ALAGAPPA UNIVERSITY

[Accredited with 'A+' Grade by NAAC (CGPA:3.64) in the Third Cycle
and Graded as Category–I University by MHRD-UGC]

(A State University Established by the Government of Tamil Nadu)

KARAIKUDI – 630 003

Directorate of Distance Education

B.Sc. [Mathematics]

VI - Semester

113 64

COMBINATORICS

Authors:

N Ch S N Iyengar, Professor, Deptt of Computer Applications, Vellore Institute of Technology, Vellore
V M Chandrasekaran, Asstt Professor, Deptt of Mathematics, Vellore Institute of Technology, Vellore
K A Venkatesh, Head - Deptt of Computer Applications, Alliance Business Academy, Bangalore
P S Arunachalam, Senior Lecturer, Department of Mathematics, SRM Engineering College, Chennai
Units (1.0 - 1.4, 2, 5, 11, 14.0-14.2)

C R Kothari, Former Principal, University Commerce College, Jaipur
Unit (3)

Vikas®Publishing House, Units (1.5 - 1.11, 4, 6, 7, 8, 9, 10, 12, 13, 14.3-14.8)

"The copyright shall be vested with Alagappa University"

All rights reserved. No part of this publication which is material protected by this copyright notice may be reproduced or transmitted or utilized or stored in any form or by any means now known or hereinafter invented, electronic, digital or mechanical, including photocopying, scanning, recording or by any information storage or retrieval system, without prior written permission from the Alagappa University, Karaikudi, Tamil Nadu.

Information contained in this book has been published by VIKAS® Publishing House Pvt. Ltd. and has been obtained by its Authors from sources believed to be reliable and are correct to the best of their knowledge. However, the Alagappa University, Publisher and its Authors shall in no event be liable for any errors, omissions or damages arising out of use of this information and specifically disclaim any implied warranties or merchantability or fitness for any particular use.



Vikas® is the registered trademark of Vikas® Publishing House Pvt. Ltd.

VIKAS® PUBLISHING HOUSE PVT. LTD.
E-28, Sector-8, Noida - 201301 (UP)
Phone: 0120-4078900 • Fax: 0120-4078999
Regd. Office: A-27, 2nd Floor, Mohan Co-operative Industrial Estate, New Delhi 1100 44
• Website: www.vikaspublishing.com • Email: helpline@vikaspublishing.com

Work Order No.AU/DDE/DE12-27/Preparation and Printing of Course Materials/2020 Dated 12.08.2020 Copies - 500

SYLLABI-BOOK MAPPING TABLE

Combinatorics

Syllabi	Mapping in Book
BLOCK I: COMBINATION OF NUMBERS AND GENERATING FUNCTIONS	
Unit 1 Basic Combinatorial Numbers – Stirling Numbers of the First Kind – Stirling Numbers of the Second Kind.	**Unit 1:** Combination and Stirling Number **(Pages 1-25);**
Unit 2 Recurrence Formula for Snm – Recurrence formula for Pnm.	**Unit 2:** Recurrence Formula **(Pages 26-44);**
Unit 3 Patterns of Distributions-Problems.	**Unit 3:** Patterns of Distributions **(Pages 45-78);**
Unit 4 Generating Functions – The Algebra of Formal Power Series – Generating Functions for Permutations – Generating Functions for Partitions.	**Unit 4:** Generating Functions **(Pages 79-99)**
BLOCK II: RECURRENCE RELATIONS AND SYMMETRIC FUNCTIONS	
Unit 5 Inventory of Maps – Recurrence Relations.	**Unit 5:** Recurrence Relations **(Pages 100-106);**
Unit 6 Symmetric Functions – The Monomial Symmetric Functions – The Complete Homogeneous Symmetric Functions.	**Unit 6:** Symmetric Functions **(Pages 107-113);**
Unit 7 The Elementary Symmetric Functions – The Power Sum Symmetric Function.	**Unit 7:** Elementary Symmetric Functions **(Pages 114-118);**
Unit 8 Multinomial - Basic Concepts- Problems.	**Unit 8:** Multinomials **(Pages 119-125)**
BLOCK III: PRINCIPLES, PERMUTATIONS AND POLYA THEORY	
Unit 9 Inclusion and Exclusion Principle –Theorems and Problems.	**Unit 9:** Inclusion and Exclusion Principles **(Pages 126-132);**
Unit 10 Permutations with Forbidden Positions – The Menage Problem.	**Unit 10:** Permutations **(Pages 133-139);**
Unit 11 Problem of Fibonacci – Polya's Theory – Problems.	**Unit 11:** Problem of Fibonacci and Polya's Theory **(Pages 140-164)**
BLOCK IV: PERMUTATION GROUPS	
Unit 12 Necklace Problem and Burnside's Lemma – Cyclic Index of a Permutation Group.	**Unit 12:** Necklace Problem and Burnside's Lemma **(Pages 165-172);**
Unit 13 Polya's Theorems and Their Immediate Applications – Related Problems.	**Unit 13:** Polya's Theorems and their Immediate Applications **(Pages 173-180);**
Unit 14 Binary Operations on Permutation Groups.	**Unit 14:** Binary Operations and Permutation Groups **(Pages 181-192)**

