^t)^k - 1}{e^t - 1}$ $S_n = \frac{a(r^n - 1)}{r - 1}$, r>1

$\because r = e^t > 1$

$= \frac{e^t(e^{tk} - 1)}{k(e^t - 1)}$ Ans

For mean and variance

Differentiate (1) both side w.r.t 't'

$M_x'(t) = \frac{1}{k} [e^t + e^{2t}(2) + e^{3t}(3) + \ldots\ldots\ldots\ldots + e^{kt}(k)]$

$= \frac{1}{k} [e^t + e^{2t}(2) + e^{3t}(3) + \ldots\ldots\ldots\ldots + k\, e^{kt}]$.....................(2)

Put t=0

$\mu_1' = M_x'(t)\,|_{t=0} = \frac{1}{k}[1+2+3 \ldots\ldots . k]$

$= \frac{1}{k} \frac{k(k+1)}{2} = \frac{(k+1)}{2}$

Mean $\mu = \mu_1' = \frac{k+1}{2}$

Var(x) = $\mu_2' - \mu_1'^2$ ------------------------(3)

Differentiate (2) both side w.r.t.'t'

Mx''(t) = $\frac{1}{k}$ $[e^t + 2^2e^{2t} + 3^2e^{3t} + \ldots\ldots\ldots\ldots + k^2e^{kt}]$

$\mu_2' = Mx''(t)\,|_{t=0} = \frac{1}{k}[1^2+2^2+3^2 \ldots\ldots\ldots\ldots + k^2]$

$= \frac{1}{k}\frac{k(k+1)(2k+1)}{6}$

$= \frac{(k+1)(2k+1)}{6}$

Use in (3)

Var(x) = $\frac{(k+1)(2k+1)}{6} - (\frac{k+1}{2})^2$

$= \frac{k+1}{2}[\frac{2k+1}{3} - \frac{k+1}{2}]$

$= \frac{k+1}{2}[\frac{4k+2-3k-3}{6}]$

$= \frac{k+1}{2}[\frac{k-1}{6}]$

$= \frac{k^2-1}{12} = \frac{k^2-1}{12}$

Example The first,second,third moments about point 2 are -1,13,20 respectively.

Find three moments

(i) about origin (ii) about mean

Solution: It is given first momentabout 2 = $E(x-2)^1] = -1$ -----(i)

2^{nd} moment about 2 = $E[(x-2)^2] = 13$ --------------------------(ii)

3rd moment about 2 = $E[(x-2)^3] = 20$ ------------------------------(iii)

From (ii), $E(x-2) = -1$

$E(x)-2=-1$

$E(x)= -1+2$ $\Rightarrow E(x)=1 = \mu_1' = 1$

From 2) , $E(x^2+2^2-4x] = 13$

$E(x^2)+ E(4) - E(4x) = 13$

$E(x^2)+ 4-4 E(x) = 13$

But $E(x)=1$

$E(x^2)+4-4(1) = 13$

$E(x^2) = 13$

i.e $\mu'_2 = E(x^2) = 13$

<u>From III)</u>

$E(x^3-2^3-6x^2+12x) = 20$

$E(x^3) - 8 - 6E(x^2)+12E(x) = 20.$ Use values

$E(x^3) - 8 - 6(13)+12(1) = 20$

$E(x^3) = 20+8+78-12$

$E(x^3) = 94$

$\mu_3' = E(x^3) = 94$

$\mu_1 = 0$

ii) we know

$\mu_2 = \mu_2' - \mu_1'^2$ (formula)

$\mu_2 = 13 - (1)^2 = 12$

$\boxed{\mu_2 = 12}$

$\mu_3 = \mu_3' - 3\mu_2\mu_1' + 2\mu_1'^3$

$= 94 - 3(13)(1) + 2(1)^3$

$\boxed{\mu_3 = 57}$

5.4.Examination Oriented Exercise/ Lesson End Exercise

Q.1. Find variance of no. of divisors in integer from 1,2,3.....1

Q.2. Find variance of no.of dots appearing on top of six dice w

Q.3. A gambler is playing with a dice. He is promised

(i) the sum

(ii) the product of number thrown

Q.4. Find mean & variance of random variable xthat takes val

each with probability $\frac{1}{n}$

Q.5. If 'x' has p.d.f. given as

$$f(x)=\begin{cases}\frac{x+1}{2}, -1 < x < 1\\ 0, elsewhere\end{cases}$$

then find mean of x

Q.6. Show that mean of x doesnot exist where x is r,v with p.d

$$f(x)=\begin{cases}\frac{1}{x^2}, -1 < x < 1\\ 0, \ otherwise\end{cases}$$

Q.7. If the variance x exist, show $E[(x^2)] \geq E[(x)]^2$

Q.7. If the variance x exist, show $E[(x^2)] \geq E[(x)]^2$

Q.8. Let the $1^{st}, 2^{nd}, 3^{rd}$ moments of distribution about point 7 are 3,11,15 respectively. Find mean and first three moments about origin and mean.

Q.9. Find mean and variance of x whose p.d.f. is

$$f(x)=\begin{cases} x, 0 \leq x \leq 1 \\ 2-x, 1 \leq x \leq 2 \\ 0, otherwise \end{cases}$$

5.5. SUGGESTED READINGS:

The students are advised to go through following references for details.

5.6. REFERENCES

1) Mathematical Statistics by J.N. Kapur & H.C. Saxena, S.Chand & Co.

2). Introduction to mathematical Statistics by P.G. Hoel, John Wiley & sons

3). Numerical Methods, Probability & statistics by Sandip Banerjee, Books & Allied Publications Ltd. Calcutta.

4). A text Book of Statistics, Numerical Analysis & Matrices by Sunil Gupta, Anupama Gupta, Narinder Kumar, Malhotra Brothers Pacca Danga, Jammu.

5.7. MODEL TEST PAPER

Q.1. Find first three moments about origin of random variable x whose moments generating function is given by

$M_x(t) = (1-t)^{-3}, t<1$

Q.2. Define probability and distribution functions of random variable x also list important properties of probability and distribution function.

Q.3. Let y= 3x-5 .Find mean and variance of Y such that mean X=4 and variance X=2.

Q.4. If x has mean 10and variance 6, then find $E(x^2+3x)$.Also find variance $(\frac{x}{2})$.

Q.5. Let $x_1, x_2, ...x_n$ are n variables having same mean , variance , m.g.f. Find mean, variance, m.g.f. of mean of these variables.

STATISTICS

BINOMIAL DISTRIBUTION

By: Dr. Sunil Gupta

6.1. INTRODUCTION: In this lesson one discrete distribution in the form of Binomial distribution is explained. Further the properties like mean, variance, moments, moment generating functions etc. of this distribution are discussed..

6.2. OBJECTIVES: Objective of studying this lesson is to give the idea of a distribution which is defined for discrete random variable.This distribution is applied when an experiment is performed more than one time.

6.3. BINOMIAL DISTRIBUTION(B.D)

6.3.1. Definition : Let x be a discrete r.v. Then x is said to be in Binomial Distribution(B.D) if its p.f. is given as

$f(x) = {}^nc_x p^x q^{n-x}$,x=0,1,2...............n, p+q=1, where p is probability of success and q is probability of failure.

6.3.2. Remark: The above function is probability function for B.D.

Proof To see above f(x) is p.f, we show $\sum_{x=-\infty}^{\infty} f(x)=1$

Consider $\sum_{x=-\infty}^{\infty} f(x)=\sum_{x=-\infty}^{n} n_{c_x} p^x q^{n-x}$

But

${}^nc_0 = {}^nc_n = 1$ ${}^nc_1 = n$

$= {}^nc_0 p^0 q^{n-0} + {}^nc_1 p^1 q^{n-1} + {}^nc_2 p^2 q^{n-2} + \ldots\ldots + {}^nc_n p^n q^{n-n}$

$= q^n + {}^nc_1 pq^{n-1} + {}^nc_2 p^2 q^{n-2} + \ldots\ldots + p^n$

$=(q+p)^n = (1)^n = 1.$

6.3.3. Article Find mean of Binomial Distribution (B.D)

Proof Let x be a discrete r.v with p.f.f(x) = $^nc_x p^x q^{n-x}$, x=0,1,2...............................n.

Mean (x) = E(x) = $\sum_{x=-\infty}^{\infty} xf(x)$

= $\sum_{x=0}^{n} x.\ ^nc_x p^x q^{n-x}$

=$0\times\ ^nc_0 p^0 q^{n-0} + 1\times\ ^nc_1 p^1 q^{n-1} + 2\times\ ^nc_2 p^2 q^{n-2} + 3\times\ ^nc_3 p^3 q^{n-3} +$ ....$+n\ ^nc_n p^n q^{n-n}$

= $npq^{n-1} + \frac{2.n(n-1)}{2.1} p^2 q^{n-2} + \frac{3.n(n-1)(n-2)}{3.2.1} p^3 q^{n-3} +$..$+ np^n$

=$np[q^{n-1} + (n-1)pq^{n-2} + \frac{(n-1)(n-2)}{2.1} p^2 q^{n-3} +$...................................$+p^{n-1}]$

=$np[(q+p)^{n-1}]$ But q+p =1

=$np\ (1)^{n-1}$ =np.

6.3.4. Article Find var of B.D

Proof Let x be in B.D with p.f.

f(x) =$^nc_x p^x q^{n-x}$, x=0,1,2,................n

var(x) = $\mu_2' - \mu_1'^2$

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

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

=$E(x^2)-(np)^2$ ....(1) ∵ E(x) = np

Now $E(x^2) = E[x(x-1)+x]$ #

= E[x(x-1)]+E(x)

=E(x(x-1))+np ---------------------------(2)

Also $E[x(x-1)] = \sum_{x=-\infty}^{\infty} x(x-1)f(x)$

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

$=0(0-1).\ ^nc_0p^0q^{n-0} + 1(1-1)^nc_1p^1q^{n-1} + 2(2-1)\ ^nc_2p^2q^{n-2} + 3(3-1)^nc_3\ p^3q^{n-3} + 4(4-1)\ ^nc_4\ p^4q^{n-4}$

$+ \text{------------------} + n(n-1)^nc_np^nq^{n-n}$

$= \cancel{2.1}\frac{n(n-1)}{\cancel{2.1}}p^2q^{n-2} + \frac{\cancel{3.2.}n(n-1)(n-2)}{\cancel{3.2.}1}p^3q^{n-3}$

$\cancel{4.3}\frac{n(n-1)(n-2)(n-3)}{\cancel{4.3.}2.1}\quad p^4q^{n-4} + \text{------------------} + n(n-1)p^n$

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

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

Use in (2)

$E(x^2) = n(n-1)p^2 + np$. Use in (1)

$Var(x) = n(n-1)p^2 + np - n^2p^2$

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

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

$=np(1-p) \quad = npq$

6.3.5. Article : Find m.g.f. of B.D

Proof Let x be a discrete random variable with probability function

$f(x) = {^nc_x}p^xq^{n-x}$, x=0,1,2..........n p+q=1

m.g.f of x = $M_x(t)=E[e^{tx}]$

$=\sum_{x=-\infty}^{\infty} e^{tx} f(x)$

$=\sum_{x=0}^{n} e^{tx}.{^nc_x}p^xq^{n-x} = \sum_{x=0}^{n} n_{c_x}(pe^t)^xq^{n-x}$

$={^nc_0}(pe^t)^0q^{n-0} + {^nc_1}(pe^t)^1q^{n-1} + {^nc_2}(pe^t)^2q^{n-2} + \ldots\ldots + {^nc_n}(pe^t)^nq^{n-n}$

$=q^n + {^nc_1}(pe^t)q^{n-1} + {^nc_2}(pe^t)^2q^{n-2} + \ldots\ldots + (pe^t)^n$

$=(q+pe^t)^n$

6.4.Recurrence relation of B.D.

6.4.1. State & prove recurrence relation of B.D.

Statement : If μ_r is r^{th} moment about the meanof B.D, then

μ_{r+1}= pq(nr μ_{r-1} + $\frac{d}{dp}\mu_r$)

Also find μ_2, μ_3,using μ_1=0, μ_0=1

Proof : $\mu_r = E[(x-\mu)^r] = \sum_{x=-\infty}^{\infty}(x-\mu)^r f(x)$

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

But in B.D μ=mean=np

$\mu_r = \sum_{x=0}^{n}(x-np)^r\, {}^nc_x p^x q^{n-x}$ -------(1)

$= \sum_{x=0}^{n}(x-np)^r\, {}^nc_x p^x (1-p)^{n-x}$

Diff b/s w.r.t 'p'

$\frac{d}{dp}(\mu_r) = \sum_{i=0}^{n} n_{c_x}\, [(x-np)^r p^x (n-x)(1-p)^{n-x-1}(-1) + (x-np)^r x\, p^{x-1}$

$(1-p)^{n-x} + r(x-np)^{r-1}(-n)p^x(1-p)^{n-x}]$

$=\sum_{x=0}^{n} n_{c_x}\, [(x-np)^r p^{x-1}(1-p)^{n-x-1}\{-p(n-x)+x(1-p)\} - nr(x-np)^{r-1} p^x (1-p)^{n-x}]$

$=\sum_{x=0}^{n} n_{c_x}\, [(x-np)^r p^{x-1} q^{n-x-1}\{-np+\cancel{xp}+x-\cancel{xp}\} - nr(x-np)^{r-1}p^x q^{n-x}]$

$=\sum_{x=0}^{n} n_{c_x}(x-np)^r p^{x-1} q^{n-x-1}(x-np) - nr(x-np)^{r-1}p^x q^{n-x}]$

$=\sum_{x=0}^{n}(x-np)^{r+1} n_{c_x} p^{x-1} q^{n-x-1} - \sum_{x=0}^{n} nr(x-np)^{r-1} n_{c_x} p^x q^{n-x}]$

Multiply b/s by pq

$pq\frac{d}{dp}\mu_r = pq\sum_{x=0}^{n}(x-np)^{r+1} n_{c_x} p^{x-1} q^{n-x-1} - pq\sum_{x=0}^{n} nr(x-np)^{r-1} n_{c_x} p^x q^{n-x}$

$=\sum_{x=0}^{n}(x-np)^{r+1} n_{c_x} p^x q^{n-x} - nrpq\sum_{x=0}^{n}(x-np)^{r-1} n_{c_x} p^x q^{n-x}$

$pq\frac{d}{dp}\mu_r = \mu_{r+1}$ - nrpq μ_{r-1}

$\Rightarrow \mu_{r+1}$ = nrpq μ_{r-1} +pq$\frac{d}{dp}\mu_r$

μ_{r+1} = pq(nrμ_{r-1} + $\frac{d}{dp}\mu_r$) .(*)

Put r=1 in (*)

μ_2 = pq (n.1μ_0 + $\frac{d}{dp}\mu_1$) But μ_1=0, μ_0 =1

μ_2=npq

put r=2, we get

μ_3 =pq (n.2. μ_1 +$\frac{d}{dp}\mu_2$) But μ_1 =0

$\therefore \mu_2$ =npq

put r = 2 in (*)

μ_3 = pq $\frac{d}{dp}(\mu_2)$

= pq $\frac{d}{dp} npq$

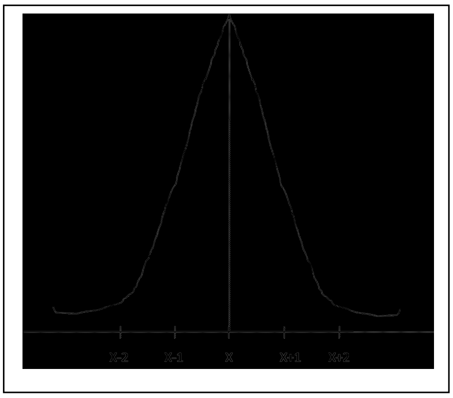
= npq $\frac{d}{dp} pq$

=npq$\frac{d}{dp}[p(1-p)]$ $\because$ p+ q =1

=npq$\frac{d}{dp}(p-p^2)$

μ_3 = npq(1-2p)

6.5.MODE OF A DISTRIBUTION



6.5.1. EXCERCISE:->Define mode and find expression for mode of B.D

Solution: Mode is the maximum value of a distribution

Let x be a discrete random variable which is in B.D

assume f(x) as mode of B.D where

$f(x)= {}^{n}c_{x}p^{x}q^{n-x}, p+q=1$

Since f(x) is assumed as mode, clearly we have

Case I

$f(x-1)\leq f(x)$

${}^{n}c_{x-1}p^{x-1}q^{n-(x-1)}\leq {}^{n}c_{x}p^{x}q^{n-x}$

$\frac{n!}{(n-(x-1))!(x-1)!}p^{x-1}q^{n-x+1}\leq \frac{n!}{(n-x)!x!}p^{x}q^{n-x}$

$\frac{n!}{(n-(x+1)!(x+1)!}p^{x+1}.q^{n-x-1}$

$\frac{n!p^{x-1}q^{n-x+1}}{(n-x+1)(n-x)!(x-1)!}\leq \frac{n!}{(n-x)!x!}p^{x}q^{n-x}$

$\frac{p^{x-1}q^{n-x}.q}{(n-x+1)!(n-x)!(x-1)!}\leq \frac{p^{x-1}.p.q^{n-x}}{(n-x)!x!}$

$\frac{q}{(n-x+1)(n-x)!(x-1)!}\leq \frac{p}{(n-x)!x(x-1)!}$

$\frac{q}{n-x+1}\leq \frac{p}{x}$

$qx \leq np-xp+p$

$qx+px \leq np+p$

$x(q+p) \leq p(n+1)$

$x \leq p(n+1)$------------(1)

Case II

$f(x)\geq f(x+1)$

${}^{n}c_{x}p^{x}q^{n-x} \geq {}^{n}c_{x+1}p^{x+1}q^{n-(x+1)}$

$\frac{n!}{(n-x)!x!}p^{x}.q^{n-x}\geq \frac{n!}{[n-(x+1)!(x+1)!}$

$\frac{p^{x}q^{n-x-1}.q}{(n-x)!x!}\geq \frac{p^{x}p.q^{n-x-1}}{(n-x-1)!(x+1)!}$

$\frac{q}{(n-x)(n-x-1)!x!}\geq \frac{p}{(n-x-1)!(x+1)!(x)!}$

$\frac{q}{n-x}\geq \frac{p}{x+1}$

$\Rightarrow qx+q \geq pn - xp$

$qx+xp\geq pn-q$

$\Rightarrow x(q+p) \geq pn-(1-p)$

$x(q+p) \geq pn-1+p$

$x \geq p(n+1)-1$

(n+1)p-1≤ x -------------------(2)

Combine (1) &(2)

(n+1)p-1 ≤ x ≤ (n+1)p

i.e if $x\epsilon$ [(n+1)p-1,(n+1)p], then given f(x) will be made.

.6. Standard Normal Variate (S.N.V)

ariable whose mean is zero and variance is 1 is called standard normal variate.

isdenoted by N(0,1)

.6.1 Exercise : Let x be in B.D.Show that z= $\frac{x-np}{\sqrt{npq}}$ is standard normal variate.

olution Let x is in B.D so variance x=npq

lean x = μ=np

lere z= $\frac{x-np}{\sqrt{npq}}$

/e show z is a S.N.V

or this, we show mean z=0 & variance z=1.

onsider

lean(z) = E[$\frac{x-np}{\sqrt{npq}}$]

$\frac{1}{\sqrt{npq}}$ [E(x)-E(np)]

$\frac{np-np}{\sqrt{npq}}$ ∵ E(x) =np

0

'ar z = var($\frac{x-np}{\sqrt{npq}}$)

var ($\frac{x}{\sqrt{npq}} - \frac{np}{\sqrt{npq}}$)

$(\frac{1}{\sqrt{npq}})^2$ var x∵ var (ax+b) = a^2 var x

$\frac{1}{\cancel{npq}}.\cancel{npq}=1$

z is S.N.V⇒z∈ N(0,1)

6.7. EXAMINATION ORIENTED EXERCISE/ LESSON END EXERCISE

Q.1. Define Binomial Distribution & show that it is a p.f.

Q.2. Find mean of B.D.

Q.3. From recurrence formulae, find m.g.f. of B.D.

Q.4. Find m.g.f. about mean of B.D.

Q.5. Find mean of n independent variables which are Binomially distributed having same mean.

Q.6. Find mean & variance of B.D.

Q.7. 3 electric bulbs are drawn at a random from a consignment of electric bulbs out of which 10% are defected. What is the probability of getting

i) No defective　　　　ii) one defective

iii) two defective bulbs　　　　iv) at least two defective

v) at most two defective

Q.8. A coin is tossed four times. What is the probability of

a) exactly two heads　　　　b) at least one head

c) more than 1 head

Q.9. Out of sixteen families with children each , how many of them are expected to have

a) no boys　　　　b) one boy

c) two boys

Also find no. of their families.

6.8. SUGGESTED READINGS:

The students are advised to go through following references for details.

6.9. REFERENCES

1). Mathematical Statistics by J.N. Kapur & H.C. Saxena, S.Chand & Co.

2). Introduction to mathematical Statistics by P.G. Hoel, John Wiley & sons

3). Numerical Methods, Probability & statistics by Sandip Banerjee, Books & Allied Publications Ltd. Calcutta.

4). A text Book of Statistics, Numerical Analysis &Matrices by Sunil Gupta,Anupama Gupta, Narinder Kumar, Malhotra Brothers Pacca Danga, Jammu.

6.10. MODEL TEST PAPER

Q.1. Find mean and variance of random variable x whose m.g.f is $(\frac{4}{5}+\frac{1}{5}e^t)^6$.

Q.2. Derive variance of B.D.

Q.3. Obtain recurrence relation of B.D.

Q.4. Define mode of a distribution & find expression for the mode of B.D.

Q.5. Define S.N.V. Let X be in B.D. Then find S.N.V. of B.D.

STATISTICS

POISSON DISTRIBUTION

By: Dr. Sunil Gupta

7.1. INTRODUCTION In this lesson another discrete distribution in the form of Poisson distribution is explained. Further the properties like mean, variance, moments, moment generating functions etc. of this distribution are discussed..

7.2. OBJECTIVES: Objective of studying this lesson is to give the idea of a distribution which is defined for discrete random variable. This distribution is applied when an experiment is performed more than one time.

7.3. POISSON DISTRIBUTION (P.D.)

7.3.1. Definition: A discrete r.v x is said to be P.D. if its probability function is

f(x) = $\frac{e^{-\lambda}x^x}{x!}$, x=0,1,2.........................., λ is any parameter

NOTE : The above function is p.f. for P.D

PROOF : To show above f(x) is p.f. we show

$\sum_{x=-\infty}^{\infty} f(x)$ =1

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

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

$=e^{-\lambda}[\frac{\lambda^0}{0!}+\frac{\lambda^1}{1!}+\frac{\lambda^2}{2!}+\cdots\ldots\ldots\ldots\ldots\ldots\ldots]$

$=\sout{e^{-\lambda}e^{\lambda}}$

(using1+ $\frac{\theta}{1!}+\frac{\theta^2}{2!}$+...= e^{θ})

=1

7.3.2.Article : Find mean of P.D

Proof : Let x be in p.d with probability function

f(x) = $\frac{e^{-\lambda}x^x}{x!}$, x=0,1,2..

Here E(x) = $\sum_{x=-\infty}^{\infty} xf(x)$

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

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

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

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

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

$=\lambda e^{-\lambda}[\frac{\lambda^0}{0!}+\frac{\lambda^1}{1!}+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}$........................]

$=\lambda e^{-\lambda}[1+\frac{\lambda^1}{1!}+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}$........................]

$=\lambda\sout{e^{-\lambda}e^{\lambda}}$

Mean x = λ

7.3.3. Article : Find m.g.f of P.D

Proof : let x be in P.D with probability function

m.g.f of x= $M_x(t) = E[e^{tx}]$

$= \sum_{x=-\infty}^{\infty} e^{tx} f(x)$

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

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

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

$=e^{-\lambda}[\frac{(\lambda e^t)^0}{0!} + \frac{(\lambda e^t)^1}{1!} + \frac{(\lambda e^t)^2}{2!} +....................]$

$=e^{-\lambda}[\ 1+ \frac{(\lambda e^t)^1}{1!} + \frac{(\lambda e^t)^2}{2!} +....................]$

$=e^{-\lambda}e^{\lambda e^t}$

$= e^{\lambda e^t - \lambda}$

m.g.f = $e^{\lambda(e^t-1)}$

7.3.4. Article Find variance of P.D

Proof let X be in P.Dwith p.f

f(x) = $\frac{e^{-\lambda}e^{xe^t}}{x!}$, x =0,1,2..................

var(x) = $\mu_2 = \mu_2' - \mu_1'^2$

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

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

Now

$E(x^2) = E[x(x-1)+x]$

$= E[x(x-1)+E(x)]$

$=E(x(x-1)] + \lambda$----------------(2)

Consider

$E[x(x-1)] = \sum_{x=-\infty}^{\infty} x(x-1)f(x)$

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

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

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

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

$= \lambda^2 e^{-\lambda} e^{\lambda}$

$= \lambda^2$. Use in (2) , we get $E[x(x-1)] = \lambda^2$. put in (1), we get var $x= \lambda$

7.4. Recurrence Relation for Poisson Distribution

7.4.1. Article: State and prove recurrence formulae of P.D

Solution: Statement If μ_r is r^{th} moment of P.D about mean, then

$\mu_{r+1} = \lambda(r\mu_{r-1} + \frac{d}{d\lambda}\mu_r)$

Findμ_2, μ_3 using $\mu_1=0$, $\mu_0=1$

Proof we know $\mu_r = E[(x-\mu)^r]$

$=\sum_{x=-\infty}^{\infty}(x-\mu)^r f(x)$

$=\sum_{x=0}^{\infty}(x-\mu)^r\frac{e^{-\lambda}\lambda^x}{x!}$ But in P.D $\mu= \lambda$

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

Difference b/s w.r.t. λ

$\frac{d}{d\lambda} \mu_r = \sum_{x=0}^{\infty} \frac{1}{x!} \frac{d}{d\lambda}[(x-\lambda)^r e^{-\lambda} \frac{d}{d\lambda}(\lambda)^x]$

$=\sum_{x=0}^{\infty} \frac{1}{x!} [(x-\lambda)^r e^{-\lambda} \frac{d}{d\lambda}(\lambda)^x + (x-\lambda)^r \lambda^x \frac{d}{d\lambda}(e^{-\lambda}) + e^{-\lambda} \lambda^x \frac{d}{d\lambda}(x-\lambda)^r]$

$\frac{d}{d\lambda}(U_r) = \sum_{x=0}^{\infty} \frac{1}{x!} [(x-\lambda)^r e^{-\lambda} x \lambda^{x-1} + (x-\lambda)^r \lambda^x (e^{-\lambda})(-1) + e^{-\lambda} \lambda^x r(x-\lambda)^{r-1}(-1)]$

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

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

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

Multiply both sides by λ

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

$= \mu_{r+1} - r \lambda \mu_{r-1}$

$\lambda \mu_{r-1} + \lambda \frac{d}{d\lambda} \mu_r = \mu_{r+1}$

$\mu_{r+1} = \lambda(r \mu_{r-1} + \frac{d}{d\lambda} \mu_r)$ -------(*)

Put r=1 in (*)

$\mu_2 = \lambda(1\mu_0 + \frac{d}{d\lambda} \mu_1)$

$= \lambda(1 + \frac{d}{d\lambda}(0))$

$\mu_2 = \lambda$

Put r=2 in (*)

$\mu_3 = \lambda (2 \mu_1 + \frac{d}{d\lambda} \mu_2)$

$= \lambda\ [2\ \mu_1 + \frac{d}{d\lambda}(\lambda))]$

$= \lambda[2\ \mu_1 + 1]$ But $\mu_1 = 0$

$= \lambda[2(0) + 1]$

$\mu_3 = \lambda$

7.4.2. Exercise Define S.N.V . if X is in P.D; then show that $\frac{x-\lambda}{\sqrt{\lambda}}$ is a standard normal variate.

i.eShow that $\frac{x-\lambda}{\sqrt{\lambda}} \in N(0,1)$

Solution A variable whose mean is zero and variance as 1 is called S.N.V.

As x is in P.D, so

Mean x = E(x) = λ

Var x = λ

Let $z = \frac{x-\lambda}{\sqrt{\lambda}}$

We show z is a S.N.V

For this, we show

Mean z=0& var z=1

Consider mean (z) = E(z) = $E[\frac{x-\lambda}{\sqrt{\lambda}}]$

$= \frac{1}{\sqrt{\lambda}}[E(x) - \lambda]$ But E(x) = λ

$= \frac{1}{\sqrt{\lambda}}(\lambda - \lambda) = 0$

Also var (z) = var $(\frac{x-\lambda}{\sqrt{\lambda}})$

$= var(\frac{x}{\sqrt{\lambda}} - \frac{\lambda}{\sqrt{\lambda}})$

$= (\frac{1}{\sqrt{\lambda}})^2$ var x

$= \frac{1}{\lambda} \lambda = 1$

$\therefore z$ is S.N.V

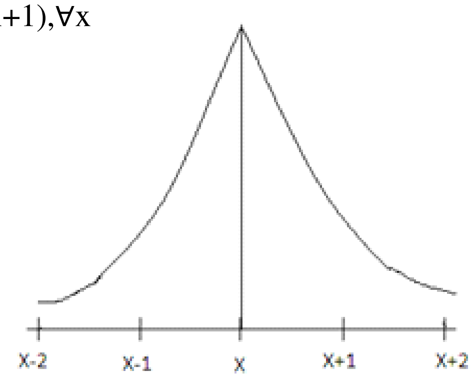
7.4.3. Exercise :- Define mode and find mode of P.D

Solution: Mode is the maximum value of the distribution . Let x be in P.D having probability function as

$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$, x = 0,1,2,........

Assume f(x) to be mode of P.D, which means

$f(x-1) \leq f(x) \geq f(x+1), \forall x$



$f(x-1) \leq f(x)$

$\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \leq \frac{e^{-\lambda}\lambda^x}{(x)!}$

$\frac{\lambda^{x-1}}{(x-1)!} \leq \frac{\lambda^x}{(x)!}$

$\frac{\lambda^{x-1}}{(x-1)!} \leq \frac{\lambda^{x-1}\lambda}{x(x-1)!}$

$1 \leq \frac{\lambda}{x}$

$x \leq \lambda$ -----------------(1)

$f(x) \geq f(x+1)$

$\frac{e^{-\lambda}\lambda^x}{(x)!} \geq \frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!}$

$\frac{\lambda^x}{(x)!} \geq \frac{\lambda^{x+1}}{(x+1)!}$

$\frac{\lambda^x}{(x)!} \geq \frac{\lambda^x\lambda}{(x+1)x!}$

$1 \geq \frac{\lambda}{x+1}$

$x+1 \geq \lambda$

$x \geq \lambda-1$

or $\lambda-1 \leq x$-------------------(2)

Combine (1) &(2)

$(\lambda -1) \leq x \leq \lambda$

$\Rightarrow x \in [\lambda-1, \lambda]$

$\therefore$ if $x \in [\lambda-1, \lambda^{-}]$, then required f(x) will be mode.

7.5. EXAMINATION ORIENTED EXERCISE/ LESSON END EXERCISE

Q.2. Derive variance of P.D.

Q.3. Obtain recurrence relation of P.D.

Q.4. Define mode of a distribution & find expression for the mode of P.D.

Q.5. Define S.N.V. Let X be in P.D. Then find S.N.V. of P.D.

Q.6. A typist on an average makes three errors per page. What is the probability of her typing a page

a) with no error b) with at least two errors.

Q.7. In P.D ,P(x=1)=P(x=2). Find P(x=4).

7.6. SUGGESTED READINGS:

The students are advised to go through following references for details.

7.7. REFERENCES

1). Mathematical Statistics by J.N. Kapur & H.C. Saxena, S.Chand & Co.

2). Introduction to mathematical Statistics by P.G. Hoel, John Wiley & sons

3). Numerical Methods, Probability & statistics by Sandip Banerjee, Books & Allied Publications Ltd. Calcutta.

4). A text Book of Statistics, Numerical Analysis & Matrices by Sunil Gupta,Anupama Gupta, Narinder Kumar, Malhotra Brothers Pacca Danga, Jammu.

7.8. MODEL TEST PAPER

Q.1. Define Poisson Distribution & show that it is a p.f.

Q.2. Find mean of P.D.

Q.3. From recurrence formulae, find m.g.f. of P.D.

Q.4. Find m.g.f. about mean of P.D.

Q.5. Find mean of n independent variables which are in Poisson variates having same mean.

STATISTICS

SIMPLE PROBLEMS BASED ON THESE DISTRIBUTIONS

By: Dr. Sunil Gupta

8.1. INTRODUCTION: In this lesson the practice questions of both lessons 06 & 07 are discussed with their solutions.

8.2 OBJECTIVES: The Objective of studying this lesson is to revise the concepts studied in both lesson 06 & 07 in the form of exercises. functions. awareness among the students about the concept of.

8.3. EXERCISES BASED ON BINOMIAL DISTRIBUTION

Q.1. Find B.D forwhich mean is 7 and variance is $\frac{28}{5}$.

Q.2. 3 electric bulbs are drawn at a random from a consignment of electric bulbs out of which 10% are defected. What is the probability of getting

i) No defective ii) one defective

iii) two defective bulbs iv) atleast two defective

v) at most two defective

Q.3. A coin is tossed four times.What is the probability of

a) exactly two heads b) atleast one head

c) more than 1 head

Q.4. Out of sixteen families with children each , how many of them are expected tohave

a) no boys b) one boy

c) two boys

Also find no.of their families.

Q.5. Find mean and variance of random variable x whose m.g.f is $(\frac{4}{5}+\frac{1}{5}e^t)^6$

Q.6. If X is in binomial variate , n=6, find p if 9 P(x=−3) = P(x=2)

Q.7. Eight coins are thrown simultaneously . Find probability of getting atleast six heads.

Q.8. Mean of B.D is 2, variance is $\frac{4}{3}$ .What is the probability of getting

a) At least two successes b) less than two successes

Q.9. If the probability of defective bolt is $\frac{1}{10}$, find

1) mean 2) variance

3) m.g.f

for the distribution of defective bolt in a total of 400.

8.4. EXERCISES BASED ON P.D

Q.1. Find mean and variance of P.D whose parameter is m.

Q.2. In P.D,P(x=1)=P(x=2). Find P(x=4)

Q.3. Calculate the probability of exactly two accidents at a busy intersection on a given day, if on the average there are 0.5 accidents per day. Also find probability of no accidents in a day.

Q.4. A typist on an average makes three errors per page. What is the probability of her typing a page

a) with no error b) with at least two errors.

Solutions of 8.3

Solution 1

$f(x) = {}^{n}c_{x}p^{x}q^{n-x}$ ----------------(1) , x=0,1,2........n

it is given mean =7 = np ----------------(2)

and variance = $\frac{28}{5}$ = npq --------------(3)

divide (3) by (2)

$\frac{npa}{np} = \frac{28/5}{7}$

$q = \frac{4}{5}$

P=1-q

$=1-\frac{4}{5} = \frac{1}{5}$

Use value of p in (2)

$n(\frac{1}{5}) = 7$

n=35

$f(x) = {}^{35}c_{x}(\frac{1}{5})^{x}(\frac{4}{5})^{35-x}$, x=0,1,2.......35.