CONTENTS

BLOCK I: COMBINATION OF NUMBERS AND GENERATING FUNCTIONS

UNIT 1 COMBINATION AND STIRLING NUMBER 1-25

1.0 Introduction
1.1 Objectives
1.2 Basic Combinatorial Numbers
1.3 Unordered Samples With/Without Repetition
1.4 Permutations with Repetitions
1.5 Stirling Number
1.6 Stirling Numbers of the First Kind and the Second Kind
1.7 Answers to Check Your Progress Questions
1.8 Summary
1.9 Key Words
1.10 Self Assessment Questions and Exercises
1.11 Further Readings

UNIT 2 RECURRENCE FORMULA 26-44

2.0 Introduction
2.1 Objectives
2.2 Recursion
2.3 Recursion and Iteration
2.4 Closed Form Expression
2.5 Recurrence Relations Obtained from Solutions
2.6 Recurrence Formula for S_n^m
2.7 Recurrence Formula for P_n^m
2.8 Answers to Check Your Progress Questions
2.9 Summary
2.10 Key Words
2.11 Self Assessment Questions and Exercises
2.12 Further Readings

UNIT 3 PATTERNS OF DISTRIBUTIONS 45-78

3.0 Introduction
3.1 Objectives
3.2 Theoretical Distributions: Discrete Distribution
  3.2.1 Binomial Distribution
  3.2.2 Poisson Distribution
  3.2.3 Limiting Form of Binomial and Poisson Fitting
3.3 Theoretical Distributions: Continuous Distribution
  3.3.1 Normal Distribution
3.4 Moment Generating Functions and Fitting of Data
3.5 Calculation of Mean and Variance from Moment Generating Functions
3.6 Answers to Check Your Progress Questions
3.7 Summary
3.8 Key Words
3.9 Self Assessment Questions and Exercises
3.10 Further Readings

UNIT 4 GENERATING FUNCTIONS **79-99**

4.0 Introduction
4.1 Objectives
4.2 Generating Functions
4.3 The Algebra of Formal Power Series
4.4 Generating Functions for Permutations
4.5 Generating Functions for Partitions
4.6 Answers to Check Your Progress Questions
4.7 Summary
4.8 Key Words
4.9 Self Assessment Questions and Exercises
4.10 Further Readings

BLOCK II: RECURRENCE RELATIONS AND SYMMETRIC FUNCTIONS

UNIT 5 RECURRENCE RELATIONS **100-106**

5.0 Introduction
5.1 Objectives
5.2 Inventory of Maps
5.3 Recurrence Relations
5.4 Answers to Check Your Progress Questions
5.5 Summary
5.6 Key Words
5.7 Self Assessment Questions and Exercises
5.8 Further Readings

UNIT 6 SYMMETRIC FUNCTIONS **107-113**

6.0 Introduction
6.1 Objectives
6.2 Symmetric Functions
6.3 The Monomial Symmetric Functions m_α
6.4 The Complete Homogeneous Symmetric Functions h_k
6.5 Answers to Check Your Progress Questions
6.6 Summary
6.7 Key Words
6.8 Self Assessment Questions and Exercises
6.9 Further Readings

UNIT 7 ELEMENTARY SYMMETRIC FUNCTIONS **114-118**

7.0 Introduction
7.1 Objectives
7.2 The Elementary Symmetric Functions
7.3 The Power Sum Symmetric Functions
7.4 Answers to Check Your Progress Questions
7.5 Summary
7.6 Key Words
7.7 Self Assessment Questions and Exercises
7.8 Further Readings

UNIT 8 MULTINOMIALS **119-125**

8.0 Introduction
8.1 Objectives
8.2 Multinomials
8.3 Basic Concepts of Multinomials
8.4 Answers to Check Your Progress Questions
8.5 Summary
8.6 Key Words
8.7 Self Assessment Questions and Exercises
8.8 Further Readings

BLOCK III: PRINCIPLES, PERMUTATIONS AND POLYA THEORY

UNIT 9 INCLUSION AND EXCLUSION PRINCIPLES **126-132**

9.0 Introduction
9.1 Objectives
9.2 Inclusion and Exclusion Principle and Theorems
9.3 Answers to Check Your Progress Questions
9.4 Summary
9.5 Key Words
9.6 Self Assessment Questions and Exercises
9.7 Further Readings

UNIT 10 PERMUTATIONS **133-139**

10.0 Introduction
10.1 Objectives
10.2 Permutations with Forbidden Positions
10.3 The Ménage Problem
10.4 Answers to Check Your Progress Questions
10.5 Summary
10.6 Key Words
10.7 Self Assessment Questions and Exercises
10.8 Further Readings

UNIT 11 PROBLEM OF FIBONACCI AND POLYA'S THEORY **140-164**

11.0 Introduction
11.1 Objectives
11.2 Problems of Fibonacci
11.3 Polya Theory
11.4 Answers to Check Your Progress Questions
11.5 Summary
11.6 Key Words
11.7 Self Assessment Questions and Exercises
11.8 Further Readings

BLOCK IV: PERMUTATION GROUPS

UNIT 12 NECKLACE PROBLEM AND BURNSIDE'S LEMMA 165-172

12.0 Introduction
12.1 Objectives
12.2 Necklace Problem
12.3 Burnside's Lemma
12.4 Cyclic Index of a Permutation Group
12.5 Answers to Check Your Progress Questions
12.6 Summary
12.7 Key Words
12.8 Self Assessment Questions and Exercises
12.9 Further Readings

UNIT 13 POLYA'S THEOREMS AND THEIR IMMEDIATE APPLICATIONS 173-180

13.0 Introduction
13.1 Objectives
13.2 Polya's Theorems
  13.2.1 Applications
13.3 Answers to Check Your Progress Questions
13.4 Summary
13.5 Key Words
13.6 Self Assessment Questions and Exercises
13.7 Further Readings

UNIT 14 BINARY OPERATIONS AND PERMUTATION GROUPS 181-192

14.0 Introduction
14.1 Objectives
14.2 Binary Operations
  14.2.1 Binary Relation
  14.2.2 Properties of Binary Relations
  14.2.3 Closures of Relation
14.3 Permutation Groups
  14.3.1 Permutation Group
  14.3.2 Even and Odd Permutation
  14.3.3 Alternating Group
14.4 Answers to Check Your Progress Questions
14.5 Summary
14.6 Key Words
14.7 Self Assessment Questions and Exercises
14.8 Further Readings

INTRODUCTION

NOTES

Combinatorics is an area of mathematics primarily concerned with counting, both as a means and an end in obtaining results, and certain properties of finite structures. It is closely related to many other areas of mathematics and has many applications ranging from logic to statistical physics, from evolutionary biology to computer science, etc. According to H.J. Ryser, a definition of the subject combinatorics is difficult because it crosses so many mathematical subdivisions.