Solution 2N=3probability of defective bulbs = 10%

$p = \frac{10}{100} = \frac{1}{10}$

$q = 1-p = 1-\frac{1}{10} = \frac{9}{10}$

1) P[no defective bulb] =P[X=0]

$= {}^{n}c_0 p^0 q^{n-0}$

$= q^n$

$= (\frac{9}{10})^3 = \frac{729}{1000} = 729$

(ii) P[one defective] = P[x=1]

$= {}^{3}c_1 p^1 q^{n-1}$

$= {}^{3}c_1 (\frac{1}{10})^1 (\frac{9}{10})^{3-1}$

$= {}^{3}c_1 \frac{1}{10} . (\frac{9}{10})^2$

$= \frac{3!}{2!} \frac{1}{10} \frac{81}{100} = \frac{3.2!}{2!} \frac{81}{1000}$

$= \frac{243}{1000} = .243$

(iii) P[two defective] =P[x=2]

$= {}^{n}c_2 p^2 q^{n-2}$

$= {}^{3}c_2 p^2 q^{3-2}$

$= \frac{3!}{2!} (\frac{1}{10})^2 (\frac{9}{10})$

$= \frac{3.2!}{2!} \frac{1}{100} . \frac{9}{10}$

$= \frac{27}{1000} = 0.027$

(iv) P[at least 2] = P[x≥2]

=P[x=2]+P[x=3]

=0.027 +0.1

=0.127

(v) P[at most two defective]

$=P[x \leq 2] = P[x=0,1,2]$

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

$=0.729 +0.243+0.027$

Solution 3 n=4Probability of head = $\frac{1}{2}$ i.e p = $\frac{1}{2}$

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

P[exactly two heads] $=P[x=2] = {}^{n}c_2 p^2 q^{n-2}$ n=4

$= {}^{4}c_2 p^2 q^2$

$= \frac{4.3}{2.1}.(\frac{1}{2})^2(\frac{1}{2})^2$

$=\frac{4.3}{2}.\frac{1}{4}.\frac{1}{4}$

$= \frac{3}{8}$

ii) P[at least one head] $=p[x\geq 1]$

$=P[x=1,2,3,4]$

$=1-P[x=0]$

$=1- {}^{n}c_0 p^0 q^{n-0}$

$=1- {}^{4}c_0 p^0 q^{4-0}$

$=1- q^4$ $= 1- (\frac{1}{2})^4 = \frac{15}{16}$

Solution 7 n=8

Probability of head $p=\frac{1}{2}$

Probability of tail $q=1-p=1- \frac{1}{2}= \frac{1}{2}$

probability of at least six head

$P[x \geq 6]$

$= P[x=6] + P[x=7] + P[x=8]$

$= {}^{n}c_6 p^6 q^{n-6} + {}^{n}c_7 p^7 q^{n-7} + {}^{n}c_8 p^8 q^{n-8}$, n=0,1,2............

$= {}^{8}c_6 (\frac{1}{2})^6 (\frac{1}{2})^2 + {}^{8}c_7 (\frac{1}{2})^7 (\frac{1}{2})^1 + {}^{8}c_8 p^8 q^0$

$= \frac{8.7}{2} \frac{1}{64} . \frac{1}{4} + \cancel{8} . \frac{1}{\cancel{128}} . \frac{1}{2} + 1(\frac{1}{2})^8$

$= \frac{7}{64} + \frac{1}{32} + \frac{1}{256}$

$= \frac{28+8+1}{256} = \frac{37}{256}$

Number of families N=16

No.of children n=4

Probability of boy = $\frac{1}{2}$

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

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

(i) P[no boy] =P[x=0]

$= {}^{n}c_0 p^0 q^{n-0}$

$= q^n = (\frac{1}{2})^4 = \frac{1}{16}$

No of families with no boy

$= N \times P$

$= 16 \times \frac{1}{16} = 1$

(ii) P[one boy] = P[x=1]

$= {}^{n}c_1 p^1 q^{n-1}$

$= {}^{4}c_1 p^1 q^3$

$= \frac{4!}{3!1!} (\frac{1}{2})^1 (\frac{1}{2})^3$

$= \frac{4.3!}{3!} \frac{1}{16} = \frac{1}{4}$

No of families with one boy

= N×P

$= 16 \times \frac{1}{4} = 4$

(iii) P[two boys] = P[x=2]

$= {}^{n}c_2 p^2 q^{n-2}$

$= {}^{4}c_2 p^2 q^2$

$= \frac{4.3!}{2} . \frac{1}{16}4 = \frac{3}{8}$

No of families with two boys = N×P

$= {}^{2}16 \times \frac{3}{8} = 6$

<u>Solution 5</u> let X be in B.D whose m.g.f is

$M_x(t) = (\frac{4}{5} + \frac{1}{5} e^t)^6$

$=(q+p\, e^t)^n$

$\Rightarrow q = \frac{4}{5} \quad p = \frac{1}{5}, n=6$

$\therefore$ mean =np = $6 \times \frac{1}{5} = \frac{6}{5}$

Variance =npq = $6 \times \frac{1}{5} \times \frac{4}{5} = \frac{24}{25}$

Solution 6 n=6 p=?

9P[x=4]=P[x=2]

Let X be in B.D with p.f

P[x=2] =f(x) = $^{n}c_{x}p^{x}q^{n-x}$,x=0,1,2...............n

Here n=6

& 9P[x=4] =P[x=2]

$9\,{}^{n}c_{4}p^{4}q^{n-4} = {}^{n}c_{2}p^{2}q^{n-2}$

$9\times{}^{6}c_{4}p^{4}q^{2} = {}^{6}c_{2}p^{2}q^{4}$

$9\,{}^{6}c_{4}p^{4}(1-p)^{2} = {}^{6}c_{2}p^{2}(1-p)^{4}$

$9[\frac{6.5.4.3}{4.3.2.1}\,p^{4}\,(1-p)^{2}] = \frac{6.5}{.2.1}\,p^{2}\,(1-p)^{4}$

$135\,p^{4}(1-p)^{2} = 15\,p^{2}\,(1-p)^{4}$

$9p^{2} = (1-p)^{2}$

$9p^{2} = 1+p^{2}-2p$

$8p^{2}+2p-1 = 0$

$p= \frac{-2\pm\sqrt{4+32}}{16}$

$=\frac{-2\pm 6}{16} = \frac{-2\pm 6}{16}, \frac{-2-6}{16}$

$= \frac{4}{16}, \frac{-8}{16}$

$=\frac{1}{4}, -\frac{1}{2}$

But p(being prob) $\neq$- $-\frac{1}{2}$

$\therefore p= \frac{1}{4}$

Solution 8 mean (B.D) =np =2------------------(1)

Variance = npq =$\frac{4}{3}$ -------------------------------(2)

Divide (2) by (1)

$\frac{\not{np}\,q}{\not{np}} = \frac{4/3}{2}$

q=$\frac{2}{3}$

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

p= $\frac{1}{3}$

use in (1)

np=2

n($\frac{1}{3}$) =2

n=6

1) P(exactly two successes)

=P[x=2] =$^{n}c_2p^2q^{n-2}$

=$^{6}c_2p^2q^4$

= $^{6}c_2(\frac{1}{3})^2\ (\frac{2}{3})^4$

=$\frac{{}^{2}\not{6}.5}{\not{2}.1}\frac{1}{\not{9}^{3}}.\frac{\not{16}\,8}{81}$

= $\frac{80}{243}$

(ii) P[less than two successes]

=P[x<2] =P[x=0,1]

=P[x=0]+P[x=1]

Solutions of 8.4.

Solution1: Take p.f.as f(x)

f(x)=$\frac{e^{-m}m^x}{x!}$, x=0,1,2..........

Solution 2 : in P.D

F(x)=P[x=x] = $\frac{e^{-\lambda}\lambda^x}{x!}$, x=0,1,2.............

It is given

P[x=1] =P[x=2]

$\frac{e^{-\lambda}\lambda^1}{1!} = \frac{e^{-\lambda}\lambda^2}{2!}$

$\lambda^1 = \frac{\lambda^2}{2}$

$2\lambda = \lambda^2$

λ^2-2λ =0

$\lambda(\lambda - 2)$=0

λ=0 not possible

λ=2

P[x=4] = $\frac{e^{-\lambda}\lambda^4}{4!}$

=$\frac{e^{-2}(2)^4}{4!}$

= $\frac{e^{-2}\not{16}\,2}{\not{24}\,3}$

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

Solution 3. Average accident = 0.5

Mean(λ) =0.5

f(x) = p[x=x] = $\frac{e^{-\lambda}\lambda^x}{x!}$

i) P[exactly two acc] =P[x=2]

= $\frac{e^{-\lambda}\lambda^2}{2!}$

=$\frac{e^{-(0.5)}(0.5)^2}{2!}$

ii) P[no. accident] = P[x=0]

=$\frac{e^{-\lambda}\lambda^0}{0!}$ =$e^{-\lambda}$ = $e^{-(0.5)}$ Answer

Solution4: Average error =3

i.e λ=3

f(x) = P[x=x] =$\frac{e^{-\lambda}\lambda^x}{x!}$

i) P[with no error] = P[x=0]

=$\frac{e^{-\lambda}\lambda^0}{0!}$

=$e^{-\lambda}$ =e^{-3}

ii) P[at least two errors]=P[x≥2]

=1-P[x=0,1]

=1-(P[x=0]+P[x=1])

=1-[$\frac{e^{-\lambda}\lambda^0}{0!}$ + $\frac{e^{-\lambda}\lambda^1}{1!}$]

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

=1-(e^{-3}+3e^{-3})

8.5. SUGGESTED READINGS/REFERENCES

1). Mathematical Statistics by J.N. Kapur & H.C. Saxena, S.Chand & Co.

2). Introduction to mathematical Statistics by P.G. Hoel, John Wiley & sons

3). Numerical Methods, Probability & statistics by Sandip Banerjee, Books & Allied Publications Ltd. Calcutta.

4). A text Book of Statistics, Numerical Analysis & Matrices by Sunil Gupta, Anupama Gupta, Narinder Kumar, Malhotra Brothers Pacca Danga, Jammu.

STATISTICS

NORMAL DISTRIBUTIONS

By: Dr. Sunil Gupta

9.1. INTRODUCTION: In these lessons Normal Distribution in the form of continuous random variable is explained. Further the properties like mean, variance, moments, moment generating functions etc. of this distribution are discussed..

9.2. OBJECTIVES: Objective of studying this lesson is to give the idea of a distribution which is defined for continuous random variable. This distribution is applied when an experiment is performed more than one time.

9.3. NORMAL DISTRIBUTION

9.3.1. Definition: A continuous random variable X is said to be in normally distributed if its p.d.f is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \quad -\infty<x<\infty, \mu>0, \sigma>0$$

NOTE $\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx = 1$

9.3.2. Article Find mean of N.D

Proof let x be in N.D whose p.d.f. is

f(x) =$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$, , -∞<x<∞

Mean=E(x)= $\int_{-\infty}^{\infty} xf(x)dx$

=$\int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}}(\frac{x-\mu}{\sigma})^2$ dx--------------------(1)

Put $\frac{x-\mu}{\sigma}$ = z

x- $\mu = \sigma$z

x= $\mu + \sigma$z

Differentiate both side w.r.t.z

$\frac{dx}{dz}$ = 0+σ (1)

$\frac{dx}{dz} = \sigma$

$\frac{dx}{dz} = \sigma$dz

Use in (1)

Mean=$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(\mu + \sigma z).e^{-\frac{1}{2}}$ z^2~~σ~~dz

=$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\mu + \sigma z)e^{\frac{-z^2}{2}}$dz

=$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\mu e^{\frac{-z^2}{2}} + \sigma ze^{\frac{-z^2}{2}})$dz

$=\frac{1}{\sqrt{2\pi}}\ [\mu \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}} dz + \sigma \int_{-\infty}^{\infty} z e^{\frac{-z^2}{2}} dz]$

$= \mu.1 + \frac{\sigma}{\sqrt{2\pi}} \times 0$ (why 0, proof see below in note)

$= \mu$

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

NOTE:- (1) if f(x) =-f(x) then f(x) is said to be odd function

f(z)=z$e^{\frac{-z^2}{2}}$

f(-z) =(-z)$e^{\frac{-(-z)^2}{2}}$

= -[z$e^{\frac{-z^2}{2}}$]

=-f(z)

Hence f(-z) is said to be odd function

(2) when interval of limit is same with opposite signs and functions is odd

then its value is always zero.

9.3.3.Article Find variance of N.D.

Proof let xbea continuous r.v.with p.d.f.

f(x) = $\frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}$, -∞<x<∞

var(x) =μ_2 = E[(x- μ)^2]

=$\int_{-\infty}^{\infty} (x - \ \mu)^2 . \frac{1}{\sqrt{2\pi}\,\sigma} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2} dx$ ----------------(1)

Put $\frac{x-\mu}{\sigma}$=z ⇒x- $\mu = \sigma$ z ⇒ x= $\mu + \sigma$z

Differentiate both side w.r.t.z

$\frac{dx}{dz}$=σ.(1)

dx= σdz

if x=∞ , z=$\frac{\infty-\mu}{\sigma}$=∞

if x=$-\infty$, z =$\frac{-\infty-\mu}{\sigma}$=$-\infty$

use in (1)

var(x) = $\frac{1}{\sqrt{2\pi}\,\not\sigma}\int_{-\infty}^{\infty}(\sigma z)^2\, e^{\frac{-1}{2}z^2}\,(\not\sigma\, dz)$

$=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(z)^2\, e^{-\frac{z^2}{2}}\, dz$

$=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z(\, z\, e^{-\frac{-z^2}{2}})\, dz$

I II

$=\frac{\sigma^2}{\sqrt{2\pi}}[|z\int ze^{-\frac{-z^2}{2}}|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{d}{dz}(z)\int ze^{-\frac{-z^2}{2}}\, dz)dz]$

$= \frac{\sigma^2}{\sqrt{2\pi}}[z(-e^{-\frac{-z^2}{2}})|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} 1(-e^{-\frac{-z^2}{2}})dz]$

$=\frac{\sigma^2}{\sqrt{2\pi}}[-(0-0)+\int_{-\infty}^{\infty} e^{-\frac{-z^2}{2}}\, dz \because e^{-\infty} = \frac{1}{e^{\infty}} = \frac{1}{\infty} =0$

$=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}\, dz$

$=\sigma^2.\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{-z^2}{2}}\, dz$

$=\sigma^2 \times 1$ $\qquad (\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}\, dz = 1$

Var(x) = $\mu_2 = \sigma^2$

9.3.4 .ArticleFind m.g.f of N.D.

Proof let x be a continuous r.v. with p.d.f. f(x)= $\frac{1}{\sqrt{2\pi}\,\sigma} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}$, -∞<x<∞

m.g.f of x = $M_x(t)$ = E[e^{tx}]

$=\int_{-\infty}^{\infty} e^{tx}$ f(x)dx

$=\int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}$ dx

$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{tx}$ - $(\frac{x-\mu}{2\sigma^2})^2$ dx

$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[(x-\mu)^2-2\sigma^2 tx]}$dx ------------------(1)

Now$(x-\mu)^2-2\sigma^2 tx$

$=(x-(\mu+\sigma^2 t))^2 - 2\sigma^2\mu t - \sigma^4 t^2$

Use in (1)

m.g.f of x

$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}[x-(\mu+t\sigma^2)^2]} - 2\sigma^2\mu t - \sigma^4 t^2$ dx

$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[\{x-(\mu+t\sigma)\}^2} e^{-\frac{1}{2\sigma^2}(-2\sigma^2\mu t-\sigma^2 t^2\}}$ dx

$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[\frac{x-(\mu+t\sigma^2}{\sigma}]^2}. e^{\mu t+\frac{\sigma^2 t^2}{2}}$ dx

$=e^{\mu t+\frac{\sigma^2 t^2}{2}}$

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

i.e. $\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[(x-\mu)^2 dx}$ =1)

9.3.5.ArticleFind expression for even moments about mean of N.D.

or

If μ_{2n} is even moment of N.D about mean, then

μ_{2n} = (2n-1)σ^2 μ_{2n-2} , for n=1,2,3..........................

Proof let x be a continuous random variable with p.d.f.

f(x) = $\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}$, -∞<x<∞

μ_{2n} = E[(x- μ)^{2n}] = $\int_{-\infty}^{\infty}(x-\mu)^{2n}$f(x) dx

=$\int_{-\infty}^{\infty}(x-\mu)^{2n} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}$ dx

=$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-\mu)^{2n} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}$dx ....(1)

Put $\frac{x-\mu}{\sigma} = z$ => x- μ =σz

Differentiate both side w.r.t.z

$\frac{dx}{dz} = \sigma$

=> dx = σdz

If x=∞, then z=∞

If x=-∞,then z=-∞

Use in (1)

$\mu_{2n} = \frac{1}{\sqrt{2\pi}\not\sigma}\int_{-\infty}^{\infty}(\sigma z)^{2n} e^{\frac{-1}{2}z^2} \not\sigma$ dz

= $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma z)^{2n} e^{\frac{-z^2}{2}}$ dz

=$\frac{\sigma^{2n}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^{2n} e^{\frac{-z^2}{2}}$ dz --------------(2)

=$\frac{\sigma^{2n}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^{2n-1} (ze^{\frac{-z^2}{2}})dz$

I II

=$\frac{\sigma^{2n}}{\sqrt{2\pi}}$ [$z^{2n-1}(-e^{\frac{-z^2}{2}})_{-\infty}^{\infty}$ - $\int_{-\infty}^{\infty}(2n-1)z^{2n-2}(-e^{\frac{-z^2}{2}})$dz]

$=\frac{\sigma^{2n}}{\sqrt{2\pi}}$ [$-z^{2n-1}(e^{-\infty}$ - e^{∞}) +(2n-1)$\int_{-\infty}^{\infty} z^{2n-2}\, e^{\frac{-z^2}{2}}$dz)

But $e^{-\infty} = \frac{1}{e^{\infty}} = \frac{1}{\infty} = 0$

$=\frac{\sigma^{2n}}{\sqrt{2\pi}}$ [$-z^{2n-1}$(0-0)+(2n-1)$\int_{-\infty}^{\infty} z^{2n-2}\, e^{\frac{-z^2}{2}}$dz]

$=\frac{\sigma^{2n}}{\sqrt{2\pi}}$(2n-1)$\int_{-\infty}^{\infty} z^{2n-2}\, e^{\frac{-z^2}{2}}$dz

=(2n-1)$\sigma^2\, \mu_{2n-2}$ using (2)

9.3.6.Article Find expression for odd moments about mean of N.D.

or

Show that all odd moments about mean of N.D are zero

Proof let x be a continuous random variable with p.d.f.

f(x)=$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$,-∞<x<∞

odd moments about mean

μ_{2n-1} = E[(x- $\mu)^{2n-1}$]

$=\int_{-\infty}^{\infty}(x-\mu)^{2n-1}\ f(x)dx$

$\mu_{2n-1} =\int_{-\infty}^{\infty}(x-\mu)^{2n-1}\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}$dx

$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-\mu)^{2n-1}\, e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}$ dx ---------------------(1)

Put $\frac{x-\mu}{\sigma}$ = z

$x-\mu = \sigma$z

Differentiate b/s w.r.t.z

$\frac{dx}{dz} = \sigma(1)$

dx = σdz

if x=∞, then z=∞

x=−∞, then z=−∞

$\mu_{2n-1} = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(\sigma z)^{2n-1}\, e^{\frac{-z^2}{2}}\, \sigma dz$

$=\frac{\sigma^{2n-1}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^{2n-1}\, e^{\frac{-z^2}{2}} dz$

$= \frac{\sigma^{2n-1}}{\sqrt{2\pi}} \times 0 = 0$

$\mu_{2n-1} = 0$ [$\because$ integrad is odd f(x) of z, so $\int_{-a}^{a} odd\ f(x) = 0$]

9.3.7.ArticleFind mode of N.D.

Proof Let x be in continuous random variable with p.d.f

$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}$, $-\infty < x < \infty$

Assume f(x) to be mode

let $y = f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}$ -------(1)

Differentiate both side w.r.t.x

$\frac{dy}{dx} = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2} \frac{d}{dx}[-\frac{1}{2\sigma^2}(x-\mu)^2]$

$= \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2} [\frac{-1}{2\sigma^2}.\, 2(x-\mu).1]$

$= \frac{-(x-\mu)e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}}{\sqrt{2\pi}\sigma^3}$

Differentiate again w.r.t.x

$\frac{d^2y}{dx^2} = \frac{1}{\sqrt{2\pi}\sigma^3}\frac{d}{dx}[-(x-\mu)e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}]$

$= -\frac{1}{\sqrt{2\pi}\sigma^3}[(x-\mu)e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \frac{d}{dx}[\frac{-1}{2\sigma^2}(x-\mu)^2 + e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.1]$

$= -\frac{1}{\sqrt{2\pi}\sigma^3}[(x-\mu)e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.\, (\frac{-1}{2\sigma^2}.2(x-\mu).1) + e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}]$------------------(2)

For mode (i.e maxima)

$\frac{dy}{dx} = 0$

$-\frac{1}{\sqrt{2\pi}\sigma^3}(x-\mu)e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}=0$

$(x-\mu)e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}=0$

But $e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \neq 0$

$\therefore x-\mu=0$

$x=\mu$

use in (2)

$\frac{d^2y}{dx^2} = -\frac{1}{\sqrt{2\pi}\sigma^3}(0+1)$

<0

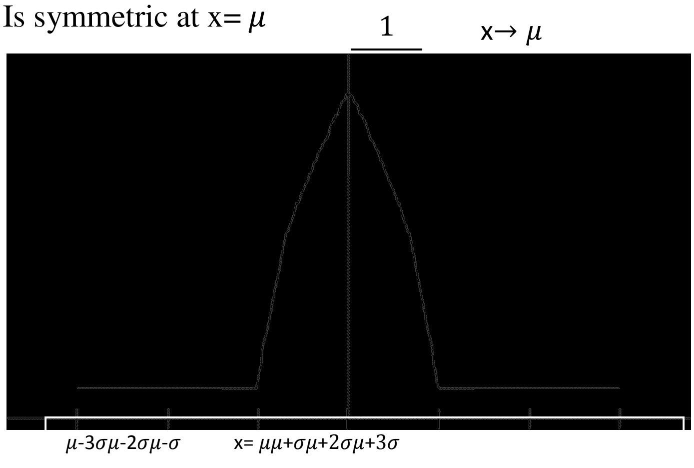
$\therefore x=\mu$ is mode of N.D

9.3.8.Article:**List important properties of N.D**.

Proof: (1) the normal curve

$Y= f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty<x<\infty$

Is symmetric at $x=\mu$



(ii) the shape of normal curve is like bell

(iii)the maximum value of normal curve is at $x=\mu$ and value is $\frac{1}{\sqrt{2\pi}\,\sigma}$

(iv) Allodd moments of N.D about the mean are zero i.e

$\mu_1=0=\mu_3=\mu_5$.......................................

or$\mu_{2n-1} = 0, \forall n=1,2,3$...

(v) even moments of N.D about mean is given by

$\mu_{2n} = (2n-1)\,\sigma^2\,\mu_{2n-2}$, n=1,2,3.

(vi) The mean deviation about mean of N.D is $\frac{4}{5}\sigma$ (approx)

(vii) 68.26 % of area falls with in one standard deviation (σ)

i.e

$P[\mu - \sigma < x < \mu + \sigma] = .6826$

(viii) 95.46% of area falls with in two S.D i.e $P[\mu - 2\sigma < x < \mu + 2\sigma] = .4773$

(ix) 99.86 % of area falls wthin 3 standard deviation

$P[\mu - 3\sigma < x < \mu + 3\sigma] = .4993$

Proof of property vii)

$= P[\mu - \sigma < x < \mu + \sigma] = .6826$

L.H.S $P[\mu - \sigma < x < \mu + \sigma]$ (subtraction by μ)

$= P[\mu - \sigma - \mu < x - \mu < \mu + \sigma - \mu\ ]$

$= P[-\sigma < x - \mu < \sigma]$

$= P[-1 < \frac{x-\mu}{\sigma} < 1]$

$= P[-1 < z < 1]$, $z = \frac{x-\mu}{\sigma}$

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

$= 2P[z<1]$ $\because$ curve is symmetric

$= 2P[z=1]$

$= 2(.3413)$ from table

$= .6826$

9.3.9.Question Let x be in N.D Then show $\frac{x-\mu}{\sigma} \epsilon$ N(0,1)

or

Show $\frac{x-\mu}{\sigma}$ is a S.N.V

Solution let $z=\frac{x-\mu}{\sigma}$

As x is in N.D, So

Mean(x)= E(x) = μ

Var(x)=σ^2

To show z is S.N.V, we show its mean =0 & var =1

Mean(z)= E(z) = $E[\frac{x-\mu}{\sigma}]$

$=\frac{1}{\sigma}E(x-\mu]$

$=\frac{1}{\sigma}[E(x)-\mu]$

$=\frac{1}{\sigma}[\mu-\mu]=0$

Var (z) = var $(\frac{x-\mu}{\sigma})$

$=(\frac{1}{\sigma})^2$var (x)

$=\frac{1}{\sigma^2}\sigma^2 = 1$

∴ z is S.N.V

9.4.Article Find recurrence relation of N.D.

Statement If μ_{2n} is even moment of N.D about mean, then

$\mu_{2n+2} = \sigma^2 \mu_{2n} + \sigma^3 \frac{d}{d\sigma}\mu_{2n}$

Proof let x be a continuous random variable with p.d.f

$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, $-\infty < x < \infty$

$\mu_{2n} = E[(x-\mu)^{2n}] = \int_{-\infty}^{\infty}(x-\mu)^{2n} f(x)\,dx$

$=\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}(x-\mu)^{2n}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$ ----------------------------------(1)

Differentiate w.r.t. σ

$$\frac{d}{d\sigma}\mu_{2n} = \frac{d}{d\sigma}\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}(x-\mu)^{2n}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-\mu)^{2n}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\frac{d}{d\sigma}\left(\frac{1}{\sigma}\right) dx + \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}(x-\mu)^{2n}\frac{d}{d\sigma}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-\mu)^{2n}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\left(\frac{-1}{\sigma^2}\right)dx + \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}(x-\mu)^{2n}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\frac{d}{d\sigma}\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 dx$$

$$=\frac{-1}{\sigma^2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-\mu)^{2n}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}(x-\mu)^{2n}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\cdot -\frac{1}{2}(x-\mu)^2\frac{d}{d\sigma}\left(\frac{-1}{\sigma^2}\right) dx$$

$$=-\frac{1}{\sigma}\cdot\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-\mu)^{2n}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}(x-\mu)^{2n}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\ -\frac{1}{2}(x-\mu)^2\left(\frac{-2}{\sigma^3}\right)dx$$

$$=-\frac{1}{\sigma}\cdot\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}(x-\mu)^{2n}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \frac{1}{\sigma^3}\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}(x-\mu)^{2n+2}\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Use (1)

$$\frac{d}{d\sigma}\mu_{2n} = -\frac{1}{\sigma}\mu_{2n} + \frac{1}{\sigma^3}\mu_{2n+2}$$

$$\Rightarrow \frac{1}{\sigma^3}\mu_{2n+2} = \frac{1}{\sigma}\mu_{2n} + \frac{d}{d\sigma}\mu_{2n}$$

Multiply by σ^3 to b/s

$\mu_{2n+2} = \sigma^2\mu_{2n} + \sigma^3\frac{d}{d\sigma}\mu_{2n}$ which is recurrence relation.

FORMULA for 95% confidence interval for mean of interval is

$$\boldsymbol{\mu \in \left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}},\ \bar{x} + \frac{1.96\,\sigma}{\sqrt{n}}\right)}$$

9.5. EXAMINATION ORIENTED EXERCISE/LESSON END EXERCISE

Q.1. An observed random sample size (from normal population with mean 2 and variance 4 has a sample mean 40. Obtain 95% confidence interval for mean of population.

Q.2. In a certain examination, the grades are normally distributed with average grade of 75 and S.Deviation 8. The instructor gave a grade A to students with marks 90 and above. If 12 students receive grade A, how many students took the examination.

Q.3. The height of male students in a college are normally distributed with mean 68.50 inches and standard deviation is2.3inches

a) what is the probability that one male student of college is over 6 feet tall.

b)what is the percentage of male students who are between 70 and 72inches tall.

Q.4. Find the median

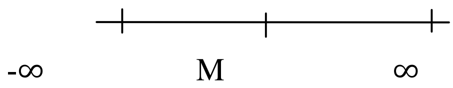
Sol 4 note that the median is the middle value of the distribution

Let X be continuous random variable with p.d.f

$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, $-\infty < x < \infty$

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

$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$



Suppose M is median

$= \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{M} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \frac{1}{\sqrt{2\pi\sigma}} \int_{M}^{\infty} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$

$\Rightarrow \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{M} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{2}$

$\Rightarrow \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\mu} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \frac{1}{\sqrt{2\pi}\,\sigma} \int_{\mu}^{M} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{2}$ --------------(1)

But $\frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\mu} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$

$= \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\mu} e^{\frac{-z^2}{2}} \sigma dz$

$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{\frac{-z^2}{2}} dz$

$= \frac{1}{2}$ [$\because \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{\frac{-z^2}{2}} dz = 1$

$\Rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{\frac{-z^2}{2}} dz = \frac{1}{2}$]

Use in (1)

z = x- μ
$\frac{dz}{dx} = \frac{1}{\sigma}$
dx= σdz
if x= -∞ , z = -∞
x= μ , z= 0

$\frac{1}{2} + \frac{1}{\sqrt{2\pi}\,\sigma}\int_{\mu}^{M} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{2}$

$\frac{1}{\sqrt{2\pi}\,\sigma}\int_{\mu}^{M} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$ =0

$\int_{\mu}^{M} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$ =0

This is possible only, if μ = M (i.e lower and upper limits of integrals are same)

Therefore median M of N.D = μ

Solution 1 : here x=9

μ =2

Var =4 = σ^2

$\bar{x}$ =40

95% confidence interval for mean of population

($\bar{x}$-1.96 $\frac{\sigma}{\sqrt{x}}$, $\bar{x} + 1.96\ \frac{\sigma}{\sqrt{x}}$)

=(40-1.96×1.96$\frac{2^2}{\sqrt{9}}$, 40+ 1.96× $\frac{2}{\sqrt{9}}$)

= 40-1.96× $\frac{2}{3}$, 40+1.96$\frac{2}{3}$

Sol: 3 in N.D

Mean (μ) = 68.50"

S.D (σ) = 2-3"

i) P[x>6ft]

ii)P[70<x<72]

Sol i) P[x>6ft] = p[x>72"]

=P[x- μ> 72- μ]

=P[$\frac{x-\mu}{\sigma}$>$\frac{72-\mu}{\sigma}$]

=P[z>$\frac{72-68.5}{2.3}$]

=P[z>$\frac{3.5}{2.3}$]

$= P[z>1.52]$

$=.5000 - P[z \leq 1.52]$

$=.5000-(.4357)$ from table

$=.0643$

(ii) $P[70<x<72]$

$=P[70-\mu<x-\mu<72-\mu]$

$=P[\frac{70-\mu}{\sigma}<\frac{x-\mu}{\sigma}<\frac{72-\mu}{\sigma}]$

$=P[\frac{70-68.5}{2.3}<z<\frac{70-68.5}{2.3}]$

$=P[\frac{1.5}{2.3}<z<1.52]$

$=P[.65<z<1.52]$

$=P[z=1.52]-P[z=.65]$ ($\because P[a<x<b]=P[x=b]-P[x=a]$)

$=.4357-.2422$

$=.1935$

Sol 2. $\sigma=8, \mu=75$

$P[x>90]$

$=P[x-\mu>90-\mu]$

$=P[\frac{x-\mu}{\sigma}>\frac{90-\mu}{\sigma}]$

$=P[z>\frac{90-75}{8}]$

$=P[z>\frac{15}{8}]$

$= P[z>1.87]$

$=0.5000-0.4693 = 0.0307$

Total no.of students who took examination is given by

$0.03N=12$

$N=400$

9.5. SUGGESTED READINGS:

The students are advised to go through following references for details.

9.6. References

1). Mathematical Statistics by J.N. Kapur & H.C. Saxena, S.Chand & Co.

2). Introduction to mathematical Statistics by P.G. Hoel, John Wiley & sons

3). Numerical Methods, Probability & statistics by Sandip Banerjee, Books & Allied Publications Ltd. Calcutta.

4). A text Book of Statistics, Numerical Analysis & Matrices by Sunil Gupta,Anupama Gupta, Narinder Kumar, Malhotra Brothers Pacca Danga, Jammu.

9.7. MODEL TEST PAPER

Q.1. In a certain examination, the grades are normally distributed with average grade of 75 and Standard Deviation 8. The instructor gave a grade A to students with marks 90 and above. If 12 students receive grade A, how many students took the examination.

Q.2. Find mean of N.D.

Q.3. Deriverecurrence formulae for N.D.

Q.4. Obtain 95% confidence interval of mean of populationofN.D.

Q.5. Find mean ofmean of n independent variables havingsame mean.

Q.6. In N.D. find mean, variance & mode.

STATISTICS

GAMMA DISTRIBUTION

By: Dr. Sunil Gupta

11.1. INTRODUCTION In this lesson special case of Normal Distribution in the form of Ã distribution as a continuous random variable is explained. Further the properties like mean, variance, moment generating functions etc. of this distribution are discussed..

11**.2. OBJECTIVES:** Objective of studying this lesson is to give the idea of a distribution which is defined for continuous random variable.

11.3. GAMMA DISTRIBUTION

11.3.1.Definition:-> A continuous random variable is said to be in Gamma (Γ) distribution if its p.d.f is given by

$$f(x) = \begin{cases} \frac{\beta^{\alpha}}{\Gamma\propto} e^{-\beta x} x^{\alpha-1}, x > 0 \\ 0, otherwise \end{cases}$$

where $\Gamma\gamma = \int_0^{\infty} e^{-x}\, e^{\gamma-1} dx$

Note:-> The above function is p.d.f. for Γ distribution.

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^{\infty} \frac{\beta^{\alpha}}{\Gamma\propto} e^{-\beta x} x^{\alpha-1}\, dx \text{ ------------------(1)}$$

Put βx= z $\Rightarrow$x=$\frac{z}{\beta}$

Differentiate b/s w.r.t.x

$\beta.1 = \frac{dz}{dx}$

dx = $\frac{dz}{\beta} \Big| \begin{matrix} if\ x=0 \Rightarrow z=0 \\ if\ x=\infty \Rightarrow z=\infty \end{matrix}$

use in (1)

$\int_{-\infty}^{\infty} f(x)$dx = $\int_0^{\infty} \frac{\beta^\alpha}{\Gamma\propto} e^{-z} (\frac{z}{\beta})^{\alpha-1} \frac{dz}{\beta}$

=$\int_0^{\infty} \frac{\beta^\alpha}{\Gamma\propto} e^{-z} \frac{z^{\alpha-1}}{\beta^{\alpha-1}.\beta}$ dz

=$\frac{\beta^\alpha}{\Gamma\propto} \int_0^{\infty} \frac{e^{-z} z^{\alpha-1}}{\beta^\alpha}$ dz

=$\frac{1}{\Gamma\propto} \int_0^{\infty} e^{-z} z^{\alpha-1}$ dz

=$\frac{1}{\Gamma\propto}.\Gamma \propto$ =1

11.3.2.Article :Find m.g.f of Gamma distribution

Proof let x be in***Γ*** distribution with p.d.f

f(x) = $\begin{cases} \frac{\beta^\alpha}{\Gamma\propto} e^{-\beta x} x^{\alpha-1}, x > 0 \\ 0, otherwise \end{cases}$

where $\Gamma \gamma = \int_0^{\infty} e^{-z} x^{l-1}$ dx

m.g.f =$M_x(t)$ = $E[e^{tx}]$

=$\int_{-\infty}^{\infty} e^{tx} f(x)$ dx

=$\int_0^{\infty} e^{tx} \frac{\beta^\alpha}{\Gamma\propto} e^{-\beta x} x^{\alpha-1}$dx

=$\frac{\beta^\alpha}{\Gamma\propto} \int_0^{\infty} e^{-(\beta-t)x} x^{\alpha-1}$dx

=$\frac{\beta^\alpha}{\Gamma\propto} \frac{\Gamma\propto}{(\beta-t)^\alpha}$

= $\frac{\beta^\alpha}{(\beta-t)^\alpha}$ --(*)

11.3.3.Cor:- find mean of Γ distribution and also find variance

Proof proceed upto (*)

Differentiating (*) w.r.t.'t'

$M_x'(t) = \beta^{\alpha} \frac{d}{dt}(\beta - t)^{-\alpha}$

$=\beta^{\alpha}(-\alpha)(\beta - t)^{-\alpha-1}(-1)$

$= \alpha\beta^{\alpha}(\beta - t)^{-\alpha-1}$ --(1)

Put t=0

$\mu_1' = M_x'(t)\Big|_{t=0} = \alpha\beta^{\alpha}(\beta)^{-\alpha-1}$

$= \alpha\beta^{\alpha}\beta^{-\alpha}(\beta)^{-1}$

Mean= $\mu_1' = \frac{\alpha}{\beta}$

<u>For variance</u> :

Differentiate (1) w.r.t.'t'

$\frac{d}{dt} M'_x(t) = \frac{d}{dt}\alpha\beta^{\alpha}(\beta - t)^{-\alpha-1}$

$= \propto \beta^{\propto}(-\propto -1)((\beta - t)^{-\propto-1-1}(-1)$

$\mu'_2 = M''_x(t)|_{t=0}$

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

$=\frac{\alpha(\alpha+1)}{\beta^2}$

Variance = $\mu_2' - \mu_1'^2$

$=\frac{\alpha(\alpha+1)}{\beta^2} - (\frac{\alpha}{\beta})^2$

$= \frac{\alpha(\alpha+1)}{\beta^2} - \frac{\alpha^2}{\beta^2} = \frac{\alpha^2+\alpha-\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}$

11.4. SUGGESTED READINGS:

The students are advised to go through following references for details.

11.5. REFERENCES

1). Mathematical Statistics by J.N. Kapur & H.C. Saxena, S.Chand & Co.

2). Introduction to mathematical Statistics by P.G. Hoel, John Wiley & sons

3). Numerical Methods, Probability & statistics by Sandip Banerjee, Books & Allied Publications Ltd. Calcutta.

4). A text Book of Statistics, Numerical Analysis & Matrices by Sunil Gupta,Anupama Gupta, Narinder Kumar, Malhotra Brothers Pacca Danga, Jammu.

11.6. MODEL TEST PAPER

Q.1. Find mean of Ã distribution.

Q.2. Derive m.g.f. of Ã distribution & derive its mean.

Q.3. Find variance of Ã distribution.

B.A. III Semester **Unit-3**

Course No. : UGM 0301 **Lesson 12**

STATISTICS

CHI-SQUARE DISTRIBUTION

By: Dr. Sunil Gupta

12.1. INTRODUCTION In this lesson Chi-square ($\div^2$) distribution as a continuous random variable is explained. Further the properties like mean, variance, moment generating functions etc. of this distribution are discussed..

12.2. Objectives: Objective of studying this lesson is to give the idea of a distribution which is defined for continuous random variable.

12.3 Chi-square ($\div^2$) distribution

<u>12.3.1.</u> <u>Definition:-></u> let X ~N (μ, σ^2) i.e X is in N.D with mean and variance & z= is S.N.V then

Square of S.N.V is

$z^2 = (\frac{x-\mu}{\sigma})^2$ is called chi- square distribution and is denoted by χ^2

i.e $\chi^2 = (\frac{x-\mu}{\sigma})^2$

NOTE :- The p.d.f of X^2 distribution is

$$f(x^2) = \frac{e^{\frac{-x^2}{2}}(x^2)^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma\frac{n}{2}}, \ 0<\chi^2<\infty$$

m.g.f of $\chi^2 = M_{\chi^2}(t) = E[e^{t\chi^2}]$

m.g.f of $\chi^2 = M_{\chi^2}(t) = E[e^{t\chi^2}]$

$$=\int_{-\infty}^{\infty} e^{t\chi^2} f(\chi^2) d\chi^2$$

$$=\int_{0}^{\infty} e^{t\chi^2} \frac{1}{2^{\frac{n}{2}}\Gamma\frac{n}{2}} e^{\frac{-\chi^2}{2}} (\chi^2)^{\frac{n}{2}-1} d\chi^2$$

$$=\frac{1}{2^{\frac{n}{2}}\Gamma\frac{n}{2}} \int_{0}^{\infty} e^{\frac{(2t-1)}{2}\chi^2} (\chi^2)^{\frac{n}{2}-1} d\chi^2$$

$$=\frac{1}{2^{\frac{n}{2}}\Gamma\frac{n}{2}} \int_{0}^{\infty} e^{\frac{-(1-2t)}{2}\chi^2} (\chi^2)^{\frac{n}{2}-1} d\chi^2$$

Use $\int_{0}^{\infty} e^{-ax} x^{l-1}\, dx = \frac{\Gamma l}{a^l}$

$$=\frac{1}{2^{\frac{n}{2}}\Gamma\frac{n}{2}} \frac{\Gamma\frac{n}{2}}{(\frac{1-2t}{2})^{\frac{n}{2}}}$$

$$=\frac{1}{2^{\frac{n}{2}}} \frac{1}{\frac{(1-2t)^{\frac{n}{2}}}{2^{\frac{n}{2}}}}$$

$M_{\chi^2}(t) = (1-2t)^{-\frac{n}{2}}$ ----------------------------(*)

$M_{\chi^2}'(t) = - -\frac{n}{2}(1-2t)^{-\frac{n}{2}-1} \frac{d}{dt}(1-2t)$

$=- -\frac{n}{2}(1-2t)^{-\frac{n}{2}-1}(-2)$

$=n(1-2t)^{-\frac{n}{2}-1}$ ....(1)

Put t=0

$\mu_1' = M_{\chi^2}'(t)|_{t=0} = n(1)^{-\frac{n}{2}-1} = n$

$\therefore$ Mean $= \mu_1' = n$

Differentiating (1) both side w.r.t. (t)

$M''_{\chi^2}(t) = n(-\frac{n}{2} - 1)(1 - 2t)^{-\frac{n}{2}-2}$ (-2)

$= n(-\frac{n-2}{\cancel{2}})\ (1 - 2t)^{-\frac{n}{2}-2}\ (-\cancel{2})$

$=-n(n+2)\ (1 - 2t)^{-\frac{n}{2}-1}\ (-\cancel{1})$

$=n(n+2)(1 - 2t)^{\frac{-n-2}{2}}$

$\mu_2' = M''_{\chi^2}(t)\ (t)|_{t=0} = n(n+2)$

$Var(\chi^2) = \mu_2' - \mu_1'^2$

$=n(n+2)-n^2$

$=n^2+2n-n^2 = 2n$

12.4. SUGGESTED READINGS:

The students are advised to go through following references for details.

12.5. REFERENCES

1). Mathematical Statistics by J.N. Kapur & H.C. Saxena, S.Chand & Co.

2). Introduction to mathematical Statistics by P.G. Hoel, John Wiley & sons

3). Numerical Methods, Probability & statistics by Sandip Banerjee, Books & Allied Publications Ltd. Calcutta.

4). A text Book of Statistics, Numerical Analysis & Matrices by Sunil Gupta, Anupama Gupta, Narinder Kumar, Malhotra Brothers Pacca Danga, Jammu.

12.6. MODEL TEST PAPER

Q.1. Find variance of Chi-square (÷^{2}) distribution.

Q.2. Derive m.g.f. of Chi-square (÷^{2}) distribution

B.A. III Semester	Unit-4
Course No. : UGM 0301	Lesson 13

NUMERICAL ANALYSIS

FINITE DIFFERENCE OPERATOR E AND D

By: Dr. Kamaljeet Kour

13.1 INTRODUCTION

The calculus of finite differences broadly deals with the changes that take place in the value of the function, the dependent variable due to changes in the independent variable. It is a study of the relations that exist between the values assumed by the function whenever the independent variable changes by finite jumps (intervals) whether equal or unequal.

The study of finite difference calculus has become important due to its application in computers and in the solutions of scientific problems.

13.2 OBJECTIVES

The main objective of this lesson is to make student familiar with the techniques of finite differences so that they are able to solve the numerical problems.

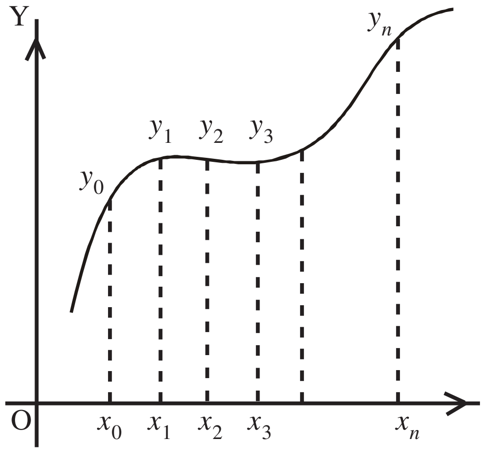
13.3. FINITE DIFFERENCE OPERATORS

Consider an interval I = $[a, b]$ and divide it into n sub-intervals (or n parts) of same size h such that $x_0 = a$, $x_1 = a + h$, $x_2 = a + 2h$, ..., $x_n = a + nh = b$. This h is called size of intervals or interval of differencing and a is the starting value.

Suppose $y = f(x)$ is a single valued continuous function of x taking y-values corresponding to values of x

as $y_0 = f(x_0)$, $y_1 = f(x_1)$, $y_2 = f(x_2)$, ..., $y_n = f(x_n)$ respectively for $x_0, x_1,$ ......, $x_n \in I$.

Then we define the following operators :



13.3.1 Forward difference operators

The first forward difference operator is denoted by Δ and is defined as

$$\Delta f(x) = f(x + h) - f(x) \text{ for all } x.$$

In particular, $\Delta y_0 = y_1 - y_0$, $\Delta y_1 = y_2 - y_1$, ... and in general $\Delta y_r = y_{r+1} - y_r$ for $0 \le r < n$.

13.3.2 Properties of Δ

(*i*) If $f(x) = k$, a constant for all x, then $\Delta f(x) = 0$.

(*ii*) Δ is linear, *i.e.* $\Delta[a f(x) + bg(x)] = a\, \Delta f(x) + b\Delta g(x)$.

(*iii*) $\Delta[f(x).\, g(x)] = f(x+h)\Delta g(x) + g(x)\Delta f(x)$.

(*iv*) $\left[\dfrac{f(x)}{g(x)}\right] = \dfrac{g(x)\Delta f(x) - f(x)\Delta g(x)}{g(x+h)\, g(x)}$.

The second forward difference operator and higher order forward differences are defined in a similar way and are given below :

$$\Delta^2 f(x) = \Delta(\Delta f(x)) = \Delta[f(x+h) - f(x)]$$

$$= \Delta f(x+h) - \Delta f(x), \text{ using linearity of } \Delta$$

$$= [f(x + 2h) - f(x+h)] - [f(x + h) - f(x)]$$

$$= f(x + 2h) - 2 f(x + h) + f(x)$$

and $\Delta^n f(x) = \Delta^{n-1} f(x+h) - \Delta^{n-1} f(x)$ for $n = 1, 2, 3,$ ....

More Properties of Δ

(*v*) $\Delta^n (\Delta^k f(x)) = \Delta^{n+k} f(x)$, where n, k positive integers.

(*vi*) If $f(x)$ is a polynomial of degree n, then $\Delta^n f(x)$ = constant and $\Delta^{n+1} f(x)$ = 0 for all n.

13.3.3 Forward difference table

Suppose the given points are (x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3) and (x_4, y_4).

By definition of Δ, we have $\Delta y_0 = y_1 - y_0$, $\Delta y_1 = y_2 - y_1$, $\Delta y_2 = y_3 - y_2$, $\Delta y_3 = y_4 - y_3$,

$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0,\ \Delta^2 y_1 = \Delta y_2 - \Delta y_1,\ \Delta^2 y_2 = \Delta y_3 - \Delta y_2,$$

$$\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0,\ \Delta^3 y_1 = \Delta^2 y_2 - \Delta^2 y_1 \text{ and } \Delta^4 y_0 = \Delta^3 y_1 - \Delta^2 y_0.$$

Then the forward difference table is given below :

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
x_0	y_0				
		Δy_0			
x_1	y_1		$\Delta^2 y_0$		
		Δy_1		$\Delta^3 y_0$	
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$
		Δy_2		$\Delta^3 y_1$	
x_3	y_3		$\Delta^2 y_2$		
		Δy_3			
x_4	y_4				

Example 1 : Construct the forward difference table for the following data :

x	:	0	1	2	3	4
y	:	7	10	13	22	43

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
0	7				
		3			
1	10		0		
		3		6	
2	13		6		0
		9		6	
3	22		12		
		21			
4	43				

Example 2 : Given $f(0) = 3$, $f(1) = 12$, $f(2) = 81$, $f(3) = 200$, $f(4) = 100$ and $f(5) = 8$. Find $\Delta^5 f(0)$ by using forward differnce table.

Solution : Here the given data is

$x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$, $x_5 = 5$ and $y_0 = 3$, $y_1 = 12$, $y_2 = 81$, $y_3 = 200$, $y_4 = 100$, $y_5 = 8$.

The forward difference table is constructed below :

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
0	3					
		9				
1	12		60	−10		
		69			−259	
2	81		50	−269		755
		119			496	
3	200		−219	227		
		−100				
4	100		8			
		−92				
5	8					

From the table, we note that $\Delta^5 f(0) = 755$. Also, we can have other values such as :

$\Delta^4 f(1) = 496$ and $\Delta^2 f(2) = -219$.

Example 3 : Evaluate the following

(*i*) $\Delta(\log x)$ (*ii*) $\Delta^2 (x^3)$ (*iii*) $\Delta\sin(ax + b)$

(*iv*) $\Delta(e^{ax} \log bx)$

Solution : (*i*) Here $f(x) = \log x$. Then by definition of Δ, we have

$\Delta(\log x) = \Delta f(x) = f(x + h) - f(x)$

$$= \log(x + h) - \log x = \log\left(\frac{x+h}{x}\right) = \log\left(1+\frac{h}{x}\right).$$

(*ii*) Here $f(x) = x^3$. Then $f(x + h) = (x + h)^3$

$$= x^3 + 3hx^2 + 3h^2x + h^3$$

and $f(x + 2h) = (x + 2h)^3 = x^3 + 6hx^2 + 12h^2x + 8h^3$.

By definition of Δ, we have

$\Delta^2(x^3) = \Delta^2 f(x) = f(x + 2h) - 2f(x + h) + f(x)$

$$= (x^3 + 6hx^2 + 12h^2x + 8h^3) - 2(x^3 + 3hx^2 + 3h^2x + h^3) + x^3$$

$$= 6h^2(x + h).$$

(*iii*) Here $f(x) = \sin(ax + b)$ and $f(x + h) = \sin(ax + ah + b)$.

Then by definition of Δ, we have

$\Delta \sin(ax + b) = \Delta f(x) = (x + h) - f(x)$

$$= \sin(ax + ah + b) - \sin(ax + b)$$

$$= 2 \sin \frac{ah}{2} . \cos\left(ax + b + \frac{ah}{2}\right).$$

(*iv*) Here $f(x) = e^{ax}$ and $g(x) = \log bx$. Then $f(x + h) = e^{ax}.e^{ah}$

and $g(x+h) = \log(bx+bh)$.

Now by property (*iii*) for forward difference of product, we have

$$\Delta[f(x)\, g(x)] = f(x+h)\Delta g(x) + g(x)\, \Delta f(x)$$

$$\Rightarrow \quad \Delta(e^{ax} \log bx) = e^{ax}.e^{ah}\, [g(x+h) - g(x)] + \log bx\, [f(x+h) - f(x)]$$

$$= e^{ax}.e^{ah}\, [\log(bx+bh) - \log bx] + \log bx\, [e^{ax}.e^{ah} - e^{ax}]$$

$$= e^{ax}.e^{ah} \log\left(1+\frac{h}{x}\right) + e^{ax}\,(e^{ah}-1) \log bx$$

$$= e^{ax}\, [e^{ah} \log\left(1+\frac{h}{x}\right) + (e^{ah}-1) \log bx\,].$$

13.3.4 Backward Difference Operators

The first backward difference operator is denoted by ∇ and is defined as

$$f(x) = f(x) - f(x-h) \text{ for all } x.$$

In particular, $y_1 = y_1 - y_0$, $y_2 = y_2 - y_1$, ... and in general $y_r = y_r - y_{r-1}$ for $1 \le r \le n$.
∇

The second backward difference operator and higher order backward difference operators are defined as under :

$$^2 f(x) = \;\;(\;\; f(x)) = \;\; [f(x) - f(x-h)]$$

$$= \;\; f(x) - \;\; f(x-h), \text{ by linearity of}$$

$$= [f(x) - f(x-h)] - [f(x-h) - f(x-2h)]$$

$$= f(x) - 2f(x-h) + f(x-2h)$$

and $\quad {}^n f(x) = \;\; {}^{n-1} f(x) - \;\; {}^{n-1} f(x-h)$ for $n = 1, 2, 3, ...$

13.3.5 Properties of . The backward difference operator has the following properties similar to that of Δ.

(*i*) $\quad f(x) = 0$, where $f(x)$ is a constant function.

(*ii*) $\quad$ is linear, *i.e.,* $\;\; [a f(x) + b g(x)] = a \;\; f(x) + b \;\; g(x)$.

(iii) $\quad (f(x)\, g(x)) = f(x-h) \quad g(x) + g(x) \quad f(x)$.

(iv) $$= \frac{g(x)\nabla f(x) - f(x)\nabla g(x)}{g(x-h)\, g(x)}.$$

(v) $\quad {}^{n} (\quad {}^{k} f(x)) = \quad {}^{n+k} f(x)$, where n, k positive integers.

(vi) If $f(x)$ is a polynomial of degree n, then $\quad {}^{n} f(x)$ = constant and $\quad {}^{n+1} f$ $(x) = 0$
for all n.

13.3.6 Backward difference table

For the given values (x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3) and (x_4, y_4), the backward difference table is given below :

x	y	y	${}^2y$	${}^3y$	${}^4y$
x_0	y_0				
		y_1			
x_1	y_1		${}^2y_2$		
		y_2		${}^3y_3$	
x_2	y_2		$\nabla\left[\frac{f(x)}{g(x)}\right]$ ${}^2y_3$		${}^4y_4$
		y_3		${}^3y_4$	
x_3	y_3		${}^2y_4$		
		y_4			
x_4	y_4				

Here $\quad y_4 = y_4 - y_3$, $\quad y_3 = y_3 - y_2$, $\quad y_2 = y_2 - y_1$, $\quad y_1 = y_1 - y_0$,

${}^2y_4 = \quad y_4 - \quad y_3$, $\quad {}^2y_3 = \quad y_3 - \quad y_2$, $\quad {}^2y_2 = \quad y_2 - \quad y_1$,

${}^3y_4 = \quad {}^2y_4 - \quad {}^2y_3$, $\quad {}^3y_3 = \quad {}^2y_3 - \quad {}^2y_2$ and $\quad {}^4y_4 = \quad {}^3y_4 - \quad {}^3y_3$.

Example 4 : Construct the backward difference table for the following data :

x	:	1	2	3	4	5
y	:	10.1	12.5	11.3	12.3	13.4

Solution : The backward difference table is given below :

x	y	y	^{2}y	^{3}y	^{4}y
1	10.1				
		2.4			
2	12.5		−3.6		
		−1.2		5.8	
3	11.3		2.2		−7.9
		1		−2.1	
4	12.3		0.1		
		1.1			
5	13.4				

Example 5 : If $f(x) = 4x^2 - 3x + 5$ for $x = 0, 1, 2, 3, 4$, then construct the backward difference table.

Solution : Given $f(x) = 4x^2 - 3x + 5$

$\Rightarrow$ $f(0) = 5$, $f(1) = 6$, $f(2) = 15$, $f(3) = 32$ and $f(4) = 57$.

So x and y values are :

$x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $x_3 = 3, \nabla x_4 = 4$ and $y_0 = 5$, $y_1 = 6$, $y_2 = 15$,

$y_3 = 32$, $y_4 = 57$.

The backward difference table for this data is constructed below :

x	y	y	^{2}y	^{3}y	^{4}y
0	5				
		1			
1	6		8		
		9		0	
2	15		8		0
		17		0	
3	32		8		
		25			
4	57				

From the table, we note that $\Delta^2 f(x) = 8$ (a constant) and $\Delta^3 f(x) = 0$, Δ^4

$f(x) = 0$.

Example 6 : Evaluate the following :

(i) $\cos(2x+1)$ (ii) $(\quad -2)^2 (x+2)^2$, taking $h = 1$

(iii) $(\Delta + \quad)^2 (x^3)$.

Solution : (i) Let $f(x) = \cos(2x+1)$. Then $f(x-h) = \cos(2x-2h+1)$.

By definition of , we have

$$\cos(2x+1) = \quad f(x) = f(x) - f(x-h)$$

$$= \cos(2x+1) - \cos(2x-2h+1)$$

$$= -2 \sin(2x+1-h) \sinh.$$

(ii) Let $f(x) = (x+2)^2$. Then

$f(x-h) = f(x-1) = (x+1)^2$ and $f(x-2h) = f(x-2) = x^2$ ($h = 1$).

Also $(\quad -2)^2 = \quad ^2 - 4 \quad + 4$.

∇

So $(\quad -2)^2 (x+2)^2 = (\quad ^2 - 4 \quad + 4) f(x)$

$$= \quad ^2 f(x) - 4 \quad f(x) + 4f(x)$$

$$= [f(x) - 2f(x-1) + f(x-2)] - 4[f(x) - f(x-1)] + 4f(x)$$

$$= f(x) + 2f(x-1) + f(x-2)$$

$$= (x+2)^2 + 2(x+1)^2 + x^2$$

$$= x^2 + 4x + 8 + 2(x^2 + 2x + 1) + x^2$$

$$= 4x^2 + 8x + 10.$$

(iii) Let $f(x) = x^2$. The $f(x+h) = (x+h)^3$ and $f(x-h) = (x-h)^3$.

Now $(\Delta + \quad)(x^3) = \Delta(x^3) + \quad (x^3)$

$$= [(x+h)^3 - x^3] + [x^3 - (x-h)^3]$$

$$= (x+h)^3 - (x-h)^3$$

$= a^3 - b^3$, where say $a = x+h$, $b = x-h$

$= (a-b)(a^2 + ab + b^2)$

$= 2h\,[(x+h)^2 + (x+h)(x-h) + (x-h)^2]$

$= 2h\,[3x^2 + h^2]$

and $(\Delta + \quad)^2 (x^3) = (\Delta + \quad)[(\Delta + \quad)(x^3)]$

$= (\Delta + \quad)[2h\,(3x^2 + h^2)]$

$= 2h.\,(\Delta + \quad)(3x^2 + h^2)$, by linearity

$= 2h\,[(\Delta + \quad)(3x^2) + (\Delta + \quad)(h^2)]$, by linearity

$= 2h\,[(3\Delta(x^2) + 3 \quad x^2 + 0]$, by linearity and by propperty (1)

$= 6h\,[(x+h)^2 - x^2 + x^2 - (x-h)^2]$

$= 6h\,[(x+h)^2 - (x-h)^2] = 6h\,(4xh) = 24xh^2$.

13.3.7 Other Operators

(*i*) **Shift operator :** It is denoted by E and is defined as $E[f(x)] = f(x+h)$ for all x. In particular, $Ey_0 = y_1$, $Ey_1 = y_2$, $Ey_2 = y_3$, ... and in general $Ey_r = y_{r+1}$ for $0 \le r < n$. It is forward shift, as it shifts y_r to y_{r+1} when E is applied.

Further $E^n[f(x)] = E^{n-1}[f(x+h)] = f(x+nh)$, where n is any rational number.

Similarly, we have $E^{-1}[f(x)] = f(x-h)$ for all x (called backward shift operator).

Also $\quad [f(x)] = f \quad$ and $E^{-\frac{1}{2}}[f(x)] = f\left(x - \frac{h}{2}\right)$ for all x.

In general, we have $E^n y_k = y_{n+k}$ for any rational numbers n, k.

(*ii*) **Averaging Operator :** It is denoted by μ and is defined as

$$\mu\,[f(x)] = \frac{1}{2}\left[f\left(x + \frac{h}{2}\right) + f\left(x - \frac{h}{2}\right)\right] \text{ for all } x.$$

(*iii*) **Central Difference Operator** : It is denoted by and is defined as

$$\delta\,[\,f(x)] = f\left(x+\frac{h}{2}\right) - f\left(x-\frac{h}{2}\right) \text{ for all } x.$$

Example 7 : Construct the forward difference table for the following data and show that $\Delta^2 y_2 = 2$.

x	:	1	2	3	4	5
y	:	2	5	10	17	26

Solution : Here $h = 1$, $x_0 = 1$, $x_1 = 2$, $x_2 = 3$, $x_3 = 4$, $x_4 = 5$, $y_0 = 2$, $y_1 = 5$, $y_2 = 10$, $y_3 = 17$ and $y_4 = 26$.

Central Difference table is constructed below :

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
1	2				
		3			
2	5		2		
		5		0	
3	10		2		0
		7		0	
4	17		2		
		9			
5	26				

Using table, we note that $\Delta^2 y_2 = 2$.

13.3.8 RELATION BETWEEN THE OPERATORS

(1) $\Delta = E - 1$ or $E = 1 + \Delta$.

(2) $\nabla = 1 - E^{-1}$ or $E^{-1} = 1 -$.

(3) $\mu =$ $(E^{1/2} + E^{-1/2})$.

(4) $\delta = E^{1/2} - E^{-1/2}$.

Proof : (1) By definition of Δ, we have

$$\Delta[f(x)] = f(x+h) - f(x) = \mathrm{E}\,[f(x)] - f(x)$$

$$= (\mathrm{E}-1)\, f(x) \text{ for any } f \text{ and } x$$

$$\Rightarrow \quad \Delta = \mathrm{E} - 1.$$

(2) By definition of ∇, we have

$$[f(x)] = f(x) - f(x-h) = f(x) - \mathrm{E}^{-1}\,[f(x)]$$

$$= (1 - \mathrm{E}^{-1})\, f(x) \text{ for any } f \text{ and } x$$

$$\Rightarrow \quad = 1 - \mathrm{E}^{-1}.$$

(3) By definition of μ, we have

$$\mu\,[f(x)] = \quad [f(x + \frac{h}{2}) + f(x - \frac{h}{2})]$$

$$= \frac{1}{2}[\,\mathrm{E}^{1/2}\,[f(x)] + \mathrm{E}^{-1/2}\,[f(x)]$$

$$= \frac{1}{2}(\mathrm{E}^{1/2} + \mathrm{E}^{-1/2})\;\, f(x) \text{ for any } f \text{ and } x$$

$$\Rightarrow \quad \mu = \frac{1}{2}[\mathrm{E}^{1/2} + \mathrm{E}^{-1/2}].$$

(4) By definition of δ, we have

$$\delta\,[f(x)] = f(x + \frac{h}{2}) - f(x - \frac{h}{2})$$

$$= \mathrm{E}^{1/2}\,[f(x)] - \mathrm{E}^{-1/2}\,[f(x)]$$

$$= [\mathrm{E}^{1/2} - \mathrm{E}^{-1/2}]\;\, f(x) \text{ for any } f \text{ and } x$$

$$\Rightarrow \quad \delta = \mathrm{E}^{1/2} - \mathrm{E}^{-1/2}.$$

Example 8 : Prove the following :

(*i*) $\Delta\nabla = \quad \Delta = \delta^2$.

(*ii*) $\Delta + \quad = \quad - \quad$.

(*iii*) $\Delta \quad = \nabla \quad = \delta$.

Solution : (*i*) Using $\Delta = E - 1$, $\quad = 1 - E^{-1}$ and $\delta = \quad - \quad$, we have

$$\Delta\nabla = (E-1)(1-E^{-1}) = E + E^{-1} - 2$$

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

and $\delta^2 = (\quad - \quad)^2 = E + E^{-1} - 2$.

Thus $\Delta\nabla = \quad \Delta = \delta^2 = E + E^{-1} - 2$

(*ii*) L.H.S. = $\Delta + \quad = (E-1) + (1-E^{-1}) = E - E^{-1}$

$\frac{\nabla^2}{\nabla\Delta}$

and R.H.S. = $\quad - \quad =$

$$= \frac{(E-1)^2 - (1-E^{-1})^2}{\nabla\Delta} = \frac{(E^2 - 2E + 1) - (1 - 2E^{-1} + E^{-2})}{\nabla\Delta}$$

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

$= E - E^{-1}$ = L.H.S., hence proved.

(*iii*) Using relationship for Δ, ∇, E and δ, we have

$$\Delta \quad = (E-1)E^{\frac{-1}{2}} = E^{\frac{1}{2}} -$$

and $\Delta E^{\frac{1}{2}} = (1-E^{-1}) \quad = \quad - \quad = \delta$.

Thus we note that $\Delta E^{\frac{-1}{2}} = \nabla \qquad = \delta\cdot$

Example 9 : Evaluate $\qquad f(x)$ and $\dfrac{\Delta^2[f(x)]}{E[f(x)]}$. Are these values equal?

Solution : We know $\Delta = E-1$. So

$$\frac{\Delta^2}{E} = (E-1)^2 E^{-1} = (E^2-2E+1)E^{-1} = E-2+E^{-1}.$$

Now $\quad \dfrac{\Delta^2}{E}[f(x)] = (E-2+E^{-1})f(x) = E[f(x)]-2f(x)+E^{-1}[f(x)]$

$$= f(x+h)-2f(x)+f(x-h)$$

and $\quad \dfrac{\Delta^2[f(x)]}{E f(x)} = \dfrac{(E-1)^2 f(x)}{f(x+h)} = \dfrac{(E^2-2E+1)f(x)}{f(x+h)}$

$$= \frac{f(x+2h)-2f(x+h)+f(x)}{f(x+h)}.$$

We note that the two values are different $\left(\dfrac{\Delta^2}{E}\right)$. For example, take $h = 1$ and $f(x) = x^2$, then

$$(x^2) = (x+1)^2-2x^2+(x-1)^2 = 2$$

where as $\quad \dfrac{\Delta^2(x^2)}{E(x^2)} = \dfrac{2!}{(x+1)^2} = \dfrac{2}{(x+1)^2}.$

Note: The two values of above example are equal only if $f(x)$ is a constant function.

Example 10 : Let D be the differential operator $\dfrac{d}{dx}$. Then prove that

(*i*) $\qquad E = e^{hD}$.

(*ii*) $\qquad D = \dfrac{1}{h}\left[\Delta - \dfrac{\Delta^2}{2} + \dfrac{\Delta^3}{3} - \cdots\cdots\cdots\right].$

(*iii*) D = $\frac{1}{h}\left[\nabla + \frac{\nabla^2}{2} + \frac{\nabla^3}{3} + \ldots\ldots\ldots\ldots\right]$.

(*iv*) $\nabla = 1 - e^{-hD}$

Solution : (*i*) By definition of E, we have

$E\,f(x) = f(x+h)$

$= f(x) + hf'(x) + \quad f''(x) + \frac{h^3}{3!} f'''(x) + \ldots\ldots$, by Taylor's expansion

$= f(x) + hD\,[f(x)] + \frac{h^2}{2!} D^2\,[f(x)] + \ldots\ldots$

$= \left[1 + hD + \frac{h^2 D^2}{2!} + \frac{h^3 D^3}{3!} + \ldots\ldots\right] f(x)$

$= e^{hD} f(x)$ for any f and x

⇒ $E = e^{hD}$.

(*ii*) From part (*i*), we have $E = e^{hD}$

⇒ $1 + \Delta = e^{hD}$.

Apply log on both sides, we have

$hD = \log(1+\Delta) = \Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \ldots\ldots$ by logarithmic series

⇒ D = $\frac{1}{h}\left[\Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \ldots\ldots\ldots\ldots\right]$.

(*iii*) From part (*i*), we have $E = e^{hD}$ ⇒ $E^{-1} = e^{-hD}$.

Now $e^{-hD} = E^{-1} = 1 - \nabla$.

Applying log on both sides, we have

$-hD = \log(1 - \quad) = - \quad - \quad - \frac{\Delta^3}{3} - \ldots\ldots$ by logarithmic series

$$\Rightarrow \qquad D = \frac{1}{h}\left[\nabla + \frac{\nabla^2}{2} + \frac{\nabla^3}{3} + \ldots\ldots\ldots\right].$$

(*iv*) We have $\nabla = 1 - E^{-1} = 1 - e^{-hD}$, by using part (*i*).

Example 11 : Show that

$$y_0 + \quad y_1 + \frac{x^2}{2!}y_2 + \frac{x^3}{3!}y_3 + \ldots.. = e^x\,[y_0 + x\Delta y_0 + \frac{x^2}{2!}\Delta^2 y_0 + \frac{x^3}{3!}\Delta^3 y_0 + \ldots..]$$

Solution : Using $y_1 = Ey_0$, $y_2 = E^2 y_0$, $y_3 = E^3 y_0$, ..., we have

$$\text{L.H.S.} = y_0 + \frac{x}{1!}y_1 + \frac{x^2}{2!}y_2 + \frac{x^3}{3!}y_3 + \ldots..$$

$$= y_0 + \frac{x}{1!}Ey_0 + \frac{x^2}{2!}E^2 y_0 + \frac{x^3}{3!}E^3 y_0 + \ldots..$$

$$= \left(1 + \frac{xE}{1!} + \frac{x^2E^2}{2!} + \frac{x^3E^3}{3!} + \ldots\ldots\right)y_0$$

$= e^{xE}y_0 = e^{x(1+\Delta)}\, y_0$, using $\frac{x}{1!}$ E = 1+Δ

$= e^x .\, e^{x\Delta}\, y_0$

$$= e^x\left(1 + x\Delta + \frac{x^2\Delta^2}{2!} + \frac{x^3\Delta^3}{3!} + \ldots\ldots\right)y_0$$

$$= e^x\left[y_0 + x\Delta y_0 + \frac{x^2\Delta^2}{2!}y_0 + \frac{x^3\Delta^3}{3!}y_0 + \ldots\ldots\right] = \text{R.H.S.}$$

Hence proved.

13.3.9 FACTORIAL POLYNOMIAL

If n is a positive integer, then the ***factorial polynomial function*** is denoted $x^{(n)}$ and is defined as the product of n factors :

$x, x-h, x-2h, \ldots\ldots, x-(n-1)\,h.$

Thus $x^{(n)} = x(x-h)\,(x-2h)\,.......\,[x-(n-1)h]$.

So $x^{(0)} = 1$, $x^{(1)} = x$, $x^{(2)} = x(x-h)$, $x^{(3)} = x(x-h)\,(x-2h)$, ...

Now let us find difference of factorial functions :

(*i*) $\Delta[x^{(2)}] = (x+h)^{(2)} - x^{(2)} = (x+h)x - x(x-h)$

$= x\,[(x+h)-(x-h)] = 2hx = 2hx^{(1)}$.

(*ii*) $\Delta[x^{(3)}] = (x+h)^{(3)} - x^{(3)}$

$= (x+h)\,x(x-h) - x(x-h)(x-2h)$

$= x\,(x-h)\,[(x+h)-(x-2h)]$

$= x\,(x-h)\,3h$

$= 3hx^{(2)}$.

In a similar way, we can verify the following :

(*iii*) $\Delta x^{(n)} = nhx^{(n-1)}$ for all $n \geq 1$.

(*iv*) $\Delta^2[x^{(n)}] = \Delta[nh\,x^{(n-1)}] = nh\,\Delta x^{(n-1)}$

$= nh\,(n-1)\,h\,x^{(n-2)}$

$= n\,(n-1)\,h^2\,x^{(n-2)}$,

$\Delta^3[x^{(n)}] = n(n-1)\,(n-2)\,h^3\,x^{(n-3)}$,

$\Delta^n[x^{(n)}] = n(n-1)\,(n-2)\,......\,3.2.1.h^n$

$= n!\,h^n$, (a constant)

and $\Delta^{n+1}[x^{(n)}] = 0$.