Typically, the area of combinatorics is involved with - the enumeration (counting) of specified structures, sometimes referred to as arrangements or configurations in a very general sense, associated with finite systems, the existence of such structures that satisfy certain given criteria, the construction of these structures, perhaps in many ways, and optimization, finding the 'Best' structure or solution among several possibilities, be it the 'Largest', 'Smallest' or satisfying some other optimality criterion.

According to Leon Mirsky, "Combinatorics is a range of linked studies which have something in common and yet diverge widely in their objectives, their methods, and the degree of coherence they have attained". One way to define combinatorics is, perhaps, to describe its subdivisions with their problems and techniques.

Enumerative combinatorics is the most classical area of combinatorics and concentrates on counting the number of certain combinatorial objects. Fibonacci numbers is the basic example of a problem in enumerative combinatorics. Partition theory studies various enumeration and asymptotic problems related to integer partitions, and is closely related to q-series, special functions and orthogonal polynomials. A recurrence relation defines each term of a sequence in terms of the preceding terms. In other words, once one or more initial terms are given, each of the following terms of the sequence is a function of the preceding terms.

This book, *Combinatorics*, is divided into four blocks, which are further subdivided into fourteen units. This book provides a basic understanding of the subject and helps to grasp its fundamentals. In a nutshell, it explains various aspects, such as basic combinatorial numbers, Stirling numbers of the first kind, Stirling numbers of the second kind, recurrence formula, patterns of distributions, generating functions for permutations and partitions, recurrence relations, symmetric functions, multinomial, inclusion and exclusion principle, permutations with forbidden positions, problem of Fibonacci, Polya's theory, necklace problem and Burnside's lemma, cyclic index of a permutation group, Polya's theorems and their immediate applications, and binary operations on permutation groups.

The book follows the Self-Instructional Mode (SIM) wherein each unit begins with an 'Introduction' to the topic. The 'Objectives' are then outlined before going on to the presentation of the detailed content in a simple and structured format. 'Check Your Progress' questions are provided at regular intervals to test the student's understanding of the subject. 'Answers to Check Your Progress Questions', a 'Summary', a list of 'Key Words', and a set of 'Self-Assessment Questions and Exercises' are provided at the end of each unit for effective recapitulation.

NOTES

BLOCK I
COMBINATION OF NUMBERS AND GENERATING FUNCTIONS

UNIT 1 COMBINATION AND STIRLING NUMBER

Structure

1.0 Introduction
1.1 Objectives
1.2 Basic Combinatorial Numbers
1.3 Unordered Samples With/Without Repetition
1.4 Permutations with Repetitions
1.5 Stirling Number
1.6 Stirling Numbers of the First Kind and the Second Kind
1.7 Answers to Check Your Progress Questions
1.8 Summary
1.9 Key Words
1.10 Self Assessment Questions and Exercises
1.11 Further Readings

1.0 INTRODUCTION

In mathematics, a combination is a selection of items from a collection, such that the order of selection does not matter (unlike permutations). For example, given three fruits, say an apple, an orange and a pear, there are three combinations of two that can be drawn from this set: an apple and a pear; an apple and an orange; or a pear and an orange. More formally, a *k*-combination of a set *S* is a subset of k distinct elements of *S*. If the set has *n* elements, the number of *k*-combinations is equal to the binomial coefficient $\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1}$, $\binom{S}{k}\frac{n!}{k!(n-k)!}$ Which can be written using factorials as whenever $k \leq 1$, and which is zero when $k > n$. The set of all *k*-combinations of a set *S* is often denoted by $\binom{S}{k}$.

Combinations refer to the combination of *n* things taken *k* at a time without repetition. To refer to combinations in which repetition is allowed, the terms *k*-selection, *k*-multiset, or *k*-combination with repetition are often used. If, in the above example, it was possible to have two of any one kind of fruit there would be 3 more 2-selections: one with two apples, one with two oranges, and one with two pears.

Although the set of three fruits was small enough to write a complete list of combinations, this becomes impractical as the size of the set increases. For example,

NOTES

a poker hand can be described as a 5-combination ($k = 5$) of cards from a 52-card deck ($n = 52$). The 5 cards of the hand are all distinct, and the order of cards in the hand does not matter. There are 2,598,960 such combinations, and the chance of drawing any one hand at random is 1 / 2,598,960.

In this unit, you will study the basic combinatorial numbers, unordered samples with/without repetition, permutations with repetitions, Stirling number and Stirling number of first kind.

1.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the basic concept of combinatorial numbers
- Analyse the concept of unordered samples with/without repetition
- Discuss about the study of permutation with repetition
- Explain the definition of Stirling numbers
- Understand the concept and function of Stirling number of first kind

1.2 BASIC COMBINATORIAL NUMBERS

The number of combinations of n different things taken r at a time is

$${}^{n}c_{r} = \frac{n(n-1)(n-2)...(n-r+1)}{1, 2, 3...r}.$$

Example 1.1: How many diagonals are there in a polygon of n sides?

Solution: The number of diagonals is the number of straight lines joining any two consecutive vertices excepted; and hence the required number $= {}^{n}C_{2} - n$

$$= \frac{n(n-1)}{2} - n = \frac{n^2 - 3n}{2}.$$

Example 1.2: The English language consists of 21 consonants and 5 vowels. How many 5 lettered words, consisting of atleast a vowel and two consonants, can be formed from them?

Solution: Each 5-lettered word might consists of,

(*i*) 1 Vowel and 4 consonants

(*ii*) 2 Vowels and 3 consonants

(*iii*) 3 Vowels and 2 consonants

Now, it

(*i*) Gives rises to $5C_1 \times 21C_4 \times 5!$

(*ii*) Gives rises to $5C_2 \times 21C_3 \times 5!$

(*iii*) Gives rises to $5C_3 \times 21C_2 \times 5!$

∴ The required number of words is 5,439,000.

Example 1.3: Find the number of permutations when six letters at a time are taken from the word RAMAYANAM.

Solution: There are 9 letters of 5 sorts, namely, *aaaa*; *mm*; *r*; *y*; *n*

The combination and permutation can be grouped. Then:

(*i*) 4 alike, 2 alike,

The number of combinations = 1.1 = 1

The number of permutations = $\frac{6!}{4!2!} = 15$

(*ii*) 4 alike, 2 different,

The number of combinations = 1.4 C_2= 6

The number of permutations = $6\frac{6!}{4!} = 180$

(*iii*) 3 alike, 2 alike, 1 different,

The number of combinations = 1.1.3 C_1= 3

The number of permutations = $3.\frac{6!}{3!2!} = 180$

(*iv*) 3 alike, 3 different,

The number of combinations = 1.4C_3= 4

The number of permutations = $4.\frac{6!}{3!} = 480$

(*v*) 2 alike, 2 alike, 2 different,

The number of combinations = 1.1.3C_2 =3

The number of permutations = $3\frac{6!}{2!2!} = 540.$

(*vi*) 2 alike, 4 different,

The number of combinations = 2C_1 4C_4= 2

The number of permutations = $2.\frac{6!}{2!} = 720$

The total number of permutations = 15 + 180 + 180 +480 + 540 + 720 = 2,115.

Example 1.4: I have 4 friends: in how many ways can I invite them for dinner?