13.3.10 Properties of Factorial functions

(*i*) When $h = 0$, $x^{(n)} = x^n$

(*ii*) When $h = 1$, $x^{(n)} = x(x-1)\,(x-2)\,.......\,[x-(n-1)]$

and $\Delta x^{(n)} = nx^{(n-1)}$ for all n. This further, gives $\dfrac{1}{\Delta}[x^{(n-1)}] = \dfrac{x^{(n)}}{n}$.

Thus for $h = 1$, Δ and $\frac{1}{\Delta}$ behave as derivative and integration respectively.

Example 12 : Express the function $f(x) = 3x^3 + x^2 + x + 1$ and its differences in factorial notation, taking $h = 1$.

Solution : To express the given function in factorial notation, write

$$f(x) = Ax^{(3)} + B\,x^{(2)} + Cx^{(1)} + D$$

$$\Rightarrow \quad 3x^3 + x^2 + x + 1 = Ax\,(x-1)(x-2) + Bx\,(x-1) + (Cx + D) \qquad ....(1)$$

Put $x = 0$, 1 and 2 respectively in (1), we obtain

$$D = 1,\ 6 = C + 1 \text{ or } C = 5, \text{ and}$$

$$31 = 0 + 2B + 10 + 1 \text{ or } B = 10.$$

Equating coefficients of x^3 on both sides of (1), we obtain A = 3.

Using these values of A, B, C, D, we get the factorial notation of $f(x)$ as

$$f(x) = 3x^{(3)} + 10\,x^{(2)} + 5x^{(1)} + 1.$$

Now $\Delta\,[f(x)] = 9x^{(2)} + 20\,x^{(1)} + 5$, by property (*ii*)

$\Delta^2\,[f(x)] = 18x^{(1)} + 20$,

$\Delta^3\,[f(x)] = 18$ and $\Delta^4\,[f(x)] = 0$.

Example 13 : Find a function whose first difference $x^3 + 3x^2 + 5x + 12$, by taking $h = 1$.

Solution : Let the required function be $f(x)$ whose first difference is

$$g(x) = x^3 + 3x^2 + 5x + 12.$$

Then $\Delta[f(x)] = g(x)$ or $f(x) = \frac{1}{\Delta}[g(x)]$.

To evaluate this, we have to first express $g(x)$ in factorial notation.

Write $g(x) = Ax^{(3)} + Bx^{(2)} + Cx^{(1)} + D$ ...(1)

$$\Rightarrow \quad x^3 + 3x^2 + 5x + 12 = A.x\,(x-1)(x-2) + B.x\,(x-1) + Cx + D \qquad ...(2)$$

Put $x = 0$, 1 and 2 respectively in (2), we obtain

D = 12, 21 = C + 12 or C = 9 and 42 = 2B + 18 + 12 or B = 6.

Equating coefficient of x^3 on both sides of (2), we obtain A = 1.

Put these values of A, B, C, D in (1), we have

$$g(x) = x^{(3)} + 6x^{(2)} + 9x^{(1)} + 12.$$

$$\text{Now } f(x) = \frac{1}{\Delta}[g(x)] = \frac{1}{\Delta}[x^{(3)} + 6x^{(2)} + 9x^{(1)} + 12]$$

$$= \frac{1}{4}x^{(4)} + 2x^{(3)} + \frac{9}{2}x^{(2)} + 12x^{(1)} + c$$

$$= \frac{1}{4}[x(x-1)(x-2)(x-3) + 8x(x-1)(x-2) + 18x(x-1) + 48x] + c$$

$$= \frac{1}{4}[x(x-1)(x-2)\{x-3+8\} + 18x(x-1) + 48x] + c$$

$$= \frac{1}{4}[x(x-1)(x-2)(x+5) + 18x(x-1) + 48x] + c$$

$$= \frac{1}{4}[x(x-1)\{x^2 + 3x - 10 + 18\} + 48x] + c$$

$$= \frac{1}{4}[x^4 + 3x^3 + 8x^2 - x^3 - 3x^2 - 8x + 48x] + c$$

$$= \frac{1}{4}(x^4 + 2x^3 + 5x^2 + 40x) + c$$

13.4 SUMMARY

In this lesson, we studied forward difference, backward difference and shift operators alongwith construction of their difference tables. Also we studied factorial polynomial function alongwith some solved problems.

13.5 EXERCISE

1. Construct forward difference table for the following set of values :

x	:	6	7	8	9	10
y	:	2.449	2.646	2.828	3	3.162

2. Construct backward difference table for the following set of values and also find the value of $\nabla^2 y_5$:

x	:	1	2	3	4	5
y	:	2	5	10	17	26

3. Evaluate the following, taking $h = 1$:

(*i*) $\Delta(e^{2x} \log 3x)$.

(*ii*) $\Delta(\tan^{-1} ax)$.

(*iii*) $\Delta(x \log x)$.

(*iv*) $(\Delta^2 + \Delta - 1)(x^2 + 2x + 1)$.

(*v*) $(\quad + \Delta)^2 (x^2 + x)$. $\quad \frac{\nabla\Delta^2 (x^3)}{EE\, x^3}$

4. Evaluate : $\sin(x + h) + \frac{\Delta^2 \sin(x+h)}{E \sin(x+h)}$.

5. Evaluate : $\Delta(x^2 e^x)$ and $\nabla(e^{2x} \log 3x)$.

6. Show that $\quad \neq \left(\frac{\Delta^2}{E}\right) x^3$.

7. Define the operators Δ, ∇, μ, δ, E and E^{-1}.

Also prove the following :

(*i*) $\Delta = E \quad$.

(*ii*) $\quad = E^{-1}\Delta$.

(*iii*) $\Delta\,(y_r^{\,2}) = (y_r + y_{r+1})\,\Delta y_r.$

(*iv*) $hD = -\log\,(1- \quad)$.

8. Show that :

$$y_0 - y_1 + y_2 - y_3 + \ldots\ldots = \quad y_0 - \frac{1}{4}\Delta y_0 + \frac{1}{8}\Delta^2 y_0 - \frac{1}{16}\Delta^3 y_0 + \ldots\ldots$$

9. Express the following functions and their successive differences in factorial notation, by taking $h = 1$:

(*i*) $2x^3 - 9x^2 + 3x - 10$

(*ii*) $2x^3 + 5x^2 - 7x + 5$

10. Obtain the function whose first difference is

$9x^2 + 11x + 5$

ANSWERS

2. 2.

3. (*i*) $e^{2x}\,[e^2 \log\,(1+\frac{1}{x}) + (e^2 \overset{\nabla}{\frac{1}{2}}\!-1) \log 3x]$

(*ii*) $\tan^{-1}\left(\dfrac{a}{1+a^2x+a^2x^2}\right)$

(*iii*) $x \log\,(1+\frac{1}{x}) + \log\,(x+1)$

(*iv*) $4 - x^2$

(*v*) 8

4. $2(\cos h - 1)\,[\sin\,(x+h) + 1]$

5. $e^x\,[(x+h)^2 e^h - x^2]$; $\dfrac{e^{2x}}{e^{2h}}\,[(e^{2h}-1)\log 3x - \log\,(1-\frac{h}{x})]$

9. (*i*) $2x^{(3)} - 3x^{(2)} - 4x^{(1)} - 10$

(*ii*) $2x^{(3)}+13x^{(2)}-2x^{(1)}+5$

10. $3x^3+x^2+x+c$

13.6 SUGGESTED READING/REFERENCES

1. "Introductory Methods of Numerical Analysis" by S.S.Sastry, PHI Learning Pvt Ltd., New Delhi.
2. "Numerical Methods in Engineering and Science" by B.S.Grewal, Khanna Publishers, New Delhi.
3. "A text book of Complex Variables & Numerical Methods" by Bhopinder Singh & Others, Kirti Publishers, Jammu.

13.7 MODEL TEST PAPER

1. If $f(x) = 2x^2 + 5x - 7$ for $x = 1, 2, 3, 4, 5$; then construct the forward difference table.

2. Construct the backward difference table for the following data :

x :	1.8	2.0	2.2	2.4
y :	2.9422	3.6269	4.4571	5.4662

3. If $f(0) = 7$, $f(1) = 10$, $f(2) = 13$, $f(3)=22$ and $f(4) = 43$, show that $\nabla^2 f(4) = 12$.

4. Construct the forward difference table for the following data and show that $\Delta^2 y_2 = 2$.

x :	1	2	3	4	5
y :	2	5	10	17	26

5. Taking $h = 1$, show that

(*i*) $(3\Delta + 1)^2 (x + 3)^2 = x^2 + 8x + 69$

(*ii*) $^2 [(2x + 1)(3x + 1)] = 12.$

6. Evaluate (x^3), taking $h = 1$.

7. Show that : $e^x = \left(\frac{\Delta^2}{E}\right) e^x \cdot \frac{Ee^x}{\Delta^2 (e^x)}$ taking $h = 1$.

8. Express the following functions and their successive differences in factorial notation, by taking $h = 1$.

(*i*) $x^4 - 12x^3 + 24x^2 - 30x + 9$

(*ii*) $x^4 + 3x^3 + 5x + 2$

9. Determine the function whose first forward difference is $2x^3 + 3x^2 - 5x + 4$, by taking $h = 1$.

$$\left(\frac{\nabla\Delta^2}{E}\right)$$

NUMERICAL ANALYSIS

NEWTON'S FORWARD AND BACKWARD DIFFERENCE INTERPOLATION FORMULA. LAGRANGE'S FORMULA

By: Dr. Kamaljeet Kour

14.1 INTRODUCTION

Let $y = f(x)$ be a given function of x which is single-valued and continuous. Then the value of $f(x)$ corresponding to certain given values of x, say $x_0, x_1, x_2, \ldots\ldots, x_n$ can easily be calculated. But in numerical analysis, it is otherway. That is, for a given set of $(n+1)$ values $(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots\ldots, (x_n, y_n)$, it is required to find the function $y = f(x)$ passing through the given points (x_k, y_k), $0 \leq k \leq n$.

The process of finding a function explitly in terms of x (*i.e.* a function expressed in terms of x) in the interval $[x_o, x_n]$ agreeing on the given set of points is called **interpolation**. In most of the problems, it is a polynomial function.

14.2 OBJECTIVES

The objective of this lesson is to study the techniques of interpolating the unknown value with the help of a given set of values.

14.3 NEWTON'S FORWARD INTERPOLATION FORMULA

Given a set of $(n+1)$ values : $(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots\ldots, (x_n, y_n)$ of the function $y = f(x)$, it is required to find a polynomial function $P_n(x)$ of nth degree such that y and $P_n(x)$ agree at the given values. Here $x_k - x_o = kh$ for $k = 1, 2, \ldots.., n$ and h is the size of all the sub-intervals $[x_k, x_{k+1}]$, $k = 0, 1, 2, 3, \ldots.., n-1$.

Since $P_n(x)$ is a polynomial of degree n, we can write it as

$$P_n(x) = a_0 + a_1 (x-x_0) + a_2 (x-x_0)(x-x_1) + a_3 (x-x_0)(x-x_1)(x-x_2) + \ldots\ldots$$

$$+ a_n (x-x_0)(x-x_1) \ldots\ldots (x-x_{n-1}) \qquad \ldots(1)$$

Putting $x = x_0, x_1, x_2, \ldots\ldots, x_n$ successfully in (1) and using the condition that y and $P_n(x)$ are same at the given set of values, we have

$$a_0 = y_0 \qquad \ldots.(2)$$

$y_1 = a_0 + a_1 (x_1 - x_0) = y_0 + a_1 (x_1 - x_0)$, by (2)

$\Rightarrow \quad a_1 = \dfrac{y_1 - y_0}{x_1 - x_0} = \quad , \qquad (\quad h = x_1 - x_0)$

$y_2 = a_0 + a_1 (x_2 - x_0) + a_2 (x_2 - x_0)(x_2 - x_1)$

$\quad = y_0 + a_1 (2h) + a_2 (2h)(h)$

$\Rightarrow \quad y_2 - y_0 = 2h \qquad + 2a_2 h^2$

$\Rightarrow \quad y_2 - y_0 - 2(y_1 - y_0) = 2a_2 h^2$

$\Rightarrow \quad a_2 = \dfrac{y_2 - 2y_1 + y_0}{2!\,h^2} = \dfrac{\Delta^2 y_0}{2!\,h^2}.$

Similarly, we have $a_3 = \dfrac{\Delta^3 y_0}{3!h^3}, \ldots\ldots, a_n = \dfrac{\Delta^n y_0}{n!h^n}$.

Using these values in (1), we have

$$P_n(x) = y_0 + (x-x_0)\frac{\Delta y_0}{h} + (x-x_0)(x-x_1) \qquad + \ldots\ldots..$$

$$+ (x-x_0)(x-x_1) \ldots\ldots (x-x_{n-1}) \frac{\Delta^n y_0}{n!h^n} \qquad \ldots(3)$$

Taking $x = x_0 + ph$ or $x - x_0 = ph$, we have

$x - x_1 = (x - x_0) - (x_1 - x_0) = ph - p = (p-1)h,$

$x - x_2 = (x - x_1) - (x_2 - x_1) = (p-1)h - h = (p-2)h,$

..... and similarly $x - x_{n-1} = (p-(n-1))h.$

Using these values, (3) becomes

$$y(x) = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \ldots\ldots$$

$$+ \frac{p(p-1)(p-2)\ldots\ldots(p-n+1)}{n!}\Delta^n y_0.$$

This is known as **Newton's Forward interpolation formula**. It is used for interpolating values of $y = f(x)$ near the beginning of the given set of tabulated values and also for extrapolating values of y a little backward of y_0.

14.4 NEWTON'S BACKWARD INTERPOLATION FORMULA

Given a set of $(n+1)$ values : $(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots\ldots, (x_n, y_n)$ of the function $y = f(x)$, it is required to find a polynomial function $P_n(x)$ of nth degree such that y and $P_n(x)$ agree at the given set of values. Since $P_n(x)$ is of degree n, it can be written as

$$P_n(x) = a_0 + a_1(x - x_n) + a_2(x - x_n)(x - x_{n-1}) + \ldots\ldots$$

$$+ a_n(x - x_n)(x - x_{n-1}) \ldots. (x - x_0) \qquad \ldots(1)$$

Putting $x = x_n, x_{n-1}, x_{n-2}, \ldots\ldots, x_1, x_0$ in (1), we get

$$a_0 = y_n \qquad \ldots.(ii)$$

$$y_{n-1} = a_0 + a_1(x_{n-1} - x_n) = y_n - a_1(x_n - x_{n-1})$$

$\Rightarrow$ $$a_1 = \qquad = \frac{\nabla y_n}{h}, \text{ where } h = x_n - x_{n-1}.$$

Similarly, we have $a_2 = \qquad , \ldots\ldots, a_n = \frac{\nabla^n y_n}{n!h^n}$.

Using these values in (1), we obtain

$$P_n(x) = y_n + (x - x_n)\frac{\nabla y_n}{n} + (x - x_n)(x - x_{n-1}) \qquad + \ldots..$$

$$+ (x-x_n)(x-x_{n-1}) \ldots\ldots (x-x_0)\frac{\nabla^n y_n}{n!h^n} \quad \ldots(iii)$$

Taking $x = x_n + ph$, we obtain

$x-x_n = ph,$

$x-x_{n-1} = (x-x_n) + (x_n-x_{n-1}) = ph + h = (p+1)h$

$x-x_{n-2} = (p+2)h, \ldots\ldots, x-x_0 = (p+n-1)h$

Using these values, (iii) becomes

$$y(x) = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \ldots\ldots + \frac{p(p+1)\ldots(p+n-1)}{n!}\nabla^n y_n.$$

This is known as **Newton's Backward interpolation formula**. It is used for interplating values of $y = f(x)$ near the end of the given tabulated values and also used for extrapolating values of y a little backward of y_n.

Note : The process of finding the value of y for some value of x inside the given set of values is called **interpolation** whereas for calculating some value of x outside the interval $[x_0, x_n]$ is called **extrapolation**.

Example 1 : Using Newton's forward interpolation formula, evaluate $\sqrt{12.5}$, given that $\sqrt{11} = 3.316$, $\sqrt{12} = 3.464$, $\sqrt{13} = 3.605$, $\sqrt{14} = 3.741$ and $\sqrt{15} = 3.872$.

Solution : Here $y = f(x) = \sqrt{x}$, $h = 1$, $x_0 = 11$ and $y_0 = \sqrt{11} = 3.316$. Newton's forward interpolation formula is applicable. So we take

$$x_0 + ph = 12.5 \Rightarrow \quad p = 12.5 - 11 = 1.5.$$

First we will construct the Newton's forward difference table.

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
11	**3.316**				
		0.148			
12	3.464		**−.007**		
		0.141		**.002**	
13	3.605		−.005		**−.002**
		0.136		0	
14	3.741		−.005		
		0.131			
15	3.872				

Using Newton's forward interpolation formula, we obtain

$$\sqrt{12.5} = f(x_0 + ph) = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0$$

$$+ \frac{p(p-1)(p-2)(p-3)}{4!}\Delta^4 y_0$$

$$= 3.316 + (1.5)\,(0.148) + \frac{1}{2}(1.5)\,(0.5)\,(-.007)$$

$$+ \frac{1}{6}(1.5)(0.5)(-0.5)\,(.002) + \frac{1}{24}(1.5)(0.5)\,(-0.5)\,(-1.5)(-.002)$$

$$= 3.316 + 0.222 - 0.002625 - .00005 - .000046$$

$$= 3.5352.$$

Note : Making use of the given data from the above example, we can calculate the value of $\sqrt{14.2}$, say, as follows :

Note that $p = 1$, $x_4 = 15$, $y_4 = 3.872$ so that

$14.2 = x_4 + ph \Rightarrow p = 14.2 - 15 = -0.8.$

Transfer the above data to backward difference table and applying the Newton's backward difference formula, we have

$$\sqrt{14.2} = f(14.2) = 3.872 + (-0.8)(0.131) + \frac{1}{2}(-0.8)(0.2)(-0.005)$$

$$+ 0 + \frac{1}{24}(-0.8)(0.2)(1.2)(2.2)(-0.002)$$

$$= 3.7712.$$

Example 2 : Using the following data, find the value of $y = e^x$ when x =0.38 :

x	:	0	0.1	0.2	0.3	0.4
$y = e^x$	:	1	1.1052	1.2214	1.3499	1.4918

Solution : First let us construct the backward difference table for the given data :

x	y	∇y	^2y	^3y	^4y
0	1				
		0.1052			
0.1	1.1052		0.0110		
		0.1162		0.0013	
0.2	1.2214		0.0123		**−0.0002**
		0.1285		**0.0011**	
0.3	1.3499		**0.0134**		
		0.1419			
0.4	**1.4918**				

Here $h = 0.1$ and $x_4 = 0.4$, $y_4 = 1.4918$.

Take $x_4 + ph = 0.38 \Rightarrow p = \frac{0.38 - 0.4}{0.1} = -0.2.$

As $x = 0.38$ is near the end, so we will be applying the Newton's backward interpolation formula.

Using Newton's backward interpolation formula, we have

$$f(0.38) = y_4 + p \quad y_4 + \qquad \nabla^2 y_4 + \qquad \nabla^3 y_4$$

$$+ \qquad \nabla^4 y_4$$

$$= 1.4918 - (0.2)(0.1419) + \quad (-0.2)(0.8)(.0134)$$

$$+ \frac{1}{6}(-0.2)\,(0.8)(1.8)(.0011)$$

$$+ \frac{1}{24}(-0.2)(0.8)(1.8)(2.8)(-.0002)$$

$$= 1.4918 - 0.02838 - 0.001072 - 0.0000528 + .0000067$$

$$= 1.4624.$$

Example 3 : The following table gives values of $y = \sin x$ for $15 \le x \le 40$, (x in degreees) :

$\nabla p(p+1)(p+2)(p+3)$ / $5\ 2\ 2!\ 3!\ 4!$

x	:	15	20	25	30	35	40
y	:	0.2588	0.3420	0.4226	0.5	0.5735	0.6427

Find the values of

(i) sin 22° (ii) sin 38°

Solution : Here $y = \sin x$, $h = 5$, $x_0 = 15$, $y_0 = 0.2588$ and $x_5 = 40$, $y_5 = 0.6427$

(i) **For sin 22°**, take $x_o + ph = 22$

$$\Rightarrow \quad p = \frac{22-15}{5} = \quad = 1.4.$$

The difference table for the given values is

x	y	1st difference	2nd difference	3rd difference	4th difference	5th difference
15	0.2588					
		0.0832				
20	0.3420		−0.0026			
		0.0806		−0.0006		
25	0.4226		−0.0032		−0.0001	
		0.0774		−0.0007		0.0004
30	0.5		−0.0039		0.0003	
		0.0735		−0.0004		
35	0.5735		−0.0043			
		0.0692				
40	0.6427					

Using the Newton's forward interpolation formula, we have

$$f(22) = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0$$

$$+\frac{p(p-1)(p-2)(p-3)}{4!}\frac{2}{5}\Delta^4 y_0 + \frac{p(p-1)(p-2)(p-3)(p-4)}{5!}\Delta^5 y_0$$

$$= .2588 + (1.4)(.0832) + \frac{1}{2}(1.4)\,(.4)\,(-.0026)$$

$$+\frac{1}{6}(1.4)(.4)\,(-.6)\,(-.0006) + \frac{1}{24}(1.4)(.4)(-.6)(-1.6)(-.0001)$$

$$+\frac{1}{120}(1.4)\,(.4)\,(-.6)\,(-1.6)\,(-2.6)\,(.0004)$$

$= 0.2588 + 0.11648 - 0.000728 - 0.00000224 + 0.000336 - 0.00000465$

$= 0.374578.$

(*ii*) As $x = 38$ is near the end, so Newton's Backward interpolation formula is applicable.

Take $x_5 + ph = 38 \Rightarrow \quad p = \frac{38-40}{5} = - \quad = -0.4.$

By Newton's Backward Interpolation formula, we have

$$f(38) = y_5 + p\nabla y_5 + \quad \nabla^2 y_5 + \quad \nabla^3 y_5$$

$$+ \quad \nabla^4 y_5 \quad +$$

$$\nabla^5 y_5$$

$$\Rightarrow \sin 38^o = .6427 - (0.4)(.0692) - \quad (0.4)(0.6)(-.0043)$$

$$-\frac{1}{6}(0.4)(0.6)(1.6)(-.0004) - \frac{1}{24}(0.4)(0.6)(1.6)(2.6)(.0003)$$

$$-\frac{1}{120}(0.4)(0.6)(1.6)(2.6)(3.6)(.0004)$$

$= 0.6427 - 0.02768 + 0.000516 - 0.0000256 - 0.00001248 - 0.0000119808$

$= 0.61553.$

$\frac{p(p+1)(p+2)(p+3)(p+4)}{2\ 2!\ 3!\ 4!\ 5!}$

Example 4 : Find the missing ~~term in the following data~~ :

x	:	0	1	2	3	4
y	:	1	3	9	−	81

Explain why the result differ from $3^3 = 27$?

Solution : Here $h = 1$, $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$, $y_0 = 1$, $y_1 = 3$, $y_2 = 9$, $y_3 =$ missing and $y_4 = 81$.

As four points are given, so the given data can be approximated by a polynomial of degree 3. Hence $\Delta^4 y_0 = 0 \Rightarrow (E-1)^4 y_0 = 0$, $(\because \Delta = E-1)$

$$\Rightarrow \quad (E^4 - 4E^3 + 6E^2 - 4E + 1)\, y_0 = 0$$

$$\Rightarrow \quad E^4 y_0 - 4E^3 y_0 + 6E^2 y_0 - 4E y_0 + y_0 = 0$$

$$\Rightarrow \quad y_4 - 4y_3 + 6y_2 - 4y_1 + y_0 = 0$$

⇒ $81-4y_3+54-12+1=0$

⇒ $y_3 = 31.$

The tabulated function is $y(x) = 3^x$ and the exact value of y_3 is $y(3) = 3^3 = 27$. Here error is due to the fact that the exponential function 3^x is approximated by a polynomial in x of degree 3.

Example 5 : Find the missing terms from the following data :

x	:	0	1	2	3	4	5	6
y	:	5	11	22	40	–	140	–

Solution : Let the missing terms be A and B respectively. Here $h = 1$. As five points are given, so $\Delta^5 y = 0$. Using Newton's forward difference table, we have

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
0	5					
		6				
1	11		5			
		11		2		
2	22		7	$\frac{370}{5}$	A−67	
		18		A−65		370−5A
3	40		A−58		303−4A	
		A−40		238−3A		10A+B−1001
4	A		180−2A		6A+B−698	
		140−A		3A+B−460		
5	140		A+B−280			
		B−140				
6	B					

Now $\Delta^5 y_0$ ⇒ $370-5A = 0$

$10A+B-1001 = 0$

⇒ A = = 74

and B = 1001−10A = 1001−740 = 261.

14.5 SUMMARY

Given a set of x and $y = f(x)$ values where x-values are equally spaced, we can find a polynomial function which satisfies the given values. Further it is being used to interpolate value of y near the beginning (or near the end) by using Newton's interpolation forward (or backward, respectively) formula.

14.6 EXERCISE

1. Derive Newton's forward interpolation formula and apply it to find the value of $\sqrt{21.5}$ using Newton's interpolation formula, given that $\sqrt{20} = 4.472$, $\sqrt{21} = 4.583$, $\sqrt{22} = 4.690$, $\sqrt{23} = 4.796$ and $\sqrt{24} = 4.899$.

2. Derive Newton's backward interpolation formula and apply to it find $\sqrt{155}$, given that $\sqrt{150} = 12.247$, $\sqrt{152} = 12.328$, $\sqrt{154} = 12.409$ and $\sqrt{156} = 12.489$.

3. Construct (*i*) the forward difference table, and

   (*ii*) the backward difference table
   for the function $f(x) = x^3 - 2x^2 + 5x - 8$ for $x = -2, -1, 0, 1, 2, 3, 4$.
   Using these tabular values, also find the values of $f(-1.5)$ and $f(5)$.

4. Using Newton's interpolation formula, find the value of $\sqrt{9.5}$ given that $\sqrt{6} = 2.449$, $\sqrt{7} = 2.646$, $\sqrt{8} = 2.828$, $\sqrt{9} = 3$, $\sqrt{10} = 3.162$ and $\sqrt{11} = 3.316$.

5. Compute the value of sin 38° correct to seven decimal places, given that sin 15° = 0.258819, sin 20° = 0.3420201, sin 25° = 0.4226183, sin 30° = 0.5, sin 35° = 0.5735764 and sin 40° = 0.6427876.

6. The following table gives the values *of* $f(x) = \tan x$ for x lying within 0.10 and 0.30 :

x	:	0.10	0.15	0.20	0.25	0.30
y	:	0.1003	0.1511	0.2027	0.2553	0.3093

   Find the values of (*i*) tan 0.12 (*ii*) tan 0.26

7. In the following table, the values of y are consecutive terms of a series of which the number 29 is the 4th term. Find the 1st and the 9th term of the series. Also find the polynomial function which approximates these values :

x	:	2	3	**4**	5	6	7	8
y	:	7	16	**29**	46	67	92	121

8. If $f(1.15) = 1.0723$, $f(1.20) = 1.0954$, $f(1.25) = 1.1180$ and $f(1.30) = 1.1401$, then find $f(1.28)$.

9. The area A of a circle of diameter d is given by the following values :

d	:	80	85	90	95	100
A	:	5026	5674	6382	7088	7854

Calculate the area of a circle of diameter 105.

10. The following data gives the distances in nautical miles of the visible horizon for the given heights in feet above the earth's surface. Find the value y when $x = 410$ ft;

x (height)	:	100	150	200	250	300	350	400
y (distance)	:	10.63	13.03	15.04	16.81	18.42	19.90	21.27

11. Find the missing terms from the following data :

(*i*)	x	:	1	2	3	4	5	
	y	:	2	5	7	....	32	
(*ii*)	x	:	0	5	10	15	20	25
	y	:	6	10	....	17	....	31

ANSWERS

1. 4.6368
2. 12.449
3. −23.375 ; 92
4. 3.0822
5. 0.6156614
6. (*i*) 0.1205 (*ii*) 0.2662
7. 2; 154; $2x^2 - x + 1$
8. 1.1312

9. 8666 10. 21.53 nautical miles

11. (*i*) 14 (*ii*) 13.25; 22.5

14.7 SUGGESTED READING/REFERENCES

1. "Introductory Methods of Numerical Analysis" by S.S.Sastry, PHI Learning Pvt Ltd., New Delhi.

2. "Numerical Methods in Engineering and Science" by B.S.Grewal, Khanna Publishers, New Delhi.

3. "A text book of Complex Variables & Numerical Methods" by Bhopinder Singh & Others, Kirti Publishers, Jammu.

14.8 MODEL TEST PAPER

1. Derive Newton's forward interpolation formula and apply it to evaluate $\sqrt{4.8}$, using the following data :

   $\sqrt{4} = 2$, $\sqrt{5} = 2.236$, $\sqrt{6} = 2.449$, $\sqrt{7} = 2.646$ and $\sqrt{8} = 2.828$.

2. Derive Newton's backward interpolation formula and apply to it find $\sqrt{9.5}$ given that $\sqrt{6} = 2.449$, $\sqrt{7} = 2.646$, $\sqrt{8} = 2.828$, $\sqrt{9} = 3$, $\sqrt{10} = 3.162$ and $\sqrt{11} = 3.316$.

3. Using Newton's interpolation formula, calculate $\sqrt{8.5}$, given that $\sqrt{7} = 2.646$,

   $\sqrt{8} = 2.828$, $\sqrt{9} = 3$, $\sqrt{10} = 3.162$ and $\sqrt{11} = 3.316$.

4. Apply Newton's forward interpolation formula to find the value of *f* (1.6), given that

x	:	1	1.4	1.8	2.2
f(*x*)	:	3.49	4.82	5.96	6.5

5. Find the 6th term and the general term of series :

   3, 9, 20, 38, 65, ...

6. Find the value of sin 52^o by using Newton's interpolation formula, given that sin $45^o = 0.7071$, sin $50^o = 0.766$, sin $55^o = 0.8192$, sin $60^o = 0.866$.

7. The following table gives the values of $f(x) = x^4$ for x lying within 2.1 and 2.6.

x	:	2.1	2.2	2.3	2.4	2.5	2.6
y	:	19.4481	23.4256	27.9841	33.1776	39.0625	45.6976

Using interpolation techniques, determine the values of $(2.34)^4$

8. From the following table, find the number of students who got marks between 40 and 45 :

Marks obtained	30-40	40-50	50-60	60-70	70-80
No. of students	31	42	51	35	31

9. Find the missing term from the following data :

x	:	−2	−1	0	1	2
y	:	6	0	...	0	6

10. Find the missing terms from the following data :

(*i*)	x	:	0	1	2	3	4		
	y	:	1	2	6	...	51		
(*ii*)	x	:	2	2.1	2.2	2.3	2.4	2.5	2.6
	y	:	0.135	...	0.111	0.1	....	0.082	0.074

NUMERICAL ANALYSIS

LAGRANGE’S FORMULA FOR INTERPOLATION AND INVERSE INTERPOLATION

By: Dr. Kamaljeet Kour

15.1 INTRODUCTION

In lession 14, we studied the finite differences for the function $y = f(x)$ when the arguments (*i.e.* x-values) are given at equal intervals. In numerical techniques, mostly the scientific data is not in equal intervals but is in unequal intervals. Such problems can be solved easily by Lagrange's interpolation formula which is described below.

15.2 OBJECTIVES

The objective is to make students capable of solving numerical problems arising in a scientific experiment with unequal arguments.

The advantage of this formula is that it is simple and easy to apply. Also there is no need to construct the difference table and we can interpolate the unknown value with the help of the given points (*i.e.* the values of x and y).

15.3 LAGRANGE'S FORMULA FOR UNEQUAL INTERVALS

Let $y = f(x)$ be a continuous function defined in the interval I = $[a, b]$ and take $(n+1)$ distinct points $x_0, x_1, x_2, \ldots, x_n$ on I which may or may not be equally spaced. Suppose the corresponding values of the function are $y_0, y_1, y_2, \ldots, y_n$. Then a polynomial $P_n(x)$ of degree $n \leq 1$ can be found which interpolates $f(x)$ such that $y_k = P_n(x_k)$ for $k = 0, 1, 2, \ldots, n$ and is given by

$$P_n(x) = \frac{(x-x_1)(x-x_2)......(x-x_n)}{(x_0-x_1)(x_0-x_2)......(x_0-x_n)} \cdot y_0 + \qquad y_1$$

$$+ .... + \qquad \cdot y_n.$$

To derive this, we can write $P_n(x)$ as a polynomial in x of degree n :

$$P_n(x) = a_0\,(x-x_1)\,(x-x_2)\,.....\,(x-x_n) + a_1\,(x-x_0)\,(x-x_2)\,.....\,(x-x_n)$$
$$+ a_2\,(x-x_0)\,(x-x_1)(x-x_3)\,.....\,(x-x_n) + ...$$
$$+ a_n\,(x-x_0)\,(x-x_1)\,.....\,(x-x_{n-1}) \qquad ...(1)$$

Putting $x = x_0, x_1, x_2, ..., x_n$ successively in (1) and using the condition that y and $P_n(x)$ are same at the given set of values, we have

$$y_0 = a_0\,(x_0-x_1)\,(x_0-x_2)\,....\,(x_0-x_n)$$

$$\Rightarrow \qquad a_0 = \frac{y_0}{(x_0-x_1)(x_0-x_2)......(x_0-x_n)};$$

$$\frac{(x-x_0)(x-x_1)......(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)......(x_n-x_{n-1})}$$

$$y_1 = a_1\,(x_1-x_0)\,(x_1-x_2)\,....\,(x_1-x_n)$$

$$\Rightarrow \qquad a_1 = \frac{y_1}{(x_1-x_0)(x_1-x_2)......(x_1-x_n)};$$

and similarly, we have

$$a_2 = \frac{y_2}{(x_2-x_0)(x_2-x_1)......(x_2-x_n)} \quad , ..., \text{ and}$$

$$a_n = \frac{y_n}{(x_n-x_0)(x_n-x_1)......(x_n-x_{n-1})}.$$

Put the values of a_0, a_1, a_2 ,......, a_n in (1), we have the required **Lagrenge's interpolation formula** as :

$$f(x) = \frac{(x-x_1)(x-x_2)......(x-x_n)}{(x_0-x_1)(x_0-x_2)......(x_0-x_n)} \cdot y_0 + \qquad y_1$$

$$+ \ldots\ldots + \frac{(x-x_0)(x-x_1)\ldots\ldots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\ldots\ldots(x_n-x_{n-1})} \cdot y_n.$$

Remark : Lagrange's interpolation formula is more general and can be applied to equal as well as unequal spaced arguments.