Solution: The required number is, $4C_1 + 4C_2 + 4C_3 + 4C_4 = 4 + 6 + 4 + 1 = 15.$

NOTES

1.3 UNORDERED SAMPLES WITH/WITHOUT REPETITION

NOTES

In any arrangement or sampling, if order does not matter then it is a combination. Here also we have two different cases, one in which repetition is allowed and another in which repetition is not allowed. We can think of a case of repetition being allowed is to think of coins in your pocket. Your pocket has coins of 1, 1, 1, 2, 5, 5… like that. The case when repetition is not allowed can be thought of as a sequence of number 1,2,3 which is considered one sample and here ordering is not important. If orders are taken then there are six number of arrangements like, 123, 132, 231, 213, 312, and 321. So unordered sample is like a set. It is a set of three elements 1,2,3. Here we will discus those where repetition is not allowed and so we are dealing with combination without repetition.

An ordered arrangement of in n things taken r at a time (without repetition) is given by $n!/(n-r)!$. If the same n things are arranged taken $r!$ at a time, it is given as $n!/\{(n-r)!r!\}$. Thus when restriction of ordering is removed then the number of arrangement is reduced by $r!$ times. Symbolically we represent first case of orderly arrangement as permutation, ${}^nP_r$ and in second case it is a combination ${}^nC_r$ and ${}^nP_r = r!\,{}^nC_r$.

1.4 PERMUTATIONS WITH REPETITIONS

Permutations with Repetitions: The number of permutations of n different things taken r at a time when the things can be repeated any number of times is n^r.

Note: The total number of permutations of n things taken 1, 2, 3,....., r at a time when the things may be repeated any number of times is $n + n^2 + ... + n^r = n.\ \dfrac{n^r - 1}{n-1}$.

Example 1.5: How many numbers of four digits can be formed with the digits 1, 2, 3 ? Find the sum of all such numbers.

Solution: The number of four digit numbers = 3 × 3 × 3 × 3 = 81.

If the numbers are all written down we find that any one of the digits 1, 2, 3 occurs in the units place in $\dfrac{81}{3} = 27$ times. Hence, the total sum of the digits in the units place is, 27 (1 +2 + 3) = 162. This is also the sum of all digits in the tens, hundreds and thousands places.

Hence, the required sum is, 162 + 162 × 10 + 162 × 100 + 162 × 1000 = 179982.

NOTES

Permutations when all the Things are not Different: The number of permutations of n things taken together when all the things are not different is given by $\frac{n!}{p!q!r!...}$

Here, among the n things, p things of them are of one kind; q of them are of second kind, and so on.

$$(p + q + r + ... = n)$$

Example 1.6: How many different words can be made out of the letters which form the word ALLAHABAD?

Solution: There are 9 letters of which 4 letters are of one sort (*A, A, A, A*); 2 are of second sort *(L, L)*; 1 is of third sort (*H*); 1 is of fourth sort (*B*); and 1 is of different sort (*D*).

$$\text{The required number} = \frac{9!}{4!2!}$$

Example 1.7: In how many ways can the letters of the word ENGINEERING be arranged (*i*) without changing the order of the consonants (*ii*) without changing the relative positions of the vowels and the consonants?

Solution:

(*i*) The consonants are required to be in the same order as in the given word and so, there can be no interchange of posititons among them and so, they may be replaced by letters say *c,c,c,c,c,c*. All these 6 *c*'s can be arranged in one and only way. Now, we have to find the number of permutations of 11 letters of which the 6 consonants are alike; 3 vowels *e,e,e* are alike; and the two vowels *i*, *i* are alike.

Hence, the required number $= \frac{11!}{6!3!2!} = 4{,}620.$

(*ii*) The places originally occupied by vowels must always be occupied by vowels and those occupied by consonants, always by consonants. The vowels *e, e, e, i, i* can be arranged among themselves in $\frac{5!}{3!2!} = 10$ ways and the consonants *n, n, n, g, g, r* can be arranged among themselves in $\frac{6!}{3!2!} = 60$ ways.

Hence, the required number is $10 \times 60 = 600$ ways.