Example 1 : Using Lagrange's formula, find $f(x)$ from the following data :

x	:	0	1	4	5
y	:	4	3	24	39

Solution : Here $x_0 = 0,\ x_1 = 1,\ x_2 = 4,\ x_3 = 5$

$y_0 = 4,\ y_1 = 3,\ y_2 = 24,\ y_3 = 39.$

By Lagrange's interpolation formula, we have

$$f(x) = \frac{(x-1)(x-4)(x-5)}{(0-1)(0-4)(0-5)}(4) + \frac{(x-0)(x-4)(x-5)}{(1-0)(1-4)(1-5)}(3)$$

$$+ \frac{(x-0)(x-1)(x-5)}{(4-0)(4-1)(4-5)}(24) + \frac{(x-0)(x-1)(x-4)}{(5-0)(5-1)(5-4)}(39)$$

$$= \frac{4(x-1)(x-4)(x-5)}{(-1)(-4)(-5)} + \frac{3x(x-4)(x-5)}{1(-3)(-4)}$$

$$+ \frac{24x(x-1)(x-5)}{4.3.(-1)} + \frac{39x(x-1)(x-4)}{5.4.1}$$

$$= -\frac{1}{5}(x-1)(x^2-9x+20) + \frac{1}{4}x(x^2-9x+20)$$

$$-2x(x^2-6x+5) + \frac{39}{20}x(x^2-5x+4)$$

$$= -\frac{1}{5}(x^3-10x^2+29x-20) + \frac{1}{4}(x^3-9x+20x)$$

$$-2x\,(x^3-6x^2+5x)+\frac{39}{20}\,(x^3-5x^2+4x)$$

$$=\left(-\frac{1}{5}+\frac{1}{4}-2+\frac{39}{20}\right)x^2+\left(2-\frac{9}{4}+12-\frac{39}{4}\right)x^2+\left(-\frac{29}{5}+5-10+\frac{39}{5}\right)x+4$$

$$=0.x^3+2x^2-3x+4=2x^2-3x+4$$

as the required polynomial function.

Example 2 : Using a suitable interpolation formula, find the polynomial function for the following data. Also find $f\,(2)$.

x	:	-1	0	1	3
y	:	2	1	0	-1

Solution : As spacing is unequal, so we can apply Lagrange's interpolation formula.

Here $x_0=-1,\ x_1=0,\ x_2=1,\ x_3=3,$

$y_0=2,\ y_1=1,\ y_2=0$ and $y_3=-1.$

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

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

$$=\frac{2x(x^2-4x+3)}{(-1)(-2)(-4)}+\frac{(x+1)(x^2-4x+3)}{(-1)(-3)}-\frac{x(x^2-1)}{4.3.2}$$

$$=\left(-\frac{x}{4}+\frac{x+1}{3}\right)(x^2-4x+3)-\frac{1}{24}(x^3-x)$$

$$=\frac{(x+4)}{12}(x^2-4x+3)-\frac{1}{24}(x^3-x)$$

$$=\frac{1}{24}[2x^3+8x^2-8x^2-32x+6x+24-(x^3-x)]$$

$$= \frac{1}{24}[x^3 - 25x + 24] \qquad ...(1)$$

which is required polynomial in x of degree 3.

Using (1), we can directly check that

$$f(-1) = 2, f(0) = 1, \ f(1) = 0 \ \text{ and } \ f(3) = -1.$$

$$\text{Now } f(2) = \frac{1}{24}[8-50+24]$$

$$= -\frac{18}{24} = -\frac{3}{4}.$$

Example 3 : Find the equation of cubic curve which passes through the points (4, −43) , (7, 83), (9, 327) and (12, 1053).

Solution : Given values (x_k, y_k) for $k = 0, 1, 2, 3$ are :

$x_0 = 4, \ y_0 = -43, \ x_1 = 7, \ y_1 = 83, \ x_2 = 9, \ y_2 = 327$ and $x_3 = 12, \ y_3 = 1053.$

Using Lagrange's interpolation formula, we have

$$f(x) = \frac{(x-x_1)(x-x_2)(x-x_3).y_0}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}$$

$$+ \qquad\qquad +$$

$$= \qquad\qquad + \frac{(x-4)(x-9)(x-12).83}{(7-4)(7-9)(7-12)}$$

$$+ \frac{(x-4)(x-7)(x-12).327}{(9-4)(9-7)(9-12)} + \frac{(x-4)(x-7)(x-9).1053}{(12-4)(12-7)(12-9)}$$

$$= -\frac{43(x-7)(x-9)(x-12)}{(-3)(-5)(-8)} + \frac{83(x-4)(x-9)(x-12)}{3(-2)(-5)}$$

$$+\frac{327(x-4)(x-7)(x-12)}{5.2.(-3)}+\frac{1053(x-4)(x-7)(x-9)}{8.5.3}$$

$$=\frac{43}{120}(x-7)(x-9)(x-12)+\frac{83}{30}(x-4)(x-9)(x-12)$$

$$-\frac{109}{10}(x-4)(x-7)(x-12)+\frac{351}{40}(x-4)(x-7)(x-9)$$

$$=\frac{1}{120}(x-9)(x-12)\,[43\,(x-7)+332\,(x-4)]$$

$$+\frac{1}{40}(x-4)(x-7)[\,351\,(x-9)-436(x-12)]$$

$$=\frac{1}{120}(x-9)(x-12)\,[375x-1629]+\frac{1}{40}(x-4)(x-7)[-85x+2073]$$

$$=\frac{3}{120}(125x-543).(x^2-21x+108)-\frac{1}{40}(85x-2073).(x^2-11x+28)$$

$$=\frac{1}{40}(125x^3-3168x^2+24903x-58644)$$

$$-\frac{1}{40}(85x^3-3008x^2+25183x-58044)$$

$$=\frac{1}{40}(40x^3-160x^2-280x-600)$$

$$=x^3-4x^2-7x-15,$$

which is the required cubic polynomial satisfying the tabulated values.

Example 4 : Apply Lagrange's formula to find the value of $\log_{10} 656$, given that $\log_{10} 654 = 2.8156$, $\log_{10} 658 = 2.8182$, $\log_{10} 659 = 2.8189$ and $\log_{10} 661 = 2.8202$.

Solution : Here $y = f(x) = \log_{10} x$. The given data can be written as :

x	:	654	658	659	661
$f(x)$	:	2.8156	2.8182	2.8189	2.8202

Apply Lagrauge's interpolation formula, we have

$$\log_{10} 656 = f(656) = \frac{(656-658)(656-659)(656-661)(2.8156)}{(654-658)(654-659)(654-661)}$$

+

+

+

= +

+ $+\frac{\ldots(656-661)(2.8182)}{14 \quad \ldots(658-661)}$

$$= \quad (2.8156) + \frac{5}{2}(2.8182) - 2\,(2.8189) + \frac{2}{7}(2.8202)$$

= 0.6033 + 7.0455 − 5.6378 + 0.8057

= 2.8167.

Example 5 : Express the function

$$\frac{x^2+4x+2}{(x^2-1)(x+2(x+3)}$$

as a sum of partial fractions using Lagrange's interpolation formula.

Solution : Let $\dfrac{x^2+4x+2}{(x^2-1)(x+2(x-3)} = \dfrac{f(x)}{g(x)}$, where $f(x) = x^2 + 4x + 2$ and $g(x) = (x^2 - 1)(x + 2)(x - 3)$.

Zeros of $g(x)$ or roots of the equation $g(x) = 0$ are $x = -2, -1, 1, 3$. Now we will write the value of $f(x)$ at these four points :

$f(-2) = -2, f(-1) = -1, f(1) = 7$ and $f(3) = 23$.

So we have the following data :

x	:	-2	-1	1	3
y	:	-2	-1	7	23

Applying Lagrange's interplation formula to this data, we have

$$f(x) = \frac{(x+1)(x-1)(x-3)}{(-2+1)(-2-1)(-2-3)}.(-2) + \frac{(x+2)(x-1)(x-3)}{(-1+2)(-1-1)(-1-3)}.(-1)$$

$$+ \frac{(x+2)(x+1)(x-3)}{(1+2)(1+1)(1-3)}.(7) + \frac{(x+2)(x+1)(x-1)}{(3+2)(3+1)(3-1)}.(23)$$

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

[illegible]

$$+ \frac{7(x+2)(x+1)(x-3)}{(3)(2)(-2)} + \frac{23(x+2)(x+1)(x-1)}{(5)(4)(2)}$$

$$= \frac{2}{15}(x+1)(x-1)(x-3) - \frac{1}{8}(x+2)(x-1)(x-3)$$

$$- \frac{7}{12}(x+2)(x+1)(x-3) + \frac{23}{40}(x+2)(x+1)(x-1)$$

Now dividing on both sides by $g(x)$, we obtain

$\frac{f(x)}{g(x)} = \frac{2}{15} \cdot \frac{1}{x+2} - \quad \frac{1}{x+1} - \quad \cdot \frac{1}{x-1} + \quad \cdot \frac{1}{x-3}$, which expresses the given as a sum of partial fraction.

15.4. INVERSE INTERPOLATION

In the inverse interpolation, for a given value of y, the corresponding value of x is to be determined. It is similar to the direct interpolation in which only the roles of x and y are interchanged.

So by interchanging the roles of x and y, we have the Lagrange's inverse interpolation formula for a given set of $(n + 1)$ tabulated values as :

$$x = \quad . x_0 + \quad . x_1$$

$$+ \ldots . + \quad . x_n.$$

Example 6 : Find the value of x when $y = 15$ using the following data :

x	:	5	6	9	11
y	:	12	13	14	16

Solution : Here $x_0 = 5$, $x_1 = 6$, $x_2 = 9$, $x_3 = 11$, $y_0 = 12$, $y_1 = 13$, $y_2 = 14$ and $y_3 = 16$.

Using inverse interpolation of Lagrange's formula, we have

$$x\,(15) = \frac{(15-y_1)(15-y_2)(15-y_3)}{(12-13)(12-14)(12-16)} \cdot x_0$$

$$+ \frac{(15-y_0)(15-y_2)(15-y_3)}{(13-12)(13-14)(13-16)} \cdot x_1$$

$$+ \frac{(15-y_0)(15-y_1)(15-y_3)}{(14-12)(14-13)(14-16)} \cdot x_2$$

$$+ \frac{(15-y_0)(15-y_1)(15-y_2)}{(16-12)(16-13)(16-14)} \cdot x_3$$

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

$$= \frac{5}{4} - 6 + \frac{27}{2} + \frac{11}{4} = \frac{23}{2} = 11.5.$$

Observation : From the above example, we observe that in the inverse interpolation the value of x so obtained corresponding to a given value of y may or may not be within the tabulated values.

Example 7 : The following table gives the values of x and y :

x	:	1.2	2.1	2.8	4.1	4.9	6.2
y	:	4.2	6.8	9.8	13.4	15.5	19.6

Find the value of x corresponding to $y = 12$ using Lagrange's techniques.

Solution : Here $x_0 = 1.2,\ x_1 = 2.1,\ x_2 = 2.8,\ x_3 = 4.1,\ x_4 = 4.9,\ x_5 = 6.2,$ $y_0 = 4.2$, $y_1 = 6.8$, $y_2 = 9.8$, $y_3 = 13.4$, $y_4 = 15.5$ and $y_5 = 19.6$.

Spacing is unequal and we have to calculate value of x for $y = 12$. For this, we will apply Lagrange's inverse interpolation formula as follows :

$$x = \frac{(y-y_1)(y-y_2)(y-y_3)(y-y_4)(y-y_5)}{(y_0-y_1)(y_0-y_2)(y_0-y_3)(y_0-y_4)(y_0-y_5)} \ldots$$ [illegible]

$$+ \qquad . x_1$$

$$+ \ldots\ldots\ldots + \qquad . x_5$$

$$\Rightarrow x(12) =$$

$$+$$

$$+$$

$+$

$+$

$+$

$=$

$+$

$+$

$+$ [illegible]

$+$

$+$

$=$

$$+\ \frac{9731.64192}{2846.06784} - \frac{4652.199552}{4824.75987} + \frac{2710.86816}{49105.74592}$$

$$= 0.0219315121 - 0.2340876 + 1.2518518518$$

$$+ 3.419328866 - 0.9642344235 + 0.0552047.$$

$= 3.5425347715.$

15.5 SUMMARY

In this lesson, we interpolate the given values of x and y by a polynomial function which satisfies the given values. Using Lagrange's interplation formula, we express a rational function as a sum of partial fractions. Inverse interpolation is also discussed with solved examples.

15.6 EXERCISE

1. Derive Lagrange's interpolation formula and apply it to find the value of $f(4)$ for the following data :

x	:	0	2	3	6
$f(x)$	:	−4	2	14	158

2. Using Lagrange's interpolation formula ; compute $f(7)$, given that

x	:	5	6	9	11
$f(x)$	:	12	13	14	16

3. Find the cubic polynomial in x using the data :

x	:	0	1	2	4
$f(x)$	:	1	1	2	5

4. State Lagrange's interpolation formula and use it to find value of y at $x =$ 5 from the following data :

x	:	0	1	3	8
y	:	1	3	13	123

5. Evaluate $f(9)$ using Lagrange's formula for the given data :

x	:	5	7	11	13	17
$f(x)$	:	150	392	1452	2366	5202

6. Find the cubic polynomial which takes the following values :

x	:	1	3	5	7
$f(x)$	:	24	120	336	720

Hence or otherwise calculate $f(8)$.

7. Using Lagrange's formula, express the following functions as a sum of partial fractions :

(*i*) $\dfrac{x^2+6x-1}{(x^2-1)(x-4)(x-6)}$ (*ii*) $\dfrac{x^2+5x-2}{(x^2-4)(x-3)(x-4)}$

8. Find the value of x when $y = 7$ from the following data :

x	:	1	3	4
$f(x)$	:	4	12	19

9. The following data gives the values of x and y :

x	:	0	1	4	5
y	:	2	3	66	127

$$\frac{3}{2}\cdot\frac{1}{x-2}+\frac{1}{15}\cdot\frac{1}{x+2}-\frac{22}{5}\cdot\frac{1}{x-3}+\frac{17}{6}\cdot\frac{1}{x-4}$$

Find the value of x when $y = 29$ using Lagrange's techniques.

ANSWERS

1. 40
2. 13.466
3. $\dfrac{1}{12}(12-8x+9x^2\ x^3)$
4. 38.14
5. 810
6. $x^3+6x^2+11x+6$; 990
7. (*i*) $\dfrac{1}{5}\cdot\dfrac{1}{x-1}+\dfrac{3}{35}\cdot\dfrac{1}{x+1}-\dfrac{13}{10}\cdot\dfrac{1}{x-4}+\dfrac{71}{70}\cdot\dfrac{1}{x-6}$

(*ii*)

8. 1.86

9. 15.43

15.7 SUGGESTED READING/REFERENCES

1. "Introductory Methods of Numerical Analysis" by S.S.Sastry, PHI Learning Pvt Ltd., New Delhi.

2. "Numerical Methods in Engineering and Science" by B.S.Grewal, Khanna Publishers, New Delhi.

3. "A text book of Complex Variables & Numerical Methods" by Bhopinder Singh & Others, Kirti Publishers, Jammu.

15.8 MODEL TEST PAPER

1. Explain Lagrange's interpolation formula and use it to find the polynomial $y = f(x)$, given that

x	:	0	1	2	4	5
y	:	2	3	12	78	147

2. A cubic polynomial $y = f(x)$ satisfies the following data :

x	:	0	1	3	4
$f(x)$	:	1	4	40	85

Determine the function $f(x)$ and hence $f''(2)$.

3. Using suitable interpolation formula, evaluate $f(4)$ given that

x	:	0	2	3	6
y	:	−4	2	14	158

4. Using Lagrange's formula, find the polynomial in x given that

x	:	0	2	3	6
y	:	659	705	729	804

5. Using the data of Q.4, find the value of $f(4)$.

6. Using Lagrange's formula, express the following functions as a sum of partial fractions :

(*i*) (*ii*) $\dfrac{2x^2-3x+5}{(x-1)(x+2)(x+4)}$

7. The following data gives the values of x and y :

x	:	0	1	3	8
y	:	1	3	13	123

Find the value of x when y = 29 using Lagrange's techniques.

$$\frac{x^2+x-3}{x^3-2x^2-x+2}$$

B.A. III Semester	Unit-5
Course No. : UGM 0301	Lesson 16

MATRICES

MATRICES AND THEIR TYPES

By: Dr. Kamaljeet Kour

16.1 INTRODUCTION

The concept on matrices appeared during the investigation of various problems on linear algebra which also includes the solution of system of linear equations. The study of matrices has many applications in various areas of science and technology.

By a matrix, we mean a system of numbers (or sometimes functions or operators) arranged in a rectangular array in rows and columns and bounded by the square brackets. If there are m rows and n columns, then the matrix having elements a_{ij} (where i = 1,2,.....,m ; j = 1,2,.....n) is said to be of order m x n. It is denoted by $A = [a_{ij}]_{m \times n}$ or simply by $A = [a_{ij}]$ or sometimes by $A_{m \times n}$.

16.2 OBJECTIVES :

The objective of this lesson is to study various types of matrices alongwith their results.

16.3 OPERATION ON MATRICES :

16.3.1. ADDITION OF MATRICES :

Two matrices A and B of the same order can be added and their sum A + B is given by the addition of the corresponding elements of A and B.

16.3.2. SCALAR MULTIPLE OF A MATRIX:

If k is any scalar and A is a matrix, then the matrix kA is defined as the matrix obtained by multiply each element of A by k. Similarly, we have $-$A, which is (-1) A.

16.3.3. MULTIPLICATION OR PRODUCT OF MATRICES:

Two matrices A and B can be multiplied to give product AB if the number of columns of A is equal to the number of rows of B. If A = $[a_{ij}]$ is a matrix of order m x n and B = $[b_{ij}]$ is a matrix of order n x p, then the product AB is the matrix C = $[c_{ij}]$ of order m x p, where

$$c_{ij} = a_{i1}\, b_{1j} + a_{i2}\, b_{2j} + a_{i3}\, b_{3j} + \ldots.. + a_{in}\, b_{nj}.$$

16.3.4 DEFINITIONS :

(1) Equal matrices :

Two matrices are said to **equal** if

(*i*) they are of the same order and

(*ii*) have same elements in the corresponding positions.

(2) Determinant of a matrix :

The determinant |A| of a square matrix A is the sum of products of any row with its corresponding cofactors.

For example,

(*i*) if A = $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then | A | = $ad - bc$.

(*ii*) if A = $\begin{bmatrix} a & b & c \\ l & m & n \\ p & q & r \end{bmatrix}$, then |A| = $a \begin{vmatrix} m & n \\ q & r \end{vmatrix} - b \begin{vmatrix} l & n \\ p & r \end{vmatrix} + c \begin{vmatrix} l & m \\ p & q \end{vmatrix}$

$= a\,(mr - qn) - b\,(lr - pn) + c\,(lq - pm)$.

(3) Singular and Non-singular matrices :

A square matrix A is called **singular** if | A | = 0. If | A | ≠ 0, then A is called **non-singular**.

For example, A = $\begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}$ is a singular matrix since |A| = 0, whereas A

$$= \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 3 & 1 & 1 \end{bmatrix}$$ is non-singular because its determinant is $|A| = 2 \neq 0$.

16.4 TYPES OF MATRICES :

The following are various types of matrices with important results and examples :

16.4.1. Row matrix is a matrix having only one row and any number of columns. For example, A = [2 4 3 5 1 8] is a row matrix of order 1 x 6.

16.4.2. Column matrix is a matrix having one column and any number of rows.

For example, $B = \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$ is a column matrix of order 3 x 1.

16.4.3. Square matrix is a matrix with number of rows equal to number of columns.

For example, $C = \begin{bmatrix} 1 & -2 \\ -1 & 6 \end{bmatrix}$ is a square matrix of order 2. If $A = [a_{ij}]$ is a matrix of order m x m, then the elements a_{11}, a_{22}, a_{33}, ..., a_{mm} are called **principal diagonal** elements of A and sum of these elements is called **trace** of A.

16.4.4 Diagonal matrix is a square matrix having all of its non-diagonal entries zero.

For example, $D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$ is a diagonal matrix.

16.4.5. Scalar matrix is a diagonal matrix in which all diagonal elements are equal to a same scalar.

For example, $A = \begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{bmatrix}$ is a scalar matrix.

16.4.6. Unit (or identity) matrix is a scalar matrix with all of its diagonal entries equal to unity.

For example, $I_2=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ is an identity matrix of order 2. A unit matrix of order n is denoted by I_n.

16.4.7. Zero (or null) matrix is a matrix having all of its entries equal to zero. It can be of any order.

For example $\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$, $\begin{bmatrix}0 & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}$ and $\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$ are zero matrices. It is denoted by O.

16.4.8. A square matrix all of whose elements below the principal diagonal are zero is called an **upper triangular matrix**. Similarly we have **lower triangular matrix**. For example,

$$\begin{bmatrix}1 & 2 & 3\\ 0 & 5 & 4\\ 0 & 0 & 8\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}3 & 0 & 0\\ 3 & 4 & 0\\ 1 & 5 & 6\end{bmatrix}$$

are upper and lower triangular matrices respectively.

16.4.9. The **transpose** of a matrix is obtained by interchanging the rows and the corresponding columns of a given matrix. The transpose of the matrix A_{mxn} is denoted by A'_{nxm}. If A is the matrix of order m x n, then the order of A′ will be n x m.

For example,

$$\text{let} \quad A = \begin{bmatrix}3 & 5\\ 4 & 3\\ 7 & 8\end{bmatrix}. \text{ Then } A' = \begin{bmatrix}3 & 4 & 7\\ 5 & 3 & 8\end{bmatrix}$$

Proposition 1 : If A, B are two matrices of same order and m, n are scalars, then

(*i*) $(mA + nB)' = mA' + nB'$

(*ii*) $(AB)' = B'A'$

(*iii*) $(A')' = A$.

16.4.10. Symmetric matrix : A square matrix A is called symmetric if $A' = A$, that is, $a_{ij} = a_{ji}$ for all i, j. For example,

(*i*) $A = \begin{bmatrix} -2 & 3 \\ 3 & 1 \end{bmatrix}$ and

(*ii*) $B = \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$ are symmetric matrices.

16.4.11. Skew symmetric matrix : A square matrix A is called skew-symmetric

if $A' = -A$, that is, $a_{ij} = -a_{ji}$ for all i, j.

Now $a_{ij} = -a_{ji} \Rightarrow a_{ii} = -a_{ii} \; \forall i$

$\Rightarrow$ $2a_{ii} = 0$ or $a_{ii} = 0 \; \forall i$. Thus diagonal entries are zero in a skew symmetric matrix.

For example (*i*) $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and (*ii*) $B = \begin{bmatrix} 0 & -h & -g \\ h & 0 & -f \\ g & f & 0 \end{bmatrix}$ are skew symmetric matrices.

Proposition 2 : Let A and B be two symmetric matrices of same order and m be a scalar. Then mA, $A + B$, $A - B$ and $A + A'$ are all symmetric matrices.

Proof : Let A and B be two square matrices of same order which are symmetric. Then by definition,

$$A' = A \text{ and } B' = B \qquad ...(1)$$

Also the matrices A, A′, B, B′, mA, A + B, A − B and AA′ are all of same order. Now

(*i*) $(mA)' = mA' = mA$ $\qquad (\because A' = A)$

$\Rightarrow$ mA is a symmetric matrix.

(*ii*) $(A + B)' = A' + B' = A + B$, by using (1)

$\Rightarrow$ $A + B$ is a symmetric matrix.

(*iii*) $(A - B)' = A' - B' = A - B$, by using (1)

$\Rightarrow$ $A - B$ is a symmetric matrix.

(*iv*) $(A+A')' = (A+A)'$ $\quad$ ($A' = A$)

$= (2A)' = 2A' = A' + A' = A+A'$

$\Rightarrow$ $A+A'$ is a symmetric matrix.

Theorem 3 : The necessary and sufficient conditions for a matrix A to be symmetric is that $A' = A$.

OR

Prove that a square matrix is symmetric if and only if $A' = A$.

Proof : Let the given matrix be $A = (a_{ij})$, where $i, j = 1, 2, ..., n$.

Necessary Part : Suppose A is symmetric. Then by definition, we have

$$a_{ij} = a_{ji} \text{ for all } i, j \qquad ....(1)$$

We will prove that $A' = A$, that is, they have same element in respective positions.

Now both A and A' are of same order n x n.

Also (i, j) the element of $A' = (j, i)$th element of A

$= a_{ji} = a_{ij}$, using (1)

$= (i, j)$th element of A.

Thus A and A' are matrices of same order and have same elements. So $A = A'$.

Converse Part : Conversely, suppose that $A' = A$. Then we have to prove that A is symmetric, that is, $a_{ij} = a_{ji}$ for all i, j.

Now $a_{ij} = (i, j)$th element of A.

$= (i, j)$th element of A' ($\because$ $A' = A$)

$= (j, i)$th element of A

$= a_{ji}$ for all i, j.

$\Rightarrow$ A is symmetric.

Thus we have proved that A is symmetric iff $A' = A$.

Theorem 4 : The necessary and sufficient conditions for a matrix A to be skew-symmetric is that $A' = -A$.

OR

Prove that a square matrix A is skew-symmetric if and only if $A' = -A$.

Proof : It is similar to the proof of Theorem 3.

Theorem 5 : (*i*) Every square matrix can be written as sum of a symmetric and a skew-symmetric matrix.

(*ii*) Every square matrix can be uniquely expressed as a sum of a symmetric and a skew-symmetric matrix.

Proof : (*i*) Let A be a square matrix. Then we can write it as

$$A = \frac{1}{2}(2A) = \frac{1}{2}(A + A) = \frac{1}{2}(A + A' - A' + A)$$

$$= \frac{1}{2}(A + A') + \frac{1}{2}(A - A') = B + C \text{ (say)}.$$

Here $B = \frac{1}{2}(A + A')$ and $C = \frac{1}{2}(A - A')$.

First note that A, A′, A + A′ and A − A′ are all matrices of same order. Hence A, B,C are also of same order.

Now $B' = \frac{1}{2}(A + A')' = \frac{1}{2}(A' + (A')') = \frac{1}{2}(A' + A)$

$$= \frac{1}{2}(A + A') = B$$

$\Rightarrow$ B is symmetric.

Similarly $C' = \frac{1}{2}(A - A')' = \frac{1}{2}(A' - A) = -\frac{1}{2}(A - A') = -C$

$\Rightarrow$ C is skew symmetric.

Thus any matrix A can be written as the sum B + C, where B is symmetric and C is skew-symmetric.

(*ii*) Let A be a square matrix. Then as proved in part (*i*) above, we have

$$A = B + C \qquad ...(1)$$

where $B = \frac{1}{2}(A + A')$ is a symmetric matrix and $C = \frac{1}{2}(A - A')$ is a skew-symmetric matrix. Next, we have to show that the representation in (1) is unique. For this, if possible, suppose there is another representation of A given by

$$A = P + Q, \qquad ...(2)$$

where P is symmetric and Q is skew-symmetric.

Now $A' = (P + Q)' = P' + Q' = P - Q$ ...(3)

since $P' = P$ and $Q' = -Q$.

Solving (2) and (3) for P and Q, we obtain

$$P = \frac{1}{2}(A + A') \text{ and } Q = \frac{1}{2}(A - A')$$

$\Rightarrow$ P = B and Q = C.

Hence the representation (2) is same as (1), proving that this representation (1) is unique.

Example 6. If A is a symmetric matrix of order *n* and B is a matrix of order *n*, then show that $B'AB$ is also a symmetric matrix of order *n*.

Solution : Let A and B be two matrices of order *n* in which A is symmetric. Then $A' = A$. Also A, B and B' are of order *n*, so the product $B'AB$ is also of order *n*.

Let $P = B'AB$. To show that $B'AB$ is symmetric, we have to verify that $P' = P$.

Now $P' = (B'AB)' = (AB)'(B')'$, $(\because (XY)' = Y'X')$

$= (B'A')B$, ($(B') = B$)

$= B'A'B$, by associative law

$= B'AB$, ($A' = A$)

$= P$.

Thus $B'AB$ is symmetric when A is symmetric.

16.4.12. Definitions :

(1) Let $A = [a_{ij}]$ be a matrix. Then the **conjugate** of A, denoted by $\bar{A}$, is defined as $\bar{A} = [\bar{a}_{ij}]$, where $\bar{a}_{ij}$ denotes the complex conjugate of a_{ij}. If $z = x + iy$ is a complex number, then $\bar{z} = x - iy$. For example, if A =

, then its conjugate is $\bar{A} = \begin{bmatrix} 2-4i & 2 \\ -7i & 3+i \end{bmatrix}$.

(2) The matrix A^* of a given matrix A is defined as $A^* = (\bar{A})'$, that is, A^* is the transpose of conjugate of A. For example if

$$A = \begin{bmatrix} 2+4i & 2 \\ 7i & 3-i \end{bmatrix}, \text{ then } \therefore A^* = \begin{bmatrix} 2-4i & -7i \\ 2 & 3+i \end{bmatrix}.$$

$$\begin{bmatrix} 2+4i & 2 \\ 7i & 3-i \end{bmatrix}$$

Note : Check that $(\bar{A})' = \overline{(A')}$.

16.4.13. Hermitian Matrix : A square matrix $A = [a_{ij}]$ is called **Hermitian** if $\bar{a}_{ij} = a_{ji}$ for all i and j, that is, if $A = A^*$. For example, $A = \begin{bmatrix} 2 & 1+2i \\ 1-2i & -3 \end{bmatrix}$ is a Hermitian matrix.

16.4.14. Skew - Hermitian Matrix : A square matrix $A = [a_{ij}]$ is called **skew-Hermitian** if $\bar{a}_{ij} = -a_{ji}$ for all i and j, that is, $A^* = -A$.

If $a_{ij} = x + iy$, then $\bar{a}_{ij} = -a_{ji}$ $\Rightarrow$ $x - iy = -(x + iy)$ $\Rightarrow$ $2x = 0$ $\Rightarrow$ $x = 0$. This means diagonal elements a_{ii} are purely imaginary *i.e.* $a_{ii} = 0 + iy$ or $a_{ii} = 0$, *i.e.,* all the diagonal elements in a skew Hermitian matrix are either zero or purely imaginary.

For example, $A = \begin{bmatrix} 4i & 1+2i \\ -(1-2i) & 0 \end{bmatrix}$ is a skew Hermitian matrix.

Proposition 7 : Let A and B be two square matrics of same order and m be a real or complex number. Then

(i) $(A + B)^* = A^* + B^*$

(ii) $(mA)^* = mA^*$ if m is real ;

and $(mA)^* = \bar{m}A^*$ if m is complex.

(iii) $(AB)^* = B^* A^*$

(iv) $(A^*)^* = A$.

Proof : (i) If A is a square matrix, then by definition of *, we have $A^* = (\bar{A})'$ or $\bar{A}'$.

Now using $(A + B)' = A' + B'$, we have

$$(A + B)^* = (\overline{A+B})' = (\bar{A} + \bar{B})'$$

$$= (\bar{A})' + (\bar{B})' = A^* + B^*.$$

In a similar way, we can prove (ii), (iii) and (iv).

Theorem 8 : The necessary and sufficient conditions for a matrix A to be Hermintian is that A* = A.

OR

Prove that a square matrix A is Hermintian if and only if $A^* = A$.

Theorem 9 : The necessary and sufficient conditions for a matrix A to be skew-Hermitian is that $A^* = -A$.

OR

Prove that a square matrix A is skew-Hermitian if and only if $A^* = -A$.

Proof of Theorem 9 : Suppose the given matrix is $A = [a_{ij}]$, where $i, j = 1, 2, \ldots, n$. **Necessary part** : Let the matrix A be skew-Hermitian. Then by definition, we have

$$a_{ij} = -\bar{a}_{ji} \text{ for all } i, j \qquad ...(1)$$

Now A is of order n x n

$\Rightarrow$ $\bar{A}$ is also of order n x n

$\Rightarrow$ ( )′ is also of order n x n

So A* = ( )' is of order n x n.

Thus the matrix A, $-$A and A^* are of same order.

Now (i, j)th element of A^*

= Complex conjugate of (i, j)th element of A′

= Complex conjugate of (j, i)th element of A

= $\ _{ji}$ $= -a_{ij}$, by (1).

= (i, j)th element of $(-A)$.

So the matrices A^* and $-$A have same elements. Hence by equality of matrices, we have $A^* = -A$.

Sufficient Part (or Converse Part) : Let A be a square matrix and assume that $A^* = -A$.

$\bar{A}$

This implies that A^* and $-$A are equal matrices and so they have same elements.

$\Rightarrow$ (i, j)th element of A^* = (i, j)th element of $(-A)$ $\forall i, j$

$\Rightarrow$ Complex conjugate of (i, j)th element of A′

= $-[(i, j)$th element of A$]$ $\forall\ i, j$

$\Rightarrow$ $\bar{a}_{ji} = -a_{ij}$ $\forall i, j$

This proves that A is skew-Hermitian.

Thus we have proved that A is skew-Hermitian if and only if $A^* = -A$.

Note : Proof of Theorem 8 is also similar to the proof of Theorem 9.

Theorem 10 : Every square matrix A can be uniquely expressed as P + i Q, where P and Q are Hermitian matrices and $i = \sqrt{-1}$.

Proof : Let A be a square matrix. Then $\bar{A}$, A^*, $A + A^*$ and $A - A^*$ are all matrices of same order. We can write A as

$$A = \frac{1}{2}(2A) = \frac{1}{2}(A+A) = \frac{1}{2}(A+A^*-A^*+A)$$

$$= \frac{1}{2}(A+A^*) + \frac{1}{2}(A-A^*)$$

$$= \frac{1}{2}(A+A^*) + i\,.\frac{1}{2i}(A-A^*)$$

$$= P + iQ \text{ (say), where } i = \sqrt{-1}$$

Here $P = \frac{1}{2}(A+A^*)$ and $Q = \frac{1}{2i}(A-A^*)$

Now $P^* = \frac{1}{2}(A+A^*)^* = \frac{1}{2}(A^*+A^{**}) = \frac{1}{2}(A^*+A) = P$ (as $A^{**}=A$)

$\Rightarrow$ P is Hermitian.

Also $Q^* = -\frac{1}{2i}(A-A^*)^* = -\frac{1}{2i}(A^*-A) = \frac{1}{2i}(A-A^*) = Q$

$\Rightarrow$ Q is Hermitian.

Thus $A = P + i\,Q$ ...(1)

where both P and Q are Hermintian matrices.

For uniqueness, if possible, suppose

$A = B + iC$ ...(2)

where B and C are Hermitian matrices, that is, $B^* = B$ and $C^* = C$.

Then $A^* = B^* - iC^* = B - iC$, ...(3)

since $B^* = B$ and $C^* = C$.

Solving (2) and (3) for B and C, we obtain

$$B = \frac{1}{2}(A+A^*) \text{ and } 2iC = A - A^* \Rightarrow C = \frac{1}{2i}(A-A^*)$$

Thus B = P and C = Q.

Hence representation in (1) is unique.

Example 11: If A is a Hermitian matrix, then prove that iA is a skew-Hermitian matrix.

Solution : Let A be a Hermitian matrix and $i = \sqrt{-1}$.

Then $\quad A^* = (\overline{A})' = A$ and $\quad = -i$.

Now $(iA)^* = (\overline{iA})' = (-i\overline{A})' = (-i)(\quad)' = -iA^*$

$= -iA, \qquad (\quad A^* = A)$

$= -(iA)$

$\Rightarrow \qquad i$ A is a skew-Hermitian matrix.

Example 12 : If A and B are two Hermitian square matrices of same order, then slow that AB−BA is a skew-Hermitian matrix.

Solution : Let A and B be two Hermitian square matrices of same order. Then A^*=A, B^*=B and P = AB−BA are all matrices of same order.

We have to show that $P^* = -P$.

Now $P^* = (AB-BA)^* = (AB)^* - (BA)^*$

$= B^*A^* - A^*B^*$

$= BA - AB, \qquad (\qquad A^* = A,\ B^* = B)$

$= -(AB-BA)$

$= -P.$

Thus AB−BA is a skew-Hermitian matrix.

Example 13 : Let A and B be two Hermitian matrices of same order. Then AB is Hermitian if and only AB = BA.

Solution : Let A and B be two Hermitian matrices of same order. Then

$$A^* = A \text{ and } B^* = B \qquad \text{....(1)}$$

Direct Part : Suppose first that the product AB is Hermitian. Then

$$(AB)^* = AB \qquad \text{...(2)}$$

We have to show that AB = BA.

Now AB = $(AB)^*$, by using (2)

$= B^* A^*$

$= BA$, by using (1).

So AB is Hermitian $\Rightarrow$ AB = BA.

Conserve Part : Now suppose AB = BA, ...(3)

where A and B are Hermitian matrices.

We have to prove that AB is Hermitian.

Now $(AB)^* = B^* A^* = BA$, by using (1)

$= AB$, by using (3)

$\Rightarrow$ AB is a Hermitian matrix.

Thus we have proved that AB is Hermitian if and only if AB = BA.

Example 14 : Let A be a Hermitian matrix. Then A can uniquely expressed as the sum P + iQ, where P and Q are symmetric and skew-symmetric matrices of real numbers respectively. Also show that A*A is real if and only if PQ = −QP.

Solution : Let A be a Hermitian matrix. Then $A^* = A$

$\Rightarrow \quad (\quad)' = A$ or $\quad = A'$ ...(1)

Now write A as $\frac{\bar{A}}{2i}$

$$A = \quad (2A) = \frac{1}{2}(A+A) = \frac{1}{2}(A+\bar{A}) + \quad (A- \quad)$$

$= P + iQ$ (say).

Here $P = \quad (A+\bar{A})$ and $Q = \quad (A-\bar{A})$.

Then both P and Q are real matrices.

Now that $P' = \quad (A+\bar{A})' = \quad (A' + (\bar{A})')$

$= \quad (\bar{A} + A^*) = \quad (\bar{A} + A)$, by using (1)

$= P$

$\Rightarrow$ P is symmetric.

Also $Q' = \quad (A - \bar{A})' = \quad (A' - (\bar{A})') = \quad (\bar{A} - A^*)$

$= \quad (\bar{A} - A)$ by using (1)

$= - \quad (A - \bar{A}) = -Q$

$\Rightarrow$ Q is skew-symmetric.

Thus a Hermitian matrix A can be expressed as A = P + iQ ...(2)

where $P = \quad (A + \bar{A})$ and $Q = \quad (A - \bar{A})$ are respectively symmetric and skew-symmetric matrices.

To show that representation (2) is unique, suppose, if possible, A = B + iC ...(3)

where B and C are respectively symmetric and skew-symmetric matrices of real numbers, that is, $B' = B$ and $C' = -C$ ...(4)

$\frac{A}{2i}$

Now from (1), we have

$= A' = (B + iC)' = B' + iC' = B - iC$ ...(5)

by using (4). Solving (3) and (5) for B and C, we obtain

$B = \quad (A + \bar{A}) = P$ and $C = \quad (A - \bar{A}) = Q.$

Hence representation (1) is unqiue.

Next, we will show that A^*A is a matrix of real numbers if and only if PQ = $-$QP for above A, P and Q.

First suppose A^*A is real. Then its imaginary part is zero. Now

$A^*A = AA = (P + iQ)(P + iQ)$

$= P^2 - Q^2 + i\,(PQ + QP)$

Since A^* A is real, so PQ + QP = O, being imaginary part of A^*A.

But $PQ + QP = 0 \Rightarrow PQ = -QP$.

Conversely, suppose $PQ = -QP$ or $PQ + QP = O$.

Then $A^*A = P^2 - Q^2 + i\,(PQ + QP) = P^2 - Q^2$

$= PP - QQ$

which is a real matrix because P and Q are real matrices.

16.5 DEFINITIONS :

16.5.1. Orthogonal Matrix :

A square matrix A is called **orthogonal** if $AA' = I$.

We note that A is orthogonal $\Rightarrow AA' = I$

$\Rightarrow \quad |AA'| = |I| \quad \Rightarrow \quad |A|\,|A'| = 1$

$\Rightarrow \quad |A|^2 = 1 \quad\quad (\;|A| = |A'|)$

$\Rightarrow \quad |A| = \pm 1$.

Example : $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is an orthogonal matrix.

16.5.2. A square matrix A is called **unitary** if $A^*A = I$.

Example : $A = \begin{bmatrix} \frac{1}{2}(1+i) & \frac{1}{2}(i-1) \\ \frac{1}{2}(1+i) & \frac{1}{2}(1-i) \end{bmatrix}$ is a unitary matrix.

Proposition 15 : If A is an orthogonal matrix, then A^{-1} and A' are also orthogonal.

Proof : Suppose A is an orthogonal matrix. Then by definition, we have $AA' = I$, where I is the unit matrix. Also $|A| = \pm 1 \Rightarrow A^{-1}$ exists.

Now $AA' = I \quad \Rightarrow \quad A^{-1} = A'$ and $A'A = I$.

(*i*) To prove A^{-1} is orthgonal, first note that $(A^{-1})' = (A')^{-1}$.

Now $A^{-1}.(A^{-1})' = A^{-1}(A')^{-1} = (A'A)^{-1} = (I)^{-1} = I$.

⇒ A^{-1} is orthogonal.

(*ii*) To prove A' is orthogonal, consider

$A'(A')' = (A'A)' = I' = I$

⇒ A' is orthogonal.

Proposition 16 : If A and B are orthogonal matrices, then AB is also orthogonal.

Proof : Given A, B are orthogonal ⇒ $AA' = I$ and $BB' = I$.

Also $(AB)' = B'A'$. Now

$(AB)(AB)' = (AB)(B'A') = A(BB')A'$, by associative law

$= A(I)A' = AA' = I$

⇒ AB is orthogonal.

Proposition 17 : If A is a unitary matrix, then A', A^{-1}, A^* and $\bar{A}$ are also unitary matrices.

Proof : Let A be a unitary matrix. Then by definition, we have

$AA^* = I = A^*A$, where $A^* = (\bar{A})'$. This also implies that $A^{-1} = A^*$.

(*i*) **A' is unitary :**

$\overline{(AA^{-1})}$

$A'(A')^* = A'(A^*)' = (A^*A)' = I' = I$

⇒ A' is unitary.

(*ii*) **A^{-1} is unitary :**

$A^{-1}(A^{-1})^* = A^{-1}(A^*)^* = A^{-1}A$ ∵ $(A^*)^* = I$

$= I$

⇒ A^{-1} is unitary.

(*iii*) **A^* is unitary :**

$A^*(A^*)^* = (A^*A)^*$ ($(AB)^* = B^*A^*$)

$= I^* = I$

⇒ A^* is unitary.

(*iv*) **is unitary :**

$(\ )^* = (\ ^*) = (\ ^{-1}) = \ = \bar{I} = I$

⇒ $\bar{A}$ is unitary.

Example 18 : If A is symmetric and B is orthogonal, then show that $B^{-1}AB$ is symmetric.

Solution : Suppose A is symmetric and B is orthogonal. Then

$$A' = A \text{ and } BB' = I = B'B.$$

Also $B^{-1} = B'$.

Now $(B^{-1}AB)' = B'A'(B^{-1})' = B'A(B')' = B'AB,$ $\quad (\because (B')' = B)$

$= B^{-1}AB$ $\quad (\because B^{-1} = B')$.

This shows that $B^{-1}AB$ is symmetric matrix.

Example 19 : Let A and B be two non-singular matrices of order n satisfying $AA' = BB'$. Then prove that there is an orthgonal matrix P of order n satisfying $A = BP$.

Solution : Suppose A and B are two non-singular matrices of order n. Then $|A| \neq 0$ and $|B| \neq 0$. So A^{-1} and B^{-1} exist. Also A' and B' are non-singular so that

$$(A')^{-1} = (A^{-1})' \text{ and } (B')^{-1} = (B^{-1})'$$

$\Rightarrow$ $\quad (A')^{-1}$ and $(B')^{-1}$ exist.

Now given $AA' = BB'$ $\quad \because$ ...(1)

Multiplying (1) on both sides by $(A')^{-1}$ from right side, we have

$$A = BB'(A')^{-1} = BP \text{ (say)}.$$

Here $P = B'(A')^{-1}$. We have to prove that P is an orthogonal matrix of order n. Clearly P is of order n as it is product of two matrices of order n.

Now $\quad PP' = [B'(A')^{-1}][B'(A')^{-1}]' = B'(A')^{-1}(A^{-1})''B''$

$= B'(A')^{-1}A^{-1}B$ $\quad (\because B'' = B)$

$= B'(AA')^{-1}B$

$= B'(BB')^{-1}B$ $\quad (\because AA' = BB')$

$= B'((B')^{-1}B^{-1})B$

$= B'(B^{-1})'(B^{-1}B)$

$= (B^{-1}B)'I$ $\quad B^{-1}B = I$

$= I'I = I.$

Thus P is orthogonal. Hence A = BP.

Example 20 : Show that the matrices and $\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$ are unitary.

Try Your Self

Example 21 : Show that $\begin{bmatrix} 0 & 2m & n \\ l & m & -n \\ l & -m & n \end{bmatrix}$ is an orthgonal matrix, where

$$l = \frac{1}{\sqrt{2}},\ m = \frac{1}{\sqrt{6}},\ n = \frac{1}{\sqrt{3}}.$$

Solution : Let A = $\begin{bmatrix} 0 & 2m & n \\ l & m & -n \\ l & -m & n \end{bmatrix}$. Then $A' = \begin{bmatrix} 0 & l & l \\ 2m & m & -m \\ n & -n & n \end{bmatrix}$

$\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}$

Then $AA' = \begin{bmatrix} 0 & 2m & n \\ l & m & -n \\ l & -m & n \end{bmatrix} \begin{bmatrix} 0 & l & l \\ 2m & m & -m \\ n & -n & n \end{bmatrix}$

$$= \begin{bmatrix} 4m^2 + n^2 & 2m^2 - n^2 & -2m^2 + n^2 \\ 2m^2 - n^2 & l^2 + m^2 + n^2 & l^2 - m^2 - n^2 \\ -2m^2 + n^2 & l^2 - m^2 - n^2 & l^2 + m^2 + n^2 \end{bmatrix}$$

Now $\quad 4m^2 + n^2 = \frac{4}{6} + \frac{1}{3} = \frac{6+2}{6} = \quad = 1.$

$$2m^2 - n^2 = \frac{2}{6} - \frac{1}{3} = 0, \; 2m^2 + n^2 = -\frac{2}{6} + \frac{1}{3} = 0,$$

$$l^2 + m^2 + n^2 = \frac{1}{2} + \frac{1}{6} + \frac{1}{3} = \frac{3+1+2}{6} = \quad = 1,$$

$$l^2 - m^2 - n^2 = \frac{1}{2} - \frac{1}{6} - \frac{1}{3} = \frac{3-1-2}{6} = 0.$$

Using these values, we have

$$AA' = \qquad = I_3 \;\Rightarrow\; \text{A is orthogonal.}$$

16.6 SUMMARY :

In this lesson, we studied various types of matrices viz., symmetric skew-symmetric, Hermitian, skew-Hermitian, orthogonal and unitary matrices with examples. Important results and problems are also discussed.

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

16.7 EXERCISE

1. Give an example of a non-zero matrix A such that $A^2 = O$ (zero matrix).
2. Give an example of non-zero matrices A and B such that AB = O.
3. Give examples of two 3 x 3 matrices which are non-singular.
4. Give examples one each of a 3 x 3 Hermitiain matrix and a skew-Hermitian matrix.
5. Express $A = \begin{bmatrix} 2 & -6 & 4 \\ 8 & 5 & 3 \\ 0 & 1 & -9 \end{bmatrix}$ as a sum of a symmetric and a skew-symmetric matrix.
6. Verify that result that $(AB)' = B'A'$ for non-zero matrices A and B.

7. Let A be a square matrix. Then show that AA' is symmetric.

8. Let A and B be two square symmetric matrices of same order. Then prove that AB is skew-symmetric if and only if $AB = -BA$.

9. Let A and B be two square matrices of same order. Then show that $B'AB$ is skew-symmetric if A is skew-symmetric.

10. Show that a square matrix A is Hermitian if and only if $A^* = A$.

11. If A is a skew-Hermitian matrix, then show that iA is a Hermitian matrix.

12. Let A and B be two square matrices. Then $(AB)^* = B^*A^*$.

13. Let A be a square matrix. Then prove that $A - A^*$ is a skew-Hermitian matrix.

14. Let A and B be two square Hermitian matrices of same order. Then prove that AB is skew-Hermitian if and only if $AB = -BA$.

15. Let A and B be two square matrices of same order. Then show that B^*AB is Hermitian if A is Hermitian.

16. Show that the following matrices are orthogonal :

(*i*) $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ (*ii*) $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ (*iii*) $\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$

17. Show that the following matrices are unitary :

(*i*) $\begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}$ (*ii*) $\frac{1}{2}\begin{bmatrix} 1+i & 1+i \\ -1+i & 1-i \end{bmatrix}$

18. If A is an orthogonal matrix, then show that $|A| \neq 0$ and A^{-1} is also orthogonal.

19. Show that the product of two orthogonal matrices is again an orthogonal matrix.

20. If A is a unitary matrix, then show that A^* and A^{-1} are also unitary matrices.

ANSWER

5. $\begin{bmatrix} 2 & 1 & 2 \\ 1 & 5 & 2 \\ 2 & 2 & -9 \end{bmatrix} + \begin{bmatrix} 0 & -7 & 2 \\ 7 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}$

16.8. SUGGESTED READING/REFERENCES :

1. "**A text book of Matrices**" by Shanti Narayan and P.K.Mittal, S. Chand & Company Ltd.

2. "**Theory of Matrices**" by B.S. Vatssa, Wiley Easterns Ltd.

3. "**Engineering Mathematics-II**" by Bhopinder Singh, Malhotra Brothers, Jammu.

16.9. MODEL TEST PAPER

1. Give an example of a non-zero matrix A for which $A^2 = A$.
2. Give an example of matrices A, B and C such that $A \neq O$ and $B \neq C$ but $AB = AC$.
3. Give examples of two 3 x 3 non-zero matrices which are singular.
4. Give examples one each of a 3 x 3 symmetric matrix and a skew-symmetric matrix.
5. Let $A = \begin{bmatrix} 2 & 4 & 6 \\ -5 & 3 & 7 \\ 1 & 0 & 9 \end{bmatrix}$. Then find two matrices of order 3 such that A = B + C, where B is symmetric and C is skew-symmetric.
6. Show that a square matrix A is skew-symmetric if and only if $A' = -A$.
7. Let A be a square matrix. Then prove that $A - A'$ is a skew-symmetric matrix.
8. Let A and B be two square symmetric matrices of same order. Then prove that AB is symmetric if and only if AB = BA (*i.e.*, A and B commute).
9. Let A be a square matrix. Then prove that $(A^*)^* = A$.
10. Every square matrix A can be written as $A = B + iC$, where B and C are both Hermitian matrices and $i = \sqrt{-1}$. Prove it.
11. Verify the result that $(AB)^* = B^*A^*$ for two non-zero matrices.
12. Let A be a Hermitian matrix. Then prove that $(-2i)A$ is a skew-Hermitian matrix.
13. Let A and B be two square Hermitian matrices of same order. Then prove that AB + BA is a Hermitian matrix.
14. Let A and B be two square matrices of same order. Then prove that B^*AB is skew-Hermitian if A is skew-Hermitian.
15. Show that the following matrices are orthogonal

(*i*) $\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{bmatrix}$ (*ii*) $\frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{bmatrix}$

16. Let A be an orthogonal matrix. Then show that A' is also orthogonal and $|A'| \neq 0$.

17. Let A be a skew-symmetric matrix and B be an orthogonal matrix. Then show that $B^{-1}AB$ is a skew-symmetric matrix.

18. Show that the following matrices are unitary :

(*i*) $\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$ (*ii*) $\begin{bmatrix} -i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 1 \end{bmatrix}$

19. Let A be a unitary matrix and B be a skew-Hermitian matrix. Then show that $B^{-1}AB$ is a skew-Hermitian matrix.

20. Prove that product of two unitary matrices is a unitary matrix.

MATRICES

RANK A MATRIX
CAYLEY-HAMILTON THEOREM

By: Dr. Kamaljeet Kour

17.1 INTRODUCTION :

Given a matrix of any order, its rank is the order of largest non-zero minor and so is a positive number. Cayley-Hamilton theorem is an important result for a square matrix stating that it always satisfies its characteristic equation.

17.2. OBJECTIVES :

Besides calculating rank of any matrix, the main objective is to test the consistency of system of linear equation by rank method. Cayley-Hamilton theorem is applied to calculate the inverse of non-singular matrices.

17.3 DEFINITIONS :

(*i*) The determinant of a square sub-matrix of a matrix A is called a minor of A. The minor formed by deleting the *i*th row only and the *j*th column only of a matrix A = $[a_{ij}]$ is denoted by M_{ij} . For example, if A is a matrix of order 3 x 4, then consider its square submatrices of order 3, 2 and 1 respectively. The determinants of these sub-matrices are minors of order 3, 2 and 1.

(*ii*) Let A = $[a_{ij}]$ be a square matrix of order *n* and let A_{ij} be the cofactor of the element a_{tj} in the determinant | A | . Then the matrix $[A_{ij}]'$ is called the adjoint of A and is denoted by adj A. It can be proved that if A = $[a_{ij}]$, then

$$A\,(\text{adj } A) = |A|.\, I_n = (\text{adj } A).\, A,$$

(*iii*) Let A be a square matrix of order *n.* If there is a matrix B of order *n* such that

$$AB = I_n = BA,$$

then B is called the inverse of A and is denoted by A^{-1} .

The inverse of a square matrix A exists if and only if $|A| \neq 0$, and is given by $A^{-1} = \frac{1}{|A|}(\text{adj } A)$.

(*iv*) A matrix is said to be of **rank *r*** (a positive number) if it has atleast one non-zero minor of order *r* and every minor of order higher than *r* is zero. Briefly, the rank of a matrix is the largest order of its any non-vanishing minor.

Notation : We use *r* (A) or ran (A) or $\rho(A)$ to denote the rank of a matrix A.

Observations : (*i*) The rank of a zero matrix is taken to be zero.

(*ii*) For a non-zero matrix A, $r(A) \geq 1$

(*iii*) If A is a non-singular matrix of order *n,* then $r(A) = n$. In particular $r(I_n) = n$

(*iv*) If A is a singular matrix of order *n*, then $r(A) < n$.

(*v*) If A is a matrix of order *m* x *n,* then $r(A) \leq \min\{m, n\}$.

Examples 1 : Let us consider the matrices :

(*i*) $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 0 & 2 & 2 \end{bmatrix}$. Then $|A| = 2 \neq 0$, that is, the largest order minor of order 3 is non-zero. So rank of A is 3.

(*ii*) $B = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & -2 \\ 2 & 4 & -3 \end{bmatrix}$. Then $|B| = 0$, that is, the largest order minor of B, which is B, is zero. So $r(A) \neq 3$. But the minor of order 2 :

$\begin{vmatrix} 2 & 1 \\ 0 & 3 \end{vmatrix} = 6 \neq 0$. So $r(A) = 2$.

(*iii*) $C = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \end{bmatrix}$

Here C is a matrix of order 3 x 4 and check that its all minors of order 3, and also of order 2 are zero. So $r(A) \neq 3$, $r(A) \neq 2$. But since C is a non-zero matrix, therefore its rank is 1.

Remark : The rank of a matrix depends upon the zero or non-zero nature of its minors. But when order of matrix is more than 4 x 3, it may become lengthy to calculate its minors. These calculations can be avoided while finding rank of a matrix by the use of elementary row or column transformations which are given below :

17.3.1. Elementary transformations on a matrix :

The following operation, three of which refer to rows and three to columns are known as elementary transformations :

(*i*) The interchange of any two rows (or columns).

(*ii*) The multiplication of any row (or column) by a non zero number.

(*iii*) The addition of a constant multiple of the elements of any row (or column) to the corresponding element of any other row (or column).

Notation: The elementary row transformation will be denoted by the following symbols.

(*i*) R_{ij} for the interchange of the *i*th and *j*th rows.

(*ii*) $R_i \to kR_i$ for multiplication of the *i*th row by *k,* we have $k \neq 0$.

(*iii*) $R_i \to R_i + kR_j$ for addition to the *i*th row elements with *k* times the *j*th row elements.

Similarly, we have elementary column transformations :

$$C_{ij}, \; C_i \to kC_i \text{ and } C_i \to C_i + kC_j.$$

17.3.2 Equivalent matrix :

Two matrices A and B are said to be **equivalent** if one can be obtained from the other by using elementary transformations.

Notation : If A and B are equivalent, then we write it as A ~ B.

17.3.3 Elementary matrices :

An elementary matrix is that which is obtained from a unit matrix by subjecting it to any of the elementary transformations.

For example from $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, we can obtain some elementary matrices as given below :

$$R_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = C_{23}\,, \quad kR_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix} = kC_2,$$

$$R_1 + kR_2 = \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = C_2 + kC_1, \text{ where } k \neq 0.$$

17.3.4. Effect of Elementary transformations on product of matrices :

Result : (*i*) R(AB) = R(A).B

(*ii*) C (AB) = A.C(B),

where R and C are used respectively for row and column transformations.

Let us understand it by the following two examples.

(*i*) Let $A = \begin{bmatrix} 2 & -3 & 4 \\ 3 & 2 & 1 \\ 0 & 1 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & 1 \\ 2 & -1 & 3 \\ 1 & 0 & 4 \end{bmatrix}$

Then $AB = \begin{bmatrix} 0 & 7 & 9 \\ 8 & 4 & 13 \\ 6 & -1 & 19 \end{bmatrix}$

Suppose row transformation R is $R_1 \to R_1 - R_3$. Let us apply it on A and AB as follows :

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

and $R(AB) = \begin{bmatrix} -6 & 8 & -10 \\ 8 & 4 & 13 \\ 6 & -1 & 19 \end{bmatrix}$

Now multiply matrix R(A) with B, we obtain

$$R(A).B = \begin{bmatrix} 2 & -4 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 2 & -1 & 3 \\ 1 & 0 & 4 \end{bmatrix} = \begin{bmatrix} -6 & 8 & -10 \\ 8 & 4 & 13 \\ 6 & -1 & 19 \end{bmatrix} = R(AB)$$

Similarly, we can check that if X, Y, Z are three matrices for which product XYZ is a new matrix, then R(XYZ) = R(X)YZ

(*ii*) Again consider the same matrices A and B as in above example and let us apply a column transformation C, say $C_3 \to C_3 - C_1$ on B and AB as follows :

$$C(B) = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 1 \\ 1 & 0 & 3 \end{bmatrix}$$

and $$C(AB) = \begin{bmatrix} 0 & 7 & 9 \\ 8 & 4 & 5 \\ 6 & -1 & 13 \end{bmatrix}$$

Now multiply A with matrix C(B); we obtain

$$A.C(B) = \begin{bmatrix} 2 & -3 & 4 \\ 3 & 2 & 1 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 1 \\ 1 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 7 & 9 \\ 8 & 4 & 5 \\ 6 & -1 & 13 \end{bmatrix} = C(AB).$$

Similarly, it can be varified that C(XYZ) = XY.C(Z).

Thus general result is

(*i*) **R(XYZ) = R(X).YZ**

(*ii*) **C(XYZ) = XY.C(Z)**,

for any three matrices X, Y, Z for which product XYZ is a new matrix.

This result say that the elementary row transformation on any product of matrices effects on the first left factor of the product, where as the elementary column transformation on any product of matrices effects on the last right factor of the product.

In particular, for two matrices, one any matrix A and another unit matrix, both of same order, we have

(*i*) R(IA) = R(I).A

⇒ R(A) = (Elementary matrix).A

that is, every row transformation on a matrix is equivalent to pre-multiplying A by the corresponding elementary matrix.

(*ii*) C(AI) = A C(I)

⇒ C(A) = A (Elementary matrix)

that is, every column transformation on a matrix is equivalent to post-multiplying A by the corresponding elementary matrix.

17.3.5 Inverse of a matrix by use of elementary transformations :

This method is also known as Gauss Jordan method and is as follows :

Take a non-singular matrix A of order n so that $r(A) = n$, that is, $A \sim I_n$. If we have to use only elementary row-transformations, then we write $A = I_n A$ and apply a number of row transformations to this until L.H.S becomes I_n and R.H.S product is, say BA.

Now $I_n = BA$ ⇒ $A^{-1} = B$

In a similar way, we can compute inverse of a non-singular matrix A by using column transformations too. In this case we have to start by writing $A = AI_n$ and then apply some number of elementary column transformations to have $I_n = AP$ so that $A^{-1} = P$.

Example 2 : Compute the inverse of the following matrices :

(*i*) $A = \begin{bmatrix} 1 & -1 & 1 \\ 4 & 1 & 0 \\ 8 & 1 & 1 \end{bmatrix}$ by using elementary column transformations.

(*ii*) $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$ by using elementary row transformation.

Solutions. (*i*) Here we note that $|A| = 1 \neq 0$. So that $A \sim I_3$. Write $A = AI_3$, that is

$$\begin{bmatrix} 1 & -1 & 1 \\ 4 & 1 & 0 \\ 8 & 1 & 1 \end{bmatrix} = A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Applying on both sides $C_3 \to C_3 + C_2$, we have $\begin{bmatrix} 1 & -1 & 0 \\ 4 & 1 & 1 \\ 8 & 1 & 2 \end{bmatrix} = A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

Applying on both sides $C_1 \to C_1 - 4C_3$, we have $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} = A \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 1 \\ -4 & 0 & 1 \end{bmatrix}$

Applying on both sides $C_2 \to C_2 + C_1$, we have $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} = A \begin{bmatrix} 1 & 1 & 0 \\ -4 & -3 & 1 \\ -4 & -4 & 1 \end{bmatrix}$

Applying on both sides $C_3 \to C_3 - C_2$, we have $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = A \begin{bmatrix} 1 & 1 & -1 \\ -4 & -3 & 4 \\ -4 & -4 & 5 \end{bmatrix}$

Applying on both sides $C_2 \to C_2 - C_3$, we have $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = A \begin{bmatrix} 1 & 2 & -1 \\ -4 & -7 & 4 \\ -4 & -9 & 5 \end{bmatrix}$

or $\quad I_3 = AP$, where $P = \begin{bmatrix} 1 & 2 & -1 \\ -4 & -7 & 4 \\ -4 & -9 & 5 \end{bmatrix}$

Thus $A^{-1} = B = \begin{bmatrix} 1 & 2 & -1 \\ -4 & -7 & 4 \\ -4 & -9 & 5 \end{bmatrix}$

(*ii*) Here $|A| = 0 \ -(1-9) + 2\,(1-6) = -2 \neq 0$. So A^{-1} exists.

Let us write $A = I_3 \,.\, A$

$$\Rightarrow \quad \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A$$

Apply on both sides R_{12}, we have $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} A$

Apply on both sides $R_3 \to R_3 - 3R_1$, we have $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -5 & -8 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -3 & 1 \end{bmatrix} A$

Apply on both sides $R_1 \to R_1 - 2R_2$ and $R_3 \to R_3 + 5R_2$, we have

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

Apply on both sides $R_2 \to R_2 - R_3$, we have $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 0 \\ -4 & 3 & -1 \\ 5 & -3 & 1 \end{bmatrix} A$

Apply on both sides $R_1 \to R_1 + \frac{1}{2} R_3$, we have [illegible] $= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -4 & 3 & -1 \\ 5 & -3 & 1 \end{bmatrix} A$

Apply $R_3 \to \frac{1}{2} R_3$, we have $I_3 =$ [illegible] $A = PA$

$\Rightarrow \quad A^{-1} = P$, where $P = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -4 & 3 & -1 \\ \frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix} = \frac{1}{2}$

Verification :

$AA^{-1} = \frac{1}{2}$ [illegible] $\begin{bmatrix} 1 & -1 & 1 \\ -8 & 6 & -2 \\ 5 & -3 & 1 \end{bmatrix}$

$$= \frac{1}{2} \qquad\qquad = I_3$$

Hence inverse is correct.

Observations : The following observations will be helpful while finding ranks of given matrices using elementary transformations.

(*i*) Elementary transformations do not alter the rank or order of a matrix while the values of minors may get changed.

(*ii*) Equivalent matrices have the same rank and the same order.

Examples 3 : Using elementary transformations, find ranks of the following matrices:

(*i*) $A = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 3 & 4 & 1 & 2 \\ -2 & 3 & 2 & 5 \end{bmatrix}$

(*ii*) $A = \begin{bmatrix} 1 & 3 & 4 & 3 \\ 3 & 9 & 12 & 9 \\ -1 & -3 & -4 & -3 \end{bmatrix} \qquad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

(*iii*) $\begin{bmatrix} 2 & 31 & 15 & 41 \\ 6 & 94 & 30 & 82 \\ 6 & 94 & 31 & 82 \\ 2 & 32 & 14 & 42 \end{bmatrix}$

Solutions.

(*i*) $A = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 3 & 4 & 1 & 2 \\ -2 & 3 & 2 & 5 \end{bmatrix}$

$R_2 \rightarrow R_2 - 3R_1$

$$\sim \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & -2 & 1 & 5 \\ 0 & 7 & 2 & 3 \end{bmatrix}$$

$R_3 \rightarrow R_3 + 2R_1$

$C_2 \rightarrow C_2 - 2C_1$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 5 \\ 0 & 7 & 2 & 3 \end{bmatrix}$$

$C_4 \rightarrow C_4 - C_1$

$C_2 \rightarrow C_2 + 2C_3$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 11 & 2 & -7 \end{bmatrix}$$

$C_4 \rightarrow C_4 - 5C_3$

$C_2 \rightarrow \frac{1}{11} C_2$

$$\underset{\sim}{R_{23}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = [I_3 : O].$$

So $r(A) = r([I_3 : O]) = 3$

(*ii*) $A = \begin{bmatrix} 1 & 3 & 4 & 3 \\ 3 & 9 & 12 & 9 \\ -1 & -3 & -4 & -3 \end{bmatrix}$

$$\underset{R_3 \to R_3 + R_1}{\overset{R_2 \to R_2 - 3R_1}{\sim}} \begin{bmatrix} 1 & 3 & 4 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\underset{\substack{C_3 \to C_3 - 4C_1 \\ C_4 \to C_4 - 3C_1}}{\overset{C_2 \to C_2 - 3C_1}{\sim}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = B, \text{ say.}$$

Note that B, all 3rd order and 2nd order sub-matrices of B are singular.

So $r(A) = r(B) = 1$.

(*iii*) Let $A = \begin{bmatrix} 2 & 31 & 15 & 41 \\ 6 & 94 & 30 & 82 \\ 6 & 94 & 31 & 82 \\ 2 & 32 & 14 & 42 \end{bmatrix}$

$C_1 \to \frac{1}{2} C_1$

$C_2 \rightarrow C_2 - C_1$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 31 & -15 & -41 \\ 3 & 29 & -14 & -41 \\ 1 & 3 & -1 & 1 \end{bmatrix}$$

$C_3 \rightarrow C_3 - 15C_1$

$C_4 \rightarrow C_4 - 41C_1$

$R_2 \rightarrow R_2 - 3R_1$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 31 & -15 & -41 \\ 0 & 29 & -14 & -41 \\ 0 & 3 & -1 & 1 \end{bmatrix}$$

$R_3 \rightarrow R_3 - 3R_1$

$R_4 \rightarrow R_4 - R_1$

$R_2 \rightarrow R_2 - R_3 - R_4$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 29 & -14 & -41 \\ 0 & 3 & -1 & 1 \end{bmatrix}$$

$C_4 \rightarrow C_4 - C_2$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 29 & -14 & -70 \\ 0 & 3 & -1 & -2 \end{bmatrix}$$

$R_3 \rightarrow R_3 + 29R_2$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -14 & -70 \\ 0 & 0 & -1 & -2 \end{bmatrix}$$

$R_4 \rightarrow R_4 + 3R_2$

$C_2 \rightarrow -C_2$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -14 & -42 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

$C_4 \rightarrow C_4 - 2C_3$

$R_3 \rightarrow R_3 - 14R_4$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -42 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

$R_3 \rightarrow \frac{-1}{42} R_3$ $\qquad\qquad$ R_{34}

$$C_3 \to C_3 - C_2 \quad \sim \quad \begin{bmatrix} 1 & 0 & 0 \\ 2 & -2 & 0 \\ 3 & -4 & 1 \end{bmatrix} = C, \text{ say,}$$

which is a lower triangular matrix. Note that ran A = ran C = 3.

Observation : If we reduce the given matrix to a triangular form, then rank is the number of non-zero rows (or columns).

17.3.6. Normal Forms :

Every non-zero matrix of rank 'r' can be reduced to one of the following forms by using elementary row and column transformations :

$$\begin{bmatrix} I_r \\ \cdots \\ O \end{bmatrix}, \; [\, I_r \,\vdots\, O], \; \begin{bmatrix} I_r & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & O \end{bmatrix}, \text{ or } I_r$$

where I_r is the identity matrix of order r and O stands for zero matrix of some order. The above forms are called **Normal forms** or **canonical forms** of the given matrix.

We know that elementary transformations do not alter the rank or order of the matrix. Therefore the rank of the normal form will be the same as the rank of a given matrix.

Examples 5 : Reduce the following matrices to their normal forms and then find their ranks:

(*i*) $\begin{bmatrix} 8 & 1 & 3 & 6 \\ 0 & 3 & 2 & 2 \\ -8 & -1 & -3 & 4 \end{bmatrix}$

(*ii*) $\begin{bmatrix} 0 & 1 & -3 & -1 \\ 1 & 0 & 1 & 1 \\ 3 & 1 & 0 & 2 \\ 1 & 1 & -2 & 0 \end{bmatrix}$

(*iii*) $\begin{bmatrix} 1 & 3 & 4 & 5 \\ 1 & 2 & 3 & 2 \\ 2 & 3 & 5 & 1 \end{bmatrix}$

(*iv*) $\begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & -2 & 1 \\ 1 & -1 & 4 & 0 \\ -2 & 2 & 8 & 0 \end{bmatrix}$

Solutions. (*i*) We have $A = \begin{bmatrix} 8 & 1 & 3 & 6 \\ 0 & 3 & 2 & 2 \\ -8 & -1 & -3 & 4 \end{bmatrix}$

$R_3 \rightarrow R_3 + R_1$

$\sim \begin{bmatrix} 8 & 1 & 3 & 6 \\ 0 & 3 & 2 & 2 \\ 0 & 0 & 0 & 10 \end{bmatrix}$

$C_1 \rightarrow \frac{1}{8} C_1$

$$C_4 \rightarrow \frac{1}{10} C_4$$

$$C_3 \to C_3 + 2C_1 - C_4 \quad \sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$R_3 \to R_3 - R_1 - 2R_2 \quad \sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$C_{12}, C_{34} \quad \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$C_{23} \quad \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} I_2 & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & O \end{bmatrix}$$

which is in normal form. Thus $r(A) = \text{ran} \begin{bmatrix} I_2 & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & O \end{bmatrix} = 2$.

(*iii*) $\quad A = \begin{bmatrix} 1 & 3 & 4 & 5 \\ 1 & 2 & 3 & 2 \\ 2 & 3 & 5 & 1 \end{bmatrix}$

$$\begin{matrix} C_2 \to C_2 - 3C_1 \\ \sim \\ C_3 \to C_3 - 4C_1 \\ C_4 \to C_4 - 5C_1 \end{matrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & -1 & -3 \\ 2 & -3 & -3 & -9 \end{bmatrix}$$

$C_3 \to C_3 - C_2$

$\sim$ $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 2 & -3 & 0 & 0 \end{bmatrix}$

$C_4 \to C_4 - 3C_2$

$R_2 \to R_2 - R_1$

$\sim$ $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -3 & 0 & 0 \end{bmatrix}$

$R_3 \to R_3 - 2R_1$

$R_3 \to R_3 - 3R_2$

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

$R_2 \to -R_2$

$\sim$ $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} I_2 & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & O \end{bmatrix}$

which is in normal form. Thus $r\,(A) = r \begin{bmatrix} I_2 & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & O \end{bmatrix} = 2.$

(*iv*) We have

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

$C_3 \to C_3 - 2C_1$

$\sim$ $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 \\ 1 & -1 & 2 & -1 \\ -2 & 2 & 12 & 2 \end{bmatrix}$

$C_4 \to C_4 - C_1$

$R_3 \to R_3 - R_1$

$\sim$

$R_4 \to R_4 + 2R_1$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & -1 & 2 & -1 \\ 0 & 2 & 12 & 2 \end{bmatrix}$$

$R_3 \to R_3 + R_2$

$\sim$

$R_4 \to R_4 - 2R_2$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 \end{bmatrix}$$

$C_3 \to C_3 + 2C_2$

$\sim$

$C_4 \to C_4 - C_2$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 \end{bmatrix}$$

$R_4 \to \frac{1}{16} R_4$

Since elementary matrices are non-singular, P and Q are non-singular.