1.5 STIRLING NUMBER

Stirling numbers arise in a variety of analytic and combinatorial problems. They are named after James Stirling, who introduced them in the 18th century. Two different sets of numbers bear this name: the Stirling numbers of the first kind and

NOTES

the Stirling numbers of the second kind. Additionally, Lah numbers are sometimes referred to as Stirling numbers of the third kind. Each kind is detailed in its respective article, this one serving as a description of relations between them.

A common property of all three kinds is that they describe coefficients relating three different sequences of polynomials that frequently arise in combinatorics. Moreover, all three can be defined as the number of partitions of n elements into k non-empty subsets, with different ways of counting orderings within each subset.

Notation

Several different notations for Stirling numbers are in use. Common notation for ordinary (signed) **Stirling numbers of the first kind** is,

$$s(n, k)$$

For **unsigned Stirling numbers of the first kind**, which count the number of permutations of n elements with k disjoint cycles, is:

$$\begin{bmatrix} n \\ k \end{bmatrix} = c(n,k) = |s(n,k)| = (-1)^{n-k} s(n,k)$$

And for **Stirling numbers of the second kind**, which count the number of ways to partition a set of n elements into k nonempty subsets:

$$\left\{ \begin{matrix} n \\ k \end{matrix} \right\} = S(n,k) = S_n^{(k)}$$

For example, the sum,

$$\sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} = n!$$

This is the number of all permutations, while the sum,

$$\sum_{k=0}^{n} \left\{ \begin{matrix} n \\ k \end{matrix} \right\} = B_n$$

This is the nth **Bell number**.

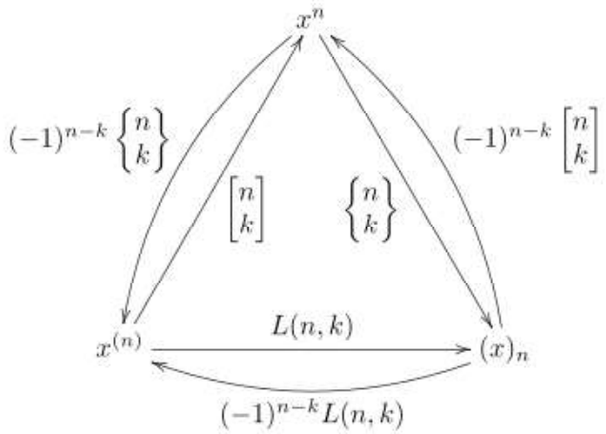
Abramowitz and Stegun use an **uppercase S** and a **blackletter S**, respectively, for the first and second kinds of Stirling number. The notation of brackets and braces, in analogy to binomial coefficients, was introduced in 1935 by Jovan Karamata and promoted later by Donald Knuth. The bracket notation conflicts with a common notation for Gaussian coefficients. The mathematical motivation for this type of notation, as well as additional Stirling number formulae, may be found on the page for Stirling numbers and exponential generating functions.

As Change of Basis Coefficients

Considering the set of polynomials in the (indeterminate) variable x as a vector space, each of the three sequences

$$x^0, x^1, x^2, x^3, \cdots \quad (x)_0, (x)_1, (x)_2, \cdots \quad x^{(0)}, x^{(1)}, x^{(2)}, \cdots$$

is a basis. That is, every polynomial in x can be written as a sum $a_0 x^{(0)} + a_1 x^{(1)} + \cdots + a_n x^{(n)}$ for some unique coefficients a_i (similarly for the other two bases). The above relations then express the change of basis between them, as summarized in the following commutative diagram:



The coefficients for the two bottom changes are described by the Lah numbers below. Since coefficients in any basis are unique, one can define Stirling numbers this way, as the coefficients expressing polynomials of one basis in terms of another, that is, the unique numbers relating x^n with falling and rising factorials as above.