Corollaries : (*i*) Every non-singular matrix can be expressed as a product of elementary matrices.

Proof : Let B be a non-singular matrix of order n and let $A = I_n$. Then by above theorem, we have $B = P\, I_n\, Q = PQ$.

(*ii*) If A is an m x n matrix of rank r, then there exists two non-singular matrices P and Q of order m and n respectively such that

$$PAQ = \begin{bmatrix} I_r & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & O \end{bmatrix}, \text{ that is, PAQ is in normal form.}$$

Proof : Since ran (A) = r = ran (I_r), we have $A \sim \begin{bmatrix} I_r & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & O \end{bmatrix}$ by applying elementary column and row transformations. But elementary row transformation and column transformation can be affected by pre- and post multiplying A by elementary matrices, we have

$$P_1 P_2 \ldots\ldots P_j A \; Q_1 Q_2 \ldots\ldots Q_k = \begin{bmatrix} I_r & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & O \end{bmatrix}$$

or $$PAQ = \begin{bmatrix} I_r & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & O \end{bmatrix}, \text{ where } P = P_1P_2 \ldots\ldots P_j \text{ and } Q = Q_1 Q_2 \ldots\ldots Q_k.$$

Examples 6 : For the following matrices, find two non-singular matrices P and Q such that PAQ is in normal form :

(*i*) $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 0 & -1 & -1 \end{bmatrix}$ (*ii*) $A = \begin{bmatrix} 3 & -1 & 2 \\ -6 & 2 & 4 \\ -3 & 1 & 2 \end{bmatrix}$

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

Solution. (*i*) Here given matrix A is of order 3 x 3. So we can write $A = I_3 A I_3$.

that is, $\begin{bmatrix} 1 & -1 & 2 \\ 1 & 2 & 3 \\ 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Applying on both sides $C_2 \to C_2 - C_1$ and $C_3 \to C_3 - 2C_1$, we have

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

Applying on both sides $R_2 \to R_2 - R_1$, we have

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

Applying on both sides $C_3 \to C_3 - C_2$, we have

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

Applying on both sides $R_3 \to R_3 + R_2$, we have

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

or $\left[\begin{array}{c:c} I_2 & O \\ \hdashline O & O \end{array}\right] = PAQ$ which is in normal form, where

$$P = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

Since $P \sim I_3$ and $Q \sim I_3$, both P and Q are non-singular.

(*ii*) Since A is order 3, so we can write $A = I_3 \, A \, I_3$,

$$\Rightarrow \quad \begin{bmatrix} 3 & -1 & 2 \\ -6 & 2 & 4 \\ -3 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Applying on both sides $R_1 \rightarrow R_1 + R_3$ and $R_2 \rightarrow R_2 - 2R_3$, we have

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

Applying on both sides $C_1 \rightarrow C_1 + 3\,C_2$ and $C_3 \rightarrow C_3 - 2C_2$, we have

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

Applying on both sides $C_3 \rightarrow \frac{1}{4} C_3$ followed by C_{13}, we have

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

Applying on both sides R_{23}, we have

$$\begin{bmatrix} I_2 & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & O \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 1 \\ -\frac{1}{2} & 1 & 3 \\ \frac{1}{4} & 0 & 0 \end{bmatrix}$$

= PAQ which is in normal form, where

$$P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} 1 & 0 & 1 \\ -\frac{1}{2} & 1 & 3 \\ \frac{1}{4} & 0 & 0 \end{bmatrix}$$

Since $P \sim I_3$ and $Q \sim I_3$, both P and Q are non-singular.

(*iii*) Here given matrix A is of order 3 x 4, so we can write $A = I_3\, A\, I_4$,

$$\Rightarrow \qquad \begin{bmatrix} 1 & 1 & 2 & 3 \\ 1 & 2 & 3 & 2 \\ 0 & -1 & -1 & 4 \end{bmatrix} = I_3 \; A \; I_4$$

Applying on both sides $R_2 \to R_2 - R_1$, we have

$$\begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & -1 & -1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} AI_4$$

Applying on both sides $R_1 \to R_1 - R_2$ and $R_3 \to R_3 + R_2$, we have

$$\begin{bmatrix} 1 & 0 & 1 & 4 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} AI_4$$

Applying on both sides $C_3 \to C_3 - C_2 - C_1$ and $C_4 \to C_4 - 3C_1 - C_3 + 2C_2$, we have

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

Applying on both sides C_{34}, we have

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

Apply on both sides $R_3 \to \frac{1}{3} R_3$, we have

$$[\, I_3 \quad O] = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} AQ = PAQ$$

$\Rightarrow$ PAQ is in normal form, where

$$P = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \text{ and } Q = \begin{bmatrix} 1 & 0 & -3 & -1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Since $P \sim I_3$ and $Q \sim I_3$, so both P and Q are non-singular.

Note : Given a matrix A, we can find another set of non-singular matrices P_1 and Q_1 such that $P_1 A Q_1$ (besides PAQ) is in normal form, that is, P and Q thus obtained in above examples are not unique.

17.4. SOLUTION OF SYSTEM OF LINEAR EQUATIONS:

Consider the following system of m linear equations in n unknown $x_1, x_2 ...... x_n$:

$$\left.\begin{array}{l} a_{11} x_1 + a_{12} x_2 + ...a_{1n} x_n = b_1 \\ a_{21} x_1 + a_{22} x_2 + ...a_{2n} x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1} x_1 + a_{m2} x_2 + ...a_{mn} x_n = b_m \end{array}\right\} \quad ...(1)$$

In the matrix notation, this system can be written as AX = B, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ and } B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

If $B \neq O$, then the system AX= B is called a system of non-homogenous linear equations, whereas AX=O is called a system of homogenous linear equations. Using matrices A and B, we write a new matrix [A : B], which is called the **Augmented matrix**.

A solution of (1) is a set of values of the unknowns $x_1, x_2, ..., x_n$ satisfying simultaneously all the equations of (1). If system (1) has a solution or many solutions, it is called consistent; otherwise it is called inconsistent. Clearly, any homogenous system of equations is always consistent since $x_i = 0$ for all $i = 1, 2, ..., n$ is a solution which is known as zero solution or trivial solution.

To determine whether the system (1) is **consistent** or **in-consistent**, we compute ranks of matrices A and [A : B] by reducing these to normal forms with the help of elementary row transformations only.

17.4.1 Theorem : For a non-homogenous system AX = B if

(*i*) $r([A : B]) \neq r(A)$, the system is inconsistent, that is, it has no solution.

(*ii*) $r([A : B]) = r(A)$ = number of unknowns, the system has a unique solution, that is, the system is consistent.

(*iii*) $r([A : B]) = r(A)$ < number of unknowns, the system is consistent but has an infinite solutions.

For a homogenous system AX = O if

(*i*) $r(A)$ = number of unknowns, the system has only the trivial solution.

(*ii*) $r(A)$ < number of unknowns, the system has infinite number of non-trivial solutions.

Examples 7 : Test for consistency and, if possible, solve the following:

(*i*) $2x - y + 3z = 4$

$x + y - 3z = -1$

$5x - y + 3z = 7$

(*ii*) $2x - 3y + 7z = 5$

$3x + y - 3z = 13$

$2x + 19y - 47z = 32$

(*iii*) $4x + 2y + 3z = 0$

$6x + 3y + 7z = 0$

$2x + y + z = 0$

Solutions. (*i*) The given system of linear equations in matrix notation can be written as AX = B, where

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 1 & -3 \\ 5 & -1 & 3 \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ and } B = \begin{bmatrix} 4 \\ -1 \\ 7 \end{bmatrix}$$

$$\text{Now } [A : B] = \left[\begin{array}{ccc:c} 2 & -1 & 3 & 4 \\ 1 & 1 & -3 & -1 \\ 5 & -1 & 3 & 7 \end{array}\right]$$

$R_2 \to R_2 + R_1$, $R_3 \to R_3 - R_1$ $\sim$
$$\begin{bmatrix} 2 & -1 & 3 & \vdots & 4 \\ 3 & 0 & 0 & \vdots & 3 \\ 3 & 0 & 0 & \vdots & 3 \end{bmatrix}$$

$R_3 \to R_3 - R_2$ $\sim$
$$\begin{bmatrix} 2 & -1 & 3 & \vdots & 4 \\ 3 & 0 & 0 & \vdots & 3 \\ 0 & 0 & 0 & \vdots & 0 \end{bmatrix}$$

From this, we note that $r([A:B]) = 2$ and $r(A) = 2$. Since $r([A:B]) = r(A) = 2$ which is less than number of unknowns, so the given system is consistent and has an infinite number of solutions. The given system of equation AX=B is equivalent to

$$\begin{bmatrix} 2 & -1 & 3 \\ 3 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix}$$

which implies $2x - y + 3z = 4$

and $3x = 3$

$\Rightarrow \quad x = 1,\ y = -4 + 2x + 3z = -2 + 3z.$

Choose $z = k$, where $k \neq 0$. Then $y = 3k - 2$. Thus $x = 1$, $y = 3k - 2$, $z = k$ for all k gives an infinite number of solutions of the given system.

(*ii*) The given system in matrix notation can be written as AX = B, where

$$A = \begin{bmatrix} 2 & -3 & 7 \\ 3 & 1 & -3 \\ 2 & 19 & -47 \end{bmatrix},\ X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ and } B = \begin{bmatrix} 5 \\ 13 \\ 32 \end{bmatrix}$$

Now
$$[A:B] = \begin{bmatrix} 2 & -3 & 7 & \vdots & 5 \\ 3 & 1 & -3 & \vdots & 13 \\ 2 & 19 & -47 & \vdots & 32 \end{bmatrix}$$

$R_2 \to 2R_2 - 3R_1$, $R_3 \to R_3 - R_1$ $\sim$
$$\begin{bmatrix} 2 & -3 & 7 & \vdots & 5 \\ 0 & 11 & -27 & \vdots & 11 \\ 0 & 22 & -54 & \vdots & 27 \end{bmatrix}$$

$R_3 \rightarrow R_3 - 2R_2$

$$\sim \begin{bmatrix} 2 & -3 & 7 & \vdots & 5 \\ 0 & 11 & -27 & \vdots & 11 \\ 0 & 0 & 0 & \vdots & 5 \end{bmatrix}$$

From this, we note that r ([A : B]) = 3 and r (A) = 2. Since r ([A : B]) ≠ r(A), the given system of linear equations is inconsistent.

(*iii*) The given system in matrix notation can be written as AX = O, where

$$A = \begin{bmatrix} 4 & 2 & 3 \\ 6 & 3 & 7 \\ 2 & 1 & 1 \end{bmatrix} \text{ and } X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$\text{Now} \quad A = \begin{bmatrix} 4 & 2 & 3 \\ 6 & 3 & 7 \\ 2 & 1 & 1 \end{bmatrix}$$

$R_1 \rightarrow R_1 - 2R_3$, $R_2 \rightarrow R_2 - 3R_3$

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

$R_2 \rightarrow R_2 - 4R_1$

$$\sim \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 2 & 1 & 1 \end{bmatrix}$$

R_{23}

$$\sim \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

⇒ r (A) = 2 which less than 3, the number of unknowns. So the system is consistent and has an infinite number of non-trivial solutions. The given system AX = O is equivalent to

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

which implies $z = 0$

and $2x + y + z = 0$

$\Rightarrow$ $y = — z — 2x = —2x$. Choose $x = -k$, where $k \neq 0$. Then $y = 2k$. Thus $x = —k$, $y = 2k$, $z = 0$ for all k is an infinite set of non-zero solutions of the given system.

17.5 CAYLEY-HAMILTON THEOREM

17.5.1 Matrix polynomial :

An expression of the form $F[x] = A_o x^o + A_1 x^1 + A_2 x^2 + \ldots\ldots A_n x^n$, $A_n \neq 0$, where A_0, A_1, ... A_n are all square matrices of the same order, is called a **matrix polynomial** of degree n in the indeterminate x.

17.5.2 Characteristic matrix and polynomial

Let A be a square matrix of order n. Then the matrix polynomial of first degree $A - \lambda I$ is called the characteristic matrix of A, and the determinant of characteristic matrix
$A - \lambda I$ is called the **characteristic polynomial** of A.

For example, consider the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Then the characteristic matrix of A is

$$A - \lambda I_2 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix}$$

and the characteristic polynomial is

$$|A - \lambda I_2| = \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = (1-\lambda^2) - 4 = \lambda^2 - 2\lambda - 3.$$

17.5.3 Characteristic Equation : The equation $|A — \lambda I| = 0$ is called the characteristic equation of the square matrix A. For example, the characteristic equation of the matrix

$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ is $|A - \lambda I_2| = 0$, that is , $\lambda^2 — 2\lambda — 3 = 0$.

Theorem (Cayley-Hamilton Theorem). Every square matrix satisfies its characteristic equation, that is, if A is a square matrix of order n and

$$|A - \lambda I| = a_o + a_1\lambda + a_2\lambda^2 + \ldots\ldots + a_n\lambda^n,$$

then $a_oI + a_1A + a_2A^2 + \ldots\ldots + a_nA^n = O.$

Proof : Let B = *adj* $(A - \lambda I)$. Then the elements of B are polynomials in λ of degree $(n - 1)$ or less. Thus we can write matrix B in the form.

$$B = B_o + B_1\lambda + B_2\lambda^2 + \ldots\ldots\ldots + B_{n-1}\lambda^{n-1}$$

where $B_o, B_1, B_2, \ldots\ldots\ldots\ldots, B_{n-1}$ are all matrices of order n and whose elements are polynomials.

Now $(A - \lambda I).B = |A - \lambda I|.I$

$\Rightarrow$ $(A - \lambda I)(B_o + B_1\lambda + B_2\lambda^2 + \ldots\ldots. + B_{n-1}\lambda^{n-1}) = (a_o + a_1\lambda + \ldots\ldots.\ a_n\lambda^n)\ I,$

which on equating coefficients of various powers of λ gives

$$\left.\begin{aligned} AB_0 &= a_0 \\ AB_1 - B_0 &= a_1I \\ AB_2 - B_1 &= a_2I \\ \vdots \quad & \quad \vdots \\ B_{n-1} &= a_1I \end{aligned}\right\} \qquad \ldots(1)$$

Multiplying equation in (1) from left by I, A, A^2, ..., A^n respectively and then adding, we obtain

$$a_oI + a_1A + a_2A^2 + \ldots\ldots\ldots + a_nA^n = 0$$

This completes the proof of the theorem.

Examples 8 : Verify Cayley-Hamilton for the following matrices :

(*i*) $A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$

(*ii*) $A = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$

Solutions. (*i*) Since A is of order 3, we take unit marix 1 of order 3 and have matrix

$$A-\lambda I = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2-\lambda & -1 & 1 \\ -1 & 2-\lambda & -1 \\ 1 & -1 & 2-\lambda \end{bmatrix}$$

Now the characteristic equation of A is given by

$$0 = |A-\lambda I| = \begin{vmatrix} 2-\lambda & -1 & 1 \\ -1 & 2-\lambda & -1 \\ 1 & -1 & 2-\lambda \end{vmatrix} \text{ (expand by } R_1)$$

$$= (2-\lambda)[(2-\lambda)^2 - 1] + [-(2-\lambda) + 1] + [1 - (2-\lambda)]$$

$$= -(\lambda-2)^3 - 2 + \lambda - 1 + \lambda - 1 + \lambda$$

$$= -(\lambda^3 - 6\lambda^2 + 12\lambda - 8) + 3\lambda - 4$$

$$= -(\lambda^3 - 6\lambda^2 + 9\lambda - 4)$$

or $\quad \lambda^3 - 6\lambda^2 + 9\lambda - 4 = 0$ ...(1)

We have to verify that $A^3 - 6A^2 + 9A - 4I = O$ ...(2)

$$\text{Now } A^2 = A.A. = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix}$$

$$A^3 = A^2 . A = \begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 22 & -21 & 21 \\ -21 & 22 & -21 \\ 21 & -21 & 22 \end{bmatrix}$$

L.H.S. of (2) = $A^3 - 6A^2 + 9A - 4I$

$$= \begin{bmatrix} 22 & -21 & 21 \\ -21 & 22 & -21 \\ 21 & -21 & 22 \end{bmatrix} - 6 \begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix} + 9 \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix} - 4 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = O = \text{R.H.S. of (2)}$$

Thus Cayley-Hamilton theorem is verified.

(*ii*) Since A is of order 3, we take unit matrix 1 of order 3 and have matrix

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

$$= \begin{bmatrix} 1-\lambda & 3 & 3 \\ 1 & 4-\lambda & 3 \\ 1 & 3 & 4-\lambda \end{bmatrix}$$

Now the characteristic equation of A is given by

$$0 = |A - \lambda I| = \begin{vmatrix} 1-\lambda & 3 & 3 \\ 1 & 4-\lambda & 3 \\ 1 & 3 & 4-\lambda \end{vmatrix} \quad \text{[Expand by } C_1\text{]}$$

$= (1-\lambda)\,[(4-\lambda)^2 - 9] - [\,3(4-\lambda) - 9] + [9 - 3\,(4 - \lambda)]$

$= -(\lambda-1)\,(\lambda^2 - 8\lambda + 7) - (3 - 3\lambda) + (3\lambda - 3)$

$= -(\lambda^3 - 9\lambda^2 + 15\lambda - 7) - 6 + 6\lambda$

$= -(\lambda^3 - 9\lambda^2 + 9\lambda - 1)$ or $\lambda^3 - 9\lambda^2 + 9\lambda - 1 = 0$ ...(1)

We have to verify that $A^3 - 9A^2 + 9A - I = O$ ...(2)

$$\text{Now } A^2 = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 7 & 24 & 24 \\ 8 & 28 & 27 \\ 8 & 27 & 28 \end{bmatrix}$$

$$A^3 = A^2 . A = \begin{bmatrix} 7 & 24 & 24 \\ 8 & 28 & 27 \\ 8 & 27 & 28 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 55 & 189 & 189 \\ 63 & 217 & 216 \\ 63 & 216 & 217 \end{bmatrix}$$

and L.H.S. of (2) = $A^3 - 9A^2 + 9A - I$

$$= \begin{bmatrix} 55 & 189 & 189 \\ 63 & 217 & 216 \\ 63 & 216 & 217 \end{bmatrix} - 9 \begin{bmatrix} 7 & 24 & 24 \\ 8 & 28 & 27 \\ 8 & 27 & 28 \end{bmatrix} + 9 \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = O = \text{ R.H.S. of (2)}$$

Thus Cayley-Hamilton theorem is verified.

17.5.4 Inverse of a matrix by using Cayley-Hamilton Theorem :

Let A be a non-singular matrix of order n. Since by Cayley Hamilton theorem every matrix satisfies its characteristic eqaution, we have, $f(A) = 0$, where $f(\lambda) = |A - \lambda I|$ is the characteristic polynomial of A. Since A is non-singular-matrix of order n, $f(\lambda)$ will be of degree n and of the form:

$$f(\lambda) = a_0 + a_1\lambda + ... + a_n\lambda^n$$

Thus we have $f(A) = a_0 I + a_1 A + ... + a_n A^n = O$

or $$I = -\frac{a_1}{a_0} A - \frac{a_2}{a_0} A^2 - ..... - \frac{a_n}{a_0} A^n \qquad ...(1)$$

Since $|A| \neq 0$, A^{-1} exists. Multiply (1) by A^{-1} from left side, we obtain

$$A^{-1} = -\frac{a_1}{a_0} I - \frac{a_2}{a_0} A - \frac{a_3}{a_0} A^2 \ldots - \frac{a_n}{a_0} A^{n-1}$$

Examples 9. Using Cayley-Hamilton theorem, find the inverse of following matrices:

(*i*) $$A = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$$

(*ii*) $$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Solutions. (*i*) As calculated in Example 8(*ii*) above, the characteristic equation of given matrix A is $\lambda^3 - 9\lambda^2 + 9\lambda - 1 = 0$. Using Cayley-Hamilton theorem, we have

$$A^3 - 9A^2 + 9A - I = O \qquad \ldots(1)$$

Since $|A| = 1 \neq 0$, A^{-1} exists. Multiplying (1) by A^{-1} from left side, we obtain

$$A^2 - 9A + 9I - A^{-1} = O$$

which gives $A^{-1} = A^2 - 9A + 9I$

$$= \begin{bmatrix} 7 & 24 & 24 \\ 8 & 28 & 27 \\ 8 & 27 & 28 \end{bmatrix} - 9 \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix} + 9 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

(*ii*) The characteristic equation of A is given by

$$0 = |A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 & 0 \\ 2 & -1-\lambda & 0 \\ 0 & 0 & -1-\lambda \end{vmatrix}$$

which on simplification gives $\lambda^3 + \lambda^2 - 5\lambda - 5 = 0$

Using Cayley-Hamilton theorem, we have

$$A^3 + A^2 - 5A - 5I = O \qquad ...(1)$$

Since $|A| = 5 \neq 0$, A^{-1} exists. So multiplying (1) by A^{-1} from left side, we obtain

$A^2 + A - 5I - 5A^{-1} = O$ which gives $A^{-1} = \frac{1}{5}A^2 + \quad A — I$

$$= \qquad + \frac{1}{5} \qquad — \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

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

Examples 10. Using Cayley-Hamilton theorem, find the values of

(*i*) A^8, where $A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$

(*ii*) A^{-2}, where $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

(*iii*) A^{-2}, where $A = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$

Solutions. (*i*) The characteristic equation of A is

$$0 = |A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 \\ 2 & -1-\lambda \end{vmatrix}$$

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

Using Cayley Hamilton theorem, we have $A^2 - 5I = O$ or $A^2 = 5I$ ...(1)

Now using (1), we have

$$A^8 = (A^2)^4 = (5I)^4 = 625\, I^4 = 625\, I \qquad (\because \; I^4 = I)$$

(ii) As already calculated in Example 9(ii) above, we have

$$A^3 + A^2 - 5A - 5I = O \qquad ...(1)$$

Since $|A| = 5 \neq 0$, so A^{-1} and $(A^{-1})^2 = A^{-2}$ also exists. Multiplying (1) by A^{-2} on both sides, we have $A + I - 5A^{-1} - 5A^{-2} = O$ which gives

$$A^{-2} = \frac{1}{5}A + \quad 1 - A^{-1}$$

$$= \qquad + \frac{1}{5} \qquad - \frac{1}{5}$$

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

(iii) As derived in Example 9 (i) above, we have [illegible]

$$A^3 - 9A^2 + 9A - I = O \qquad ...(1)$$

and
$$A^{-1} = \begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad ...(2)$$

Now
$$A^{-2} = A^{-1}.\, A^{-1} = \begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 55 & -24 & -24 \\ -8 & 4 & 3 \\ -8 & 3 & 4 \end{bmatrix}$$

17.6 SUMMARY

In this lesson, we studied techniques to calculate rank of matrices and inverses of non-singular matrices by rank method. Consistency of a system of linear equation is also discussed. We studied Cayley-Hamilton theorem and used it to find inverse of non-singular matrices. Related examples are given on these concepts.

17.7 EXERCISE

1. Using elementary row and column transformations, compute ranks of the following matrices:

(*i*) $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ -3 & -6 & -3 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 4 & 8 & 12 \end{bmatrix}$

(*iii*) $\begin{bmatrix} 2 & -1 & 3 & -2 & 4 \\ 4 & -2 & 5 & 1 & 7 \\ 2 & -1 & 1 & 8 & 2 \end{bmatrix}$ (*iv*) $\begin{bmatrix} 1 & 4 & -2 & 1 \\ 5 & 2 & 4 & 3 \\ 2 & -1 & 3 & 1 \end{bmatrix}$

(*v*) $\begin{bmatrix} 25 & 31 & 17 & 43 \\ 75 & 94 & 53 & 132 \\ 75 & 94 & 54 & 134 \end{bmatrix}$

2. Using elementary row transformations, reduce the following matrices to upper as well as lower triangular forms and hence find their ranks :

(*i*) $\begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & -1 \\ 5 & 2 & 3 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 0 & 1 & 0 & 2 \\ 0 & 3 & 1 & 4 \\ 4 & 1 & 0 & 2 \\ 1 & -1 & 2 & -3 \end{bmatrix}$

3. For the following matrices, find two non-singular matrices P and Q such that PAQ is in normal form :

(*i*) $\begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & -1 \\ 5 & 2 & 3 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 1 & 2 & 3 & 2 \\ 2 & 3 & 5 & 1 \\ 1 & 3 & 4 & 5 \end{bmatrix}$

4. Reduce the following matrices to normal forms, and hence find their ranks :

(*i*) $\begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 3 & 2 & 3 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 1 & 2 & 3 & 2 \\ 2 & 3 & 5 & 1 \\ 1 & 3 & 4 & 5 \end{bmatrix}$

5. Using elementary row transformations only, compute the inverse of the following matrices:

(*i*) $\begin{bmatrix} 1 & 2 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 2 & 1 & -1 \\ 0 & 2 & 1 \\ 5 & 2 & -3 \end{bmatrix}$

6. Using elementary column transformations only, compute the inverse of the following matrices :

(*i*) $\begin{bmatrix} 1 & 2 & -1 \\ -4 & -7 & 4 \\ -4 & -9 & 5 \end{bmatrix}$ (*ii*) $\frac{1}{4}\begin{bmatrix} -1 & 3 & -4 \\ 3 & -1 & 0 \\ -1 & -1 & 4 \end{bmatrix}$

7. Test for consistency and, if possible, solve the following systems of linear equations :

(*i*) $5x + 3y + 7z = 4$
$3x + 26y + 2z = 9$
$7x + 2y + 10z = 5$

(*ii*) $x + 4y + 9z = 6$
$x + 2y + 3z = 4$
$x + y + z = 3$

(*iii*) $3x + y + 2z = 3$
$2x - 3y - z = -3$
$x + 2y + z = 4$

(*iv*) $2x + 6y = -11$
$6x + 20y - 6z = -3$
$6y - 18z = -1$

8. Verify Cayley-Hamilton theorem for the following matrices :

(*i*) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

9. Using Cayley-Hamilton theorem, find the inverse of the following matrices :

(*i*) $\begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 1 & -1 & 1 \\ 4 & 1 & 0 \\ 8 & 1 & 1 \end{bmatrix}$

ANSWERS

1(*i*) 1 (*ii*) 1 (*iii*) 2 (*iv*) 2 (*v*) 3

2. (*i*) 3 (*ii*) 4

4. (*i*) 3 (*ii*) 2

5. (*i*) $\frac{1}{4}$ (*ii*) $\begin{bmatrix} 8 & -1 & -3 \\ -5 & 1 & 2 \\ 10 & -1 & -4 \end{bmatrix}$

6. (*i*) $\begin{bmatrix} 1 & -1 & 1 \\ 4 & 1 & 0 \\ 8 & 1 & 1 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 1 & 2 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix}$

7. (*i*)Consistent; $x = \frac{1}{11}(7—16k)$, $y =$ $(k+3)$, $z = k$, where k is arbitrary.

(*ii*) Consistent; $x = 2, y = 1, z$ = [illegible]

(*iii*) Consistent : $x = 1, y = 2, z$ = [illegible]

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

(*iv*) In-consistent

9. (*i*) (*ii*) $\begin{bmatrix} 1 & 2 & -1 \\ -4 & -7 & 4 \\ -4 & -9 & 5 \end{bmatrix}$

17.8. SUGGESTED READING/REFERENCES :

1. "A text book of Matrices" by Shanti Narayan and P.K.Mittal, S. Chand & Company Ltd.
2. "Theory of Matrices" by B.S. Vatssa, Wiley Easterns Ltd.
3. "Engineering Mathematics-II" by Bhopinder Singh, Malhotra Brothers, Jammu.

17.9. MODEL TEST PAPER

1. Using elementary row and column transformations, compute ranks of the following matrices:

(*i*) $\begin{bmatrix} 3 & 1 & 2 \\ 2 & 3 & 4 \\ -1 & 2 & 2 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 1 & 2 & -2 \\ 2 & 4 & 6 \\ 4 & 6 & 12 \end{bmatrix}$

(*iii*) $\begin{bmatrix} 2 & -1 & 3 & -2 & 4 \\ 4 & -2 & 5 & 1 & 7 \\ 2 & -1 & 1 & 8 & 2 \end{bmatrix}$ (*iv*) $\begin{bmatrix} 24 & 19 & 36 & -38 \\ 49 & 40 & 73 & -80 \\ 73 & 59 & 98 & -118 \end{bmatrix}$

(*v*) $\begin{bmatrix} 0 & 1 & 2 & 3 \\ 2 & 2 & 4 & 3 \\ 3 & 3 & 2 & 1 \\ 5 & 6 & 8 & 7 \end{bmatrix}$ (*vii*) $\begin{bmatrix} 1 & -1 & 2 & 3 \\ 4 & 1 & 0 & 2 \\ 0 & 3 & 0 & 4 \\ 0 & 1 & 0 & 2 \end{bmatrix}$

2. Using elementary column transformations, reduce the following matrices to upper as well as lower triangular forms and hence find their ranks :

(*i*) $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 3 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 1 & 4 & 2 \\ 2 & 1 & 3 \\ 3 & 7 & 5 \end{bmatrix}$

3. For the following matrices, find non-singular matrices P and Q such that PAQ is in normal form :

(*i*) $\begin{bmatrix} 1 & 4 & 2 \\ 2 & 1 & 3 \\ 3 & 7 & 5 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 3 & 4 & 1 \\ 4 & 1 & 2 & 0 \\ 1 & -1 & 0 & 2 \end{bmatrix}$

4. Reduce the following matrices to normal forms, and hence find their ranks :

(*i*) $\begin{bmatrix} 1 & -1 & 0 & 3 \\ 0 & 1 & 0 & -1 \\ 1 & 2 & 3 & 0 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 1 & 7 & 2 \\ 3 & 2 & 1 \\ -1 & 0 & 4 \end{bmatrix}$

5. Using elementary row transformations only, compute the inverse of the following matrices:

(*i*) $\begin{bmatrix} 1 & -1 & 0 \\ -2 & 3 & -4 \\ -2 & 3 & -3 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$

6. Using elementary colum transformation only, compute the inverse of the following matrices :

(*i*) $\begin{bmatrix} 1 & -1 & 0 \\ -2 & 3 & -4 \\ -2 & 3 & -3 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 1 & 2 & 5 \\ 3 & 1 & 4 \\ 1 & 1 & 2 \end{bmatrix}$

7. Test for consistency and, if possible, solve the following systems of linear equations :

(*i*) $x + 2y + 2z = 2$
$2x + y - 2z = 7$
$2x - 2y + z = 1$

(*ii*) $x + y + z = 1$
$x - 3y - 2z = 2$
$2x - y + z = 4$

(*iii*) $2x + y - z = 4$
$x - y + 2z = -2$
$-x + 2y - z = 2$

(*iv*) $x - y + z = 1$
$2x + y - z = 2$
$5x - 2y + 2z = 5$

B.A. III Semester	**Unit-5**
Course No. : UGM 0301	**Lesson 18**

MATRICES

EIGEN VALUES AND EIGEN VECTORS OF MATRICES

By: Dr. Kamaljeet Kour

18.1 INTRODUCTION :

Let A be a non-zero square matrix of order n and consider the equation

$$AX = \lambda X \qquad ...(1)$$

where λ is a scalar. Clearly X = O is its solution for any value of λ. A value of scalar λ for which the equation (1) has a non-zero solution, that is, $X \neq O$ is called an eigen value of A. This non-zero vector X which satisfies the equation (1) is called the corresponding eigen vector of A. The set of all eigen values of A is called the spectrum of A. The largest of the absolute values of the eigen values of A is called the spectral radius of A. The problem of finding X corresponding to λ of equation (1) is known as the eigen value problem. This has further many applications and beautiful results in the area of analysis and operator theory.

18.2. OBJECTIVES :

The main objective of this lesson is to study eigen values and eigen vectors of square matrices of various orders. Also important properties alongwith solved examples are given for better understanding of the students.

18.3. VECTORS :

By $\mathbb{R}$, we denote the real line and by $\mathbb{R}^2$, we mean $\mathbb{R} \times \mathbb{R}$, the cartesian product of $\mathbb{R}$ with itself. Similarly, we have $\mathbb{R}^3$ and in general $\mathbb{R}^n$ where n = 1, 2, 3, ... Let $\mathbb{C}$ denote the set of all complex numbers. Then we have $\mathbb{C} \times \mathbb{C} = \mathbb{C}^2$, and in general $\mathbb{C}^n$. The sets $\mathbb{R}^n$ or $\mathbb{C}^n$ are called n-spaces or n-dimensional vector spaces and their elements are known as **vectors**.

An element X in $\mathbb{R}^n$ or $\mathbb{C}^n$ is of the form $(a_1, a_2, \ldots a_n)$, called n-tuple, which can be denoted by the column matrix : $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$

18.3.1 Linearly dependent and independent vectors :

(*i*) A set of m vectors $X_1, X_2, \ldots X_m$ is said to linearly dependent if there exists scalar (constants) $c_1, c_2, \ldots c_m$ not all zero such that

$$c_1 X_1 + c_2 X_2 + \ldots + c_m X_m = 0 \qquad \ldots(1)$$

The sum $c_1 X_1 + c_2 X_2 + \ldots + c_m X_m$ is called a linear combination of m vectors $X_1, X_2, \ldots, X_m$.

Suppose $c_1 \neq 0$, then form (1), we obtain

$$X_1 = a_2 X_2 + a_3 X_3 + \ldots + a_m X_m,$$

where $a_2 = \frac{-c_2}{c_1}$, $a_3 = \frac{-c_3}{c_1}$, ..., $a_m = \frac{-c_m}{c_1}$.

In this way, we see that the vector X_1, being expressed as a linear combination of vectors $X_2, X_3, \ldots X_m$ depends upon them.

(*ii*) A set of m vectors $X_1, X_2, \ldots, X_3$ is called linearly independent if $c_1 X_1 + c_2 X_2 + \ldots + c_m X_m = 0$ implies that $c_i = 0$ for all $i = 1, 2, \ldots, m$.

Examples 1 : Let us verify the linear dependence or independence of the following vectors in $\mathbb{R}^3$:

(1) Take $X_1 = (4, -2, 6)$, $X_2 = (6, -3, 9)$.

Suppose there are scalars c_1 and c_2 such that $c_1 X_1 + c_2 X_2 = 0$.

This implies that $c_1 (4, -2, 6) + c_2 (6, -3, 9) = (0, 0, 0)$

or $(4c_1 + 6c_2, -2c_1 - 3c_2, 6c_1 + 9c_2) = (0, 0, 0)$

$\Rightarrow$ $4c_1 + 6c_2 = 0, -2c_1 - 3c_2 = 0, 6c_1 + 9c_2 = 0$

But these equations are same as

$$2c_1 + 3c_2 = 0 \text{ which gives } 2c_1 = — 3c_2.$$

Take $c_2 = —2$. Then $c_1 = 3$.

Note that $3X_1 — 2X_2 = 0$.

So given vectors X_1, X_2 are linerly dependent.

(2) Let us take three vectors :

$X_1 = (2. —3, 1), X_2 = (3, —1, 5), X_3 = (1, —4, 3)$.

Suppose that there exists three scalars c_1, c_2, c_3 such that $c_1X_1 + c_2X_2 + c_3X_3 = 0$.

This implies that

$$c_1 (2, —3, 1) + c_2 (3, — 1, 5) + c_3 (1, — 4, 3) = (0, 0, 0)$$

or $(2c_1 + 3c_2 + c_3 , —3c_1 — c_2 — 4c_3 , c_1 + 5c_2 + 3c_3) = (0, 0, 0)$

$\Rightarrow$ $2c_1 + 3c_2 + c_3 = 0$ ...(1)

$—3c_1 — c_2 — 4c_3 = 0$ ...(2)

$c_1 + 5c_2 + 3c_3 = 0$ ...(3)

From (3), we obtain $c_1 = —5c_2 — 3c_3$. Substitute this value of c_1 in (1) and (2), we have

$$—7c_2—5c_3 = 0$$

$$14c_2 + 5c_3 = 0$$

which is a homogenous system of linear equations and its only solution is $c_2= 0, c_3 = 0$.

Put these values in (3), we obtain $c_1 = 0$.

Thus $c_1 X_1 + c_2X_2 + c_3 X_3 = 0$ implies $c_1 = 0, c_2 = 0, c_3 = 0$.

Hence the vectors X_1, X_2 , X_3 are linearly independent.

18.4. Eigen values :

Let A be a square matrix. Then the roots of the characteristic equation $| A-\lambda I | = 0$ are called eigen values are characteristic roots or latent roots of the matrix A.

If the matrix A is of order n, then the characteristic equation $|A-\lambda I| = 0$ has n roots which may or may not be all distinct.

For example, consider the matrix

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}.$$

Then $|A - \lambda I| = 0 \quad \Rightarrow \quad \lambda^3 - 7\lambda^2 + 11\lambda - 5 = 0 \quad \Rightarrow \quad \lambda = 1, 1, 5$

$\therefore$ the characteristic roots or eigen values of A are 1, 1, 5.

18.4.1. Properties of Eigen values :

(1) Any square matrix A and its transpose A′ have the same eigen values.

(2) The sum of the eigen values of a matrix is equal to the trace of the matrix.

(3) The product of the eigen values of a matrix A is equal to the determinant of A.

(4) If $\lambda_1, \lambda_2, \lambda_3, ..., \lambda_n$ are eigen values of A, then the eigen values of

(*i*) kA are $k\lambda_1, k\lambda_2, ..., k\lambda_n$;

(*ii*) A^m are $\lambda_1^m, \lambda_2^m, ..., \lambda_n^m$;

and (*iii*) A^{-1} are $\lambda_1^{-1}, \lambda_2^{-1}, ..., \lambda_n^{-1}$, where $|A| \neq 0$

(5) The eigen values of a triangular matrix or diagonal matrix are just the diagonal elements of the matrix.

(6) If λ is an eigen value of a non-singular matrix A, then $\frac{|A|}{\lambda}$ is an eigen value of adj (A).

(7) If λ is an eigen value of A, then the eigen value of $A + kI$ is $\lambda + k$.

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

Then the characteristic equation of A is $|A-\lambda I| = 0$, which gives $\lambda^3 + \lambda^2 - 5\lambda - 5 = 0$

$\Rightarrow \qquad (\lambda + 1)(\lambda^2 - 5) = 0$

$\Rightarrow \qquad \lambda = -1, \sqrt{5}, -\sqrt{5},$

$\therefore \qquad -1, \sqrt{5}, -\sqrt{5}$ are the eigen value of A. Also its inverse, as calculated in Example 9(*ii*) of lesson 17, is given by

$$A^{-1} = \frac{1}{5}\begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -5 \end{bmatrix}.$$

So the characteristic equation of A^{-1} is $|A^{-1} - \lambda I| = 0$

or $\qquad \left| \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -5 \end{bmatrix} - 5\lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right| = 0$

or $\qquad \begin{vmatrix} 1-5\lambda & 2 & 0 \\ 2 & -1-5\lambda & 0 \\ 0 & 0 & -5-5\lambda \end{vmatrix}$

Expanding by R_3, we obtain $25(1+\lambda)(1 - 5\lambda^2) = 0$

$\Rightarrow \qquad \lambda = -1, \frac{1}{\sqrt{5}}, \frac{-1}{\sqrt{5}}$ as the eigen values of A^{-1}.

This example also verifies the property 4 (*iii*) of eigen values as stated in 18.4.1.

Now $\quad A^2 = A.A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

$$= \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and the characteristic equation of A^2 is $|A^2 - \lambda I| = 0$

$\Rightarrow \qquad \begin{vmatrix} 5-\lambda & 0 & 0 \\ 0 & 5-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = 0$

$\Rightarrow \quad (5-\lambda)\,(5-\lambda)\,(1-\lambda) = 0$

$\Rightarrow \quad \lambda = 1, 5, 5$ as eigen values of A^2.

Thus we note that the eigen values of A^2 are 1, 5, 5 which are square of the eigen values of A.

Also note that $|A| = 5$ which is same as the product of eigen values $-1 . \sqrt{5} . -\sqrt{5}$ of A. Again note that the sum of eigen values of A is — 1, which is same as the trace of A.

In a same way, we can verify other properties of eigen values as stated above for a given matrix.

Examples 3. Find all the eigen values of the following matrices :

(*i*) $\quad A^3$, where $A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$

(*ii*) $\quad A^5$, where $A = \begin{bmatrix} 4 & 1 & -1 \\ 2 & 3 & -1 \\ -2 & 1 & 5 \end{bmatrix}$

(*iii*) $\quad A^5$, where $A = \begin{bmatrix} -1 & 2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$

Solutions : (*i*) The characteristic equation of A is given by

$$0 = |A - \lambda I| = \begin{vmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & -2 & -\lambda \end{vmatrix}$$

$= -(\lambda^3 + \lambda^2 - 21\lambda - 45)$, on expansion

$\Rightarrow \quad \lambda^3 + \lambda^2 - 21\lambda - 45 = 0$

$\Rightarrow \quad (\lambda + 3)\,(\lambda^2 - 2\lambda - 15) = 0$

$\Rightarrow \quad (\lambda + 3)(\lambda + 3)(\lambda - 5) = 0$

$\Rightarrow \quad \lambda = -3, -3, 5$ as eigen values of A.

So, by property 4(*ii*) stated above in 18.4.1, the eigen values of A^3 are

$$\lambda^3 = (-3)^3, (-3)^3, 5^3 = -27, -27, 125.$$

(*ii*) The characteristic equation of A is given by

$$0 = |A - \lambda I| = \begin{vmatrix} 4-\lambda & 1 & -1 \\ 2 & 3-\lambda & -1 \\ -2 & 1 & 5-\lambda \end{vmatrix}$$

$= -(\lambda^3 - 12\lambda^2 + 44\lambda - 48)$, on expansion

$\Rightarrow \quad \lambda^3 - 12\lambda^2 + 44\lambda - 48 = 0$

$\Rightarrow \quad (\lambda - 2)(\lambda^2 - 10\lambda + 24) = 0$

$\Rightarrow \quad (\lambda - 2)(\lambda - 4)(\lambda - 6) = 0$

$\Rightarrow \quad \lambda = 2, 4, 6$ as eigen values of A.

So, by property 4(*ii*) stated above in 18.4.1, the eigen values of A^5 are

$\lambda^5 = 2^5, 4^5, 6^5 = 32, 1024, 7776.$

(*iii*) The characterstic equation of A is given by

$0 = |A - \lambda I|$

$= -(\lambda^3 - \lambda^2 - 5\lambda + 5)$

$\Rightarrow \quad (\lambda - 1)(\lambda^2 - 5) = 0.$

$\Rightarrow \quad \lambda = 1, \sqrt{5}, -\sqrt{5}$ as eigen values of A.

So, by property 4(*ii*) stated above in 18.4.1. the eigen values of A^4 are 1, 25, 25.

18.5. Eigen vectors or Characteristic vectors :

Let A be a square matrix of order *n*. If corresponding to an eigen value λ of A, there exists a non-zero vector $X \in \mathbb{R}^n$ (or $\mathbb{C}^n$) such that $(A - \lambda I)X = 0$, then X is called the eigen vector of A.

From this definition, it is clear that corresponding to a non-zero vector X there exists a scalar λ such that

$$(A — \lambda I)\, X = 0 \qquad ...(1)$$

or $$AX = \lambda X. \qquad ...(2)$$

Then λ is the eigen value of A and vice versa.

Example 4 : Verify the following statements :

(*i*) $X = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigen vector of $A = \begin{bmatrix} 2 & -5 \\ -2 & -1 \end{bmatrix}$

(*ii*) $X = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigen vector of $A = \begin{bmatrix} -2 & 5 \\ 2 & 1 \end{bmatrix}$

Solutions : We have to check that $AX = \lambda X$ for some λ.

(*i*) $AX = \begin{bmatrix} 2 & -5 \\ -2 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \end{bmatrix} = -3 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda X$

where $\lambda = -3$. This shows that -3 is an eigen value of A and the corresponding vector is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

So the given statement is true.

(*ii*) $AX = \begin{bmatrix} 2 & -5 \\ -2 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -8 \\ -4 \end{bmatrix} = -4 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} \neq -4\,X$

So this statement is not true.

Results : (*i*) The matrix $A - \lambda I$ in (1) of 18.5 above is singular.

(*ii*) If there are *n* distinct eigen values of A, then we have n linearlyindependent eigen vectors.

(*iii*) If eigen values are repeated (equal), then it may or may not be possible to obtain linearly independent eigen vectors corresponding to repeated eigen values.

(*iv*) From (2), we note that if X satisfies (2), then *k*X also satisfies (2), where *k* is non-zero scalar. That is, eigen vector X_i corresponding to an eigen value λ_i is not unique but it is one of the vectors kX_i.

(*v*) An eigen vector cannot correspond to two different eigen values.

(*vi*) Consider the homogenous system AX = 0 of linear equations, where A is a square matrix of order *n*. Let *r* (A) = *r*. Then the number of independent solutions of the system is equal to *n* — *r*.

Examples 5. Find all the eigen values and eigen vectors of the following matrices :

(*i*) $A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$

(*ii*) $B = \begin{bmatrix} 0 & 0 & 1 \\ 8 & 0 & 2 \\ 0 & -1 & 5 \end{bmatrix}$

(*iii*) $C = \begin{bmatrix} -1 & 2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$

Solution. (*i*) The characteristic equation of given matrix A is

$$0 = |A - \lambda I| = \begin{vmatrix} 6-\lambda & -2 & 2 \\ -2 & 3-\lambda & -1 \\ 2 & -1 & 3-\lambda \end{vmatrix}$$

which on expanding by R_1 and after simplification gives

$$(2 - \lambda)(\lambda - 2)(\lambda - 8) = 0$$

$\Rightarrow \qquad \lambda = 2, 2, 8.$

Thus the eigen values of A are $\lambda_1 = 8, \lambda_2 = \lambda_3 = 2$.

Next let us determine the corres-ponding eigen vectors of A.

For $\lambda_1 = 8$, the corresponding eigen vector $X_1 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$ is given by the non-zero solution of the matrix equation :

$$(A - \lambda_1 I) X_1 = O \text{ or } (A - 8I) X_1 = O$$

or $$\begin{bmatrix} -2 & -2 & 2 \\ -2 & -5 & -1 \\ 2 & -1 & -5 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Applying row transformations $R_2 \to R_2 - R_1$ and $R_3 \to R_3 + R_1$, we have

$$\begin{bmatrix} -2 & -2 & 2 \\ 0 & -3 & -3 \\ 0 & -3 & -3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Applying row transformation $R_3 \to R_3 - R_2$, we have

$$\begin{bmatrix} -2 & -2 & 2 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad ...(1)$$

which is of the form $PX_1 = 0$

Now order of P is 3 and $r(A) = 2$. So the number of independent solutions of system (1) is $3 - 2 = 1$. From (1). we obtain

$-2a - 2b + 2c = 0 \quad \Rightarrow \quad b = -c$

$-3b - 3c = 0 \quad\quad a = -b + c = 2c$

Choose $c = 1$. Then $b = -1$ and $a = 2$.

$\therefore$ the eigen vector corresponding to eigen value $\lambda_1 = 8$ is $X_1 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$

For $\lambda_2 = \lambda_3 = 2$, the corresponding eigen vector $X = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$ is given by the non-zero solution of the matrix equation :

$$(A - 2I)\, X = O$$

or $$\begin{bmatrix} 4 & -2 & 2 \\ -2 & 1 & -1 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Applying row transformations $R_1 \to R_1 - 2R_3$ and $R_2 \to R_2 + R_3$, we have

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad ...(2)$$

which is of the form PX = 0. Here order of P is 3 and r (P) = 1. So the number of independent solutions of the system (2) are 3 — 1= 2.

From (2), we have $2a - b + c = 0$ or $2a = b - c$. Choose first $b = 0$ and $c = 2$. Then $a = -1$.

Again choose $c = 0$ and $b = 2$. Then $a = 1$.

So eigen vectors corresponding to $\lambda_2 = \lambda_3 = 2$ are $X_2 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$ and $X_3 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

(*ii*) The characteristic equation of the given matrix B is

$$0 = | B - \lambda I | = \begin{vmatrix} -\lambda & 0 & 1 \\ 8 & -\lambda & 2 \\ 0 & -1 & 5-\lambda \end{vmatrix}$$

Expanding by R_1, we obtain $-\lambda^3 + 5\lambda^2 - 2\lambda - 8 = 0$

$\Rightarrow \quad \lambda^3 - 5\lambda^2 + 2\lambda + 8 = 0$

$\Rightarrow \quad (\lambda + 1)(\lambda^2 - 6\lambda + 8) = 0$

$\Rightarrow \quad (\lambda + 1)(\lambda - 2)(\lambda - 4) = 0$

$\Rightarrow \quad \lambda = -1, 2, 4$

Thus the eigen values of B are $\lambda_1 = -1, \lambda_2 = 2, \lambda_3 = 4$.

Next let us determine the corresponding eigen vectors of B.

For λ_1 = —1, the corresponding eigen vector $X_1 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$ is given by the non-zero solution of the matrix equation :

$$(B - \lambda_1 I)\, X_1 = O$$

or $$(B + I)\, X_1 = O$$

or $$\begin{bmatrix} 1 & 0 & 1 \\ 8 & 1 & 2 \\ 0 & -1 & 6 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Applying $R_2 \to R_2 - 8R_1 + R_3$, we have

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & -1 & 6 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad ...(1)$$

This system is of the form $PX_1 = 0$. Here order of P is 3 and $r(P) = 2$. So the number of independent solutions of system (1) is 3 — 2 =1. From (1) we obtain

$$a + c = 0 \quad \text{and} \quad -b + 6c = 0$$

$\Rightarrow$ $a = -c$ and $b = 6c$

Choose $c = 1$. Then $a = -1$ and $b = 6$.

$\therefore$ the eigen vector corresponding to λ_1 = —1 is $X_1 = \begin{bmatrix} -1 \\ 6 \\ 1 \end{bmatrix}$

For $\lambda_2 = 2$, the corresponding eigen vector $X_2 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$ is given by the non-zero solution of the matrix equation: $(B - \lambda_2 I)\, X_2 = O$

or $(B - 2I)\, X_2 = O$ or $\begin{bmatrix} -2 & 0 & 1 \\ 8 & -2 & 2 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Applying $R_2 \rightarrow R_2 + 4R_1 — 2R_3$, we have

$$\begin{bmatrix} -2 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad ...(2)$$

This is of the form $AX_2 = 0$, where order of A is 3 and $r(A) = 2$.

So the number of independent solutions of system (2) is 3 — 2 = 1.

From (2), we obtain

$$-2a + c = 0 \text{ and } -b + 3c = 0$$

$$\Rightarrow \quad 2a = c \text{ and } b = 3c$$

Choose $c = 2$. Then $a = 1$ and $b = 6$.

Thus the eigen vector corresponding to $\lambda_2 = 2$ is

$$X_2 = \begin{bmatrix} 1 \\ 6 \\ 2 \end{bmatrix}$$

In a similar way, it can be found that the eigen vector corresponding to $\lambda_3 = 4$ is

$$X_3 = \begin{bmatrix} 1 \\ 4 \\ 4 \end{bmatrix}$$

(*iii*) The characteristic equation of given matrix C is

$$0 = |C - \lambda I| = \begin{vmatrix} -1-\lambda & 2 & -2 \\ 1 & 2-\lambda & 1 \\ -1 & -1 & -\lambda \end{vmatrix}$$

Expanding by R_1, we obtain

$$-(1 + \lambda)(—2\lambda + \lambda^2 + 1) — 2(—\lambda + 1) — 2(—1 + 2 — \lambda) = 0$$

which on simplification gives $(\lambda — 1)(5 — \lambda^2) = 0$

$$\Rightarrow \quad \lambda = 1, \sqrt{5}, -\sqrt{5}.$$

Thus the eigen values of C are $\lambda_1 = 1, \lambda_2 = \sqrt{5}, \lambda_3 = -\sqrt{5}$.

Next, let us determine the corresponding eigen vectors of C.

For $\lambda_1 = 1$, the corresponding eigen vector $X_1 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$ is non-zero solution of the matrix equation :

$$(C - \lambda_1 I) X_1 = O$$

or $$(C - I) X_1 = O$$

or $$\begin{bmatrix} -2 & 2 & -2 \\ 1 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Applying row transformations $R_1 \to R_1 + 2R_2$ and $R_3 \to R_3 + R_2$, we have

$$\begin{bmatrix} 0 & 4 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{...(1)}$$

This is of the form $AX_1 = 0$, where order of A is 3 and $r(A) = 2$.

So the number of independent solutions of system (1) is $3 - 2 = 1$.

From (1), we obtain

$$4b = 0 \text{ and } a + b + c = 0$$

$\Rightarrow$ $b = 0$ and $a = -b - c = -c$.

Choose $c = -1$. Then $a = 1$.

Thus the eigen vector corresponding to $\lambda_1 = 1$ is $X_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$

For $\lambda_3 = -\sqrt{5}$, the corresponding eigen vector $X_3 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$ is the non-zero solution of the matrix equation : $(C - \lambda_3 I) X_3 = O$

or $$\begin{bmatrix} -1+\sqrt{5} & 2 & -2 \\ 1 & 2+\sqrt{5} & 1 \\ -1 & -1 & \sqrt{5} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Applying $R_1 \to R_1 + (-1+\sqrt{5}) R_3$ and $R_2 \to R_2 + R_3$, we have

$$\begin{bmatrix} 0 & 3-\sqrt{5} & 3-\sqrt{5} \\ 0 & 1+\sqrt{5} & 1+\sqrt{5} \\ -1 & -1 & \sqrt{5} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Applying $R_1 \to \frac{1}{3-\sqrt{5}} R_1$ and $R_2 \to \frac{1}{1+\sqrt{5}} R_2$, we have

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & \sqrt{5} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Applying $R_1 \to R_1 - R_2$, we have

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & -1 & \sqrt{5} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{...(2)}$$

This is of the form $BX_3 = 0$, where order of B is 3 and $r(B) = 2$.

So the number of independent solutions of system (2) is 3 — 2 = I.

From (2), we have

$b + c = 0$ and $-a - b + \sqrt{5}c = 0$

$\Rightarrow \quad b = -c$ and $a = (1 + \sqrt{5})\, c$

Choose $c = 1$. Then $b = -1$ and $a = 1 + \sqrt{5}$.

Thus the eigen vector corresponding to λ_3 is $X_3 = \begin{bmatrix} 1+\sqrt{5} \\ -1 \\ 1 \end{bmatrix}$.

In a similar way, it can be found that the eigen vector corresponding to $\lambda_2 = \sqrt{5}$ is $X_2 = \begin{bmatrix} 1-\sqrt{5} \\ -1 \\ 1 \end{bmatrix}$

In the following theorem, we prove some of the properties of eigen values for a given square matrix A :

18.4.2 Theorem : Let A be a square matrix. Then

(*i*) the eigen values of A and its transpose A′ are same

(*ii*) the sum of the eigen values of A is equal to the trace of A.

Further if λ is an eigen value of A, then

(*iii*) $\lambda + k$ is an eigen value of A + k I.

(*iv*) λ^2 is an eigen value of A^2

(*v*) λ^{-1} is an eigen value of A^{-1} provided A is non-singular.

(*vi*) $\frac{|A|}{\lambda}$ is an eigen value of the matrix adj (A) provided A is non-singular.

Proof : (*i*) Consider $(A - Ix)' = A' - (I\,x)' = A' - Ix$.

Now $|A - Ix| = |(A - Ix)'| = |A' - I\,x|$

implies that the characteristic polynomials of A and A′ are identical.

Hence these matrices have the same eigen values.

(*ii*) Let us prove it for a matrix of order 3.

Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

Then $|A - I\,x| = \begin{vmatrix} a_{11}-x & a_{12} & a_{13} \\ a_{21} & a_{22}-x & a_{23} \\ a_{31} & a_{32} & a_{33}-x \end{vmatrix}$

$= -x^3 + (a_{11} + a_{22} + a_{33})\,x^2 -$ ...(1)

If $x = \lambda_1, \lambda_2, \lambda_3$ are eigen values of A, then $|A - Ix| = (\lambda_1 - x)(\lambda_2 - x)(\lambda_3 - x)$

$$= -(x - \lambda_1)(x - \lambda_2)(x - \lambda_3)$$

$$= -x^3 + (\lambda_1 + \lambda_2 + \lambda_3)x^2 - \ldots.. \qquad ...(2)$$

Comparing coefficients of x^2 on R.H.S. of (1) and (2), we obtain

$$\lambda_1 + \lambda_2 + \lambda_3 = a_{11} + a_{22} + a_{33}$$

In a same way, we can prove this property for any square matrix.

Further for proving (*iii*) — (*iv*), let $x = \lambda$ be an eigen value of A and let X be the corresponding eigen vector of A. Then we have

$$(A - \lambda I) X = 0 \text{ or } AX = \lambda X \qquad ...(3)$$

(*iii*) Now $(A + kI) X = AX + k (I X)$

$$= \lambda X + k X, \text{ by (3)}$$

$$= (\lambda + k) X \qquad ...(4)$$

Equation (4) shows that $\lambda + k$ is an eigen value of $A + k I$.

(*iv*) Now $A^2 X = (AA)X$

$$= A(AX) = A(\lambda X), \text{ by (3)}$$

$$= \lambda(AX) = \lambda(\lambda X), \text{ again by (3)}$$

$$= \lambda^2 X \qquad ...(5)$$

Equation (5) show that λ^2 is an eigen value of A^2.

Similarly, we can prove that λ^m is an eigen value of A^m provided λ is an eigen value of A, where m is positive integer.

(*v*) Since $|A| \neq 0$, A^{-1} exists. Multiplying on both sides of (3) by A^{-1}, we obtain $A^{-1}(AX) = A^{-1}(\lambda X)$

which implies $IX = \lambda(A^{-1}X)$ or $A^{-1} X = \frac{1}{\lambda} X \qquad ...(6)$

Equations (6) proves that $\frac{1}{\lambda}$ is an eigen value of A^{-1}.

(*vi*) Multiplying (3) on both sides by adj (A) and using the fact that adj (A). A= | A |. I,

we obtain (| A | . I) X = λ.adj (A).X or λ [adj (A).X] = | A | X

which implies adj (A). X = $\dfrac{|A|}{\lambda}$ X ...(7)

Equation (7) shows that $\dfrac{|A|}{\lambda}$ is an eigen value of adj. (A).

18.4.3. Diagonalization : If a square matrix A of order *m* has *m* linearly independent eigen vectors, then a matrix P of order *m* can be determined such that P^{-1} AP is a diagonal matrix.

The matrix P which diagonalizes A is called the **modal matrix** of A and the resulting diagonal matrix is called the **spectral matrix** of A.

Now let us understand this result for a matrix of order 3. Suppose A is a matrix of order 3 with eigen values $\lambda_1, \lambda_2, \lambda_3$ and let the corresponding independent eigen vectors be

$$X_1 = \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix} \qquad X_2 = \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix} \quad \text{and} \quad X_3 = \begin{bmatrix} a_3 \\ b_3 \\ c_3 \end{bmatrix}$$

Since eigen vector are non-zero solutions of the matrix equation : AX = λX,

we have $AX_1 = \lambda_1 X_2, \quad AX_2 = \lambda_2 X_2, \quad AX_3 = \lambda_3 X_3$

Let $P = [X_1, X_2, X_3] = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ and $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$

Then $AP = [AX_1, AX_2, AX_3] = [\lambda_1 X_1, \lambda_2 X_2, \lambda_3 X_3]$

$$= \begin{bmatrix} \lambda_1 a_1 & \lambda_2 a_2 & \lambda_3 a_3 \\ \lambda_1 b_1 & \lambda_2 b_2 & \lambda_3 b_3 \\ \lambda_1 c_1 & \lambda_2 c_2 & \lambda_3 c_3 \end{bmatrix}$$

$$= \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = PD.$$

This implies that $P^{-1}AP = D$.

Examples 6. Diagonalize the following matrices :

(*i*) $A = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 1 & 1 \\ -7 & 2 & -3 \end{bmatrix}$

(*ii*) $B = \begin{bmatrix} -1 & 2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$

(*iii*) $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

Soluions. (*i*) The characteristic equation of given matrix A is

$$0 = |A - \lambda I| = \begin{vmatrix} 2-\lambda & 2 & 0 \\ 2 & 1-\lambda & 1 \\ -7 & 2 & -3-\lambda \end{vmatrix}$$

Expanding by R_1, we obtain

$$-\lambda^3 + 13\lambda - 12 = 0$$

$\Rightarrow \quad \lambda^3 - 13\lambda + 12 = 0$

$\Rightarrow \quad (\lambda-1)(\lambda^2 + \lambda - 12) = 0$

$\Rightarrow \quad (\lambda-1)(\lambda+4)(\lambda-3) = 0$

$\Rightarrow \quad \lambda = 1, 3, -4.$

Thus the eigen values of A are

$\lambda_1 = 1, \lambda_2 = 3, \lambda_3 = -4.$

For $\lambda_1 = 1$, the corresponding eigen vector $X_1 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$ is non-zero solution of the matrix equation : $(A - \lambda_1 I) X_1 = O$

$$\text{or} \quad \begin{bmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \\ -7 & 2 & -4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Applying $R_3 \to R_3 - R_1 + 4R_2$, we have

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad ...(1)$$

This is of the form : $CX_1 = 0$, where order of C is 3 and $r(C) = 2$.

So the number of independent solutions of system (1) is $3 - 2 = 1$.

From (1), we have $a + 2b = 0$ and $2a + c = 0$

$\Rightarrow$ $2b = -a$ and $c = -2a$.

Choose $a = -2$. Then $b = 1$ and $c = 4$.

Thus the eigen vector corresponding to λ_1 is

$$X_1 = \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}$$

For $\lambda_2 = 3$, it can be found that the corresponding eigen vector is $X_2 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$ where as

$X_3 = \begin{bmatrix} 1 \\ -3 \\ 13 \end{bmatrix}$ is the eigen vector corresponding to $\lambda_3 = -4$.

$$\text{Let } D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -4 \end{bmatrix} \text{ and}$$

let $P = [X_1\ X_2\ \ X_3] = \begin{bmatrix} -2 & 2 & 1 \\ 1 & 1 & -3 \\ 4 & -2 & 13 \end{bmatrix}$

$$\text{Then } AP = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 1 & 1 \\ -7 & 2 & -3 \end{bmatrix} \begin{bmatrix} -2 & 2 & 1 \\ 1 & 1 & -3 \\ 4 & -2 & 13 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 6 & -4 \\ 1 & 3 & 12 \\ 4 & -6 & -52 \end{bmatrix}$$

$$\text{and } PD = \begin{bmatrix} -2 & 2 & 1 \\ 1 & 1 & -3 \\ 4 & -2 & 13 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 6 & -4 \\ 1 & 3 & 12 \\ 4 & -6 & -52 \end{bmatrix}$$

Thus we see that $AP = PD$, which implies that $P^{-1}AP = D$. Hence the modal matrix $P = \begin{bmatrix} -2 & 2 & 1 \\ 1 & 1 & -3 \\ 4 & -2 & 13 \end{bmatrix}$ diagonalizes A in the sense that $P^{-1}AP = D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -4 \end{bmatrix}$.

(*ii*) As already derived in Example 5 (*iii*) above, the eigen values of B are $\lambda_1 = 1, \lambda_2 = \sqrt{5}, \lambda_3 = -\sqrt{5}$ and corresponding eigen vectors are respectively

$$X_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 1-\sqrt{5} \\ -1 \\ 1 \end{bmatrix} \text{ and } X_3 = \begin{bmatrix} 1+\sqrt{5} \\ -1 \\ 1 \end{bmatrix}$$

Let D = $\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{5} & 0 \\ 0 & 0 & -\sqrt{5} \end{bmatrix}$

and $P = [X_1, X_2, X_3] = \begin{bmatrix} 1 & 1-\sqrt{5} & 1+\sqrt{5} \\ 0 & -1 & -1 \\ -1 & 1 & 1 \end{bmatrix}$

Then BP = $\begin{bmatrix} -1 & 2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1-\sqrt{5} & 1+\sqrt{5} \\ 0 & -1 & -1 \\ -1 & 1 & 1 \end{bmatrix}$

$= \begin{bmatrix} 1 & -5+\sqrt{5} & -5-\sqrt{5} \\ 0 & -\sqrt{5} & \sqrt{5} \\ -1 & \sqrt{5} & -\sqrt{5} \end{bmatrix}$

and PD = $\begin{bmatrix} 1 & 1-\sqrt{5} & 1+\sqrt{5} \\ 0 & -1 & -1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{5} & 0 \\ 0 & 0 & -\sqrt{5} \end{bmatrix}$

$= \begin{bmatrix} 1 & \sqrt{5}-5 & -\sqrt{5}-5 \\ 0 & -\sqrt{5} & \sqrt{5} \\ -1 & \sqrt{5} & -\sqrt{5} \end{bmatrix}$

Thus we see that BP = PD, which implies that P^{-1} BP = D.

Hence the model matrix P = $\begin{bmatrix} 1 & 1-\sqrt{5} & 1+\sqrt{5} \\ 0 & -1 & -1 \\ -1 & 1 & 1 \end{bmatrix}$ diagonalizes B in the sense that $P^{-1} BP = D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{5} & 0 \\ 0 & 0 & -\sqrt{5} \end{bmatrix}$.

(*iii*) The characteristic equation of A is

$$0 = |A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 & 1 \\ 0 & 1-\lambda & 2 \\ 0 & 0 & 1-\lambda \end{vmatrix}$$

$$= (1 - \lambda)^3$$

⇒ $(1 - \lambda)^3 = 0$

⇒ $\lambda = 1, 1, 1$ as eigen values (repeated thrice) of A.

For $\lambda = 1$, the corresponding eigen vector $X = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$ is non-zero solution of $(A - \lambda I)X = O$ or $(A - I) X = O$.

This gives $\begin{bmatrix} 0 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

which is of the form $PX = 0$. ...(1)

Here order of P is 3 and $r(P) = 2$. So the number of independent solutions of the system (1) is $3 - 2 = 1$. From (1), we have

$2b + c = 0$ and $2c = 0$

⇒ $c = 0$, $b = 0$. Choose $a = 1$.

Thus only one eigen vector corresponding to $\lambda = 1$ is $X = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

So in this example we do not have three linearly independent eigen vectors. Hence the matrix P cannot be found and so the given matrix is not diagonalizable.

18.5. SUMMARY

In this lesson, we studied eigen values of a matrix and its important properties alongwith solved problems. Also we studied eigen vectors and diagonalization of a matrix with solved examples.

18.6 EXERCISE

1. Show that the eigen values of a triangular matrix or a diagonal matrix are just the diagonal elements of the matrix.

2. Show that the product of the eigen values of the matrix A = $\begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$ is equal to |A|.

3. Verify property (1) given in 18.4.1 for the matrix :

$$\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$$

4. Verify properties (6) and (7) given in 18.4.1 for the matrix : $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

5. Find eigen values of A^3, where

(*i*) $\begin{bmatrix} 3 & 2 \\ 8 & 9 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$

6. Verify whether the following statements are true or false :

(*i*) $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigen vector of $\begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$

(*ii*) $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigen vector of $\begin{bmatrix} 2 & 1 \\ -2 & 5 \end{bmatrix}$.

7. Find all the eigen values and eigen vectors of the following matrices :

(*i*) $\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$

(*ii*) $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$

(*iii*) $\begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{bmatrix}$

(*iv*) $\begin{bmatrix} 4 & 6 & 6 \\ 1 & 3 & 2 \\ -1 & -5 & -2 \end{bmatrix}$

8. Diagonalize the following matrices :

(*i*) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -1 & 3 \end{bmatrix}$

(*ii*) $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$

(*iii*) $\begin{bmatrix} 0 & 0 & 1 \\ 8 & 0 & 2 \\ 0 & -1 & 5 \end{bmatrix}$

(*iv*) $\begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{bmatrix}$

9. Show that the vectors $X_1 = (5, 4, 3)$, $X_2 = (8, 1, 3)$ and $X_3 = (3, 3, 2)$ are linearly dependent.

10. Show that the vectors $X_1 = (1, 0, -1)$, $X_2 = (0, 1, 0)$ and $X_3 = (1, 0, 1)$ are linearly independent.

5. (*i*) 1, 1331 (*ii*) 1, 1, 27

6. (*i*) False (*ii*) True

7. (*i*) 0, 3, 15 ; $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$

(*iii*) 1, 2, 2 ; $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

(*ii*) 1, 2, 3 ; $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

(*iv*) 1, 2, 2 ; $\begin{bmatrix} 4 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$

8. (*i*) $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$, $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$

(*ii*) $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$, $P = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 0 & -1 \\ 0 & 2 & 1 \end{bmatrix}$

(*iii*) $D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$, $P = \begin{bmatrix} -1 & 1 & 1 \\ 6 & 6 & 4 \\ 1 & 2 & 4 \end{bmatrix}$

(*iv*) Not diagonalizable

18.7. SUGGESTED READING/REFERENCES :

1. "A text book of Matrices" by Shanti Narayan and P.K.Mittal, S. Chand & Company Ltd.

2. "Theory of Matrices" by B.S. Vatssa, Wiley Easterns Ltd.

3. "Engineering Mathematics-II" by Bhopinder Singh, Malhotra Brothers, Jammu.

18.8. MODEL TEST PAPER

1. Verify whether the following statements are true or false :

(*i*) $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ is an eigen vector of $\begin{bmatrix} -1 & 4 \\ 3 & 3 \end{bmatrix}$

(*ii*) $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ is an eigen vector of $\begin{bmatrix} 3 & 2 \\ 8 & 9 \end{bmatrix}$

2. Find all the eigen values and eigen vectors of the following matrices :

(*i*) $\begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$

(*ii*) $\begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$

(*iii*) $\begin{bmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{bmatrix}$

(*iv*) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix}$.

3. Diagonalize the following matrices :

(*i*) $\begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix}$

(*ii*) $\begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{bmatrix}$

(*iii*) $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

4. Verify that the product of the eigven values of the matrix A = $\begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$ is equal to |A|.

5. Verify property (1) given in 18.4.1 for the matrix :

$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$$

6. Verify properties (6) and (7) given in 18.4.1 for the matrix : $\begin{bmatrix} 3 & 2 \\ 8 & 9 \end{bmatrix}$

7. Find the eigen values of A^3, where

(*i*) $A = \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$ (*ii*) $A = \begin{bmatrix} 1 & -2 & -2 \\ 0 & -1 & 1 \\ 2 & 0 & -2 \end{bmatrix}$

8. Show that vectors $X_1 = (2, 1, 4)$, $X_2 = (7, 5, 17)$ and $X_3 = (1, 1, 3)$ are linearly dependent.

9. Show that vectors $X_1 = (2, -1, 1)$, $X_2 = (1, 2, 0)$ and $X_3 = (-1, 0, 2)$ are linearly independent.