Falling factorials define, up to scaling, the same polynomials as binomial coefficients:

$$\binom{x}{k} = \frac{(x)_k}{k!}$$

The changes between the standard basis $x^0, x^1, x^2, \ldots$ and the basis $\binom{x}{0}, \binom{x}{1}, \binom{x}{2}, \ldots$ are thus described by similar formulas:

$$x^n = \sum_{k=0}^{n} \left\{ n \atop k \right\} k! \binom{x}{k} \text{ and } \binom{x}{n} = \sum_{k=0}^{n} \frac{s(n,k)}{n!} x^k.$$

As Inverse Matrices

The Stirling numbers of the first and second kinds can be considered inverses of one another

$$\sum_{j \geq 0} s(n,j) S(j,k) = \sum_{j \geq 0} (-1)^{n-j} \left[n \atop j \right] \left\{ j \atop k \right\} = \delta_{nk}$$

And

$$\sum_{j \geq 0} S(n,j) s(j,k) = \sum_{j \geq 0} (-1)^{j-k} \left\{ n \atop j \right\} \left[j \atop k \right] = \delta_{nk},$$

NOTES

Where λ_{nk} is the Kronecker delta. These two relationships may be understood to be matrix inverse relationships. That is, let s be the lower triangular matrix of Stirling numbers of the first kind, whose matrix elements $s_{nk} = s(n,k)$. The inverse of this matrix is S, the lower triangular matrix of Stirling numbers of the second kind, whose entries are $S_{nk} = S(n,k)$. Symbolically, this is written

$s^{-1} = S$

Although s and S are infinite, so calculating a product entry involves an infinite sum, the matrix multiplications work because these matrices are lower triangular, so only a finite number of terms in the sum are nonzero.

Example 1

Expressing a polynomial in the basis of falling factorials is useful for calculating sums of the polynomial evaluated at consecutive integers. Indeed, the sum of a falling factorial is simply expressed as another falling factorial (for $k \neq -1$)

$$\sum_{0 \le i < n} (i)_k = \frac{(n)_{k+1}}{k+1}$$

This is analogous to the integral $\int_0^n x^k = \frac{n^{k+1}}{k+1}$, though the sum should be over integers i strictly less than n.

For example, the sum of fourth powers of integers up to n (this time with n included), is:

$$\sum_{i=0}^{n} i^4 = \sum_{i=0}^{n} \sum_{k=0}^{4} \left\{ {4 \atop k} \right\} (i)_k = \sum_{k=0}^{4} \left\{ {4 \atop k} \right\} \frac{(n+1)_{k+1}}{k+1} =$$
$$= \frac{\left\{ {4 \atop 1} \right\}}{2}(n+1)_2 + \frac{\left\{ {4 \atop 2} \right\}}{3}(n+1)_3 + \frac{\left\{ {4 \atop 3} \right\}}{4}(n+1)_4 + \frac{\left\{ {4 \atop 4} \right\}}{5}(n+1)_5 =$$
$$= \frac{1}{2}(n+1)_2 + \frac{7}{3}(n+1)_3 + \frac{6}{4}(n+1)_4 + \frac{1}{5}(n+1)_5$$

Here the Stirling numbers can be computed from their definition as the number of partitions of 4 elements into k non-empty unlabelled subsets.

In contrast, the sum in the standard basis is given by Faulhaber's formula, which in general is $\sum_{i=0}^{n} i$ more complex.

Lah Numbers

The Lah numbers $L(n,k) = \binom{n-1}{k-1} \frac{n!}{k!}$ are sometimes called Stirling numbers of the third kind. By convention, $L(0,0) = 1$ and $L(n,k) = 0$ if $n > k$ or $k = 0 < n$.

These numbers are coefficients expressing falling factorials in terms of rising factorials and vice versa:

$$x^{(n)} = \sum_{k=0}^{n} L(n,k)(x)_k \quad \text{and} \quad (x)_n = \sum_{k=0}^{n} (-1)^{n-k} L(n,k) x^{(k)}.$$

As above, this means they express the change of basis between the bases, $(x)_0, (x)_1, (x)_2, \cdots$ and $x^{(0)}, x^{(1)}, x^{(2)}, \cdots$ completing the diagram. In particular, one formula is the inverse of the other, thus:

$$\sum_j L(n,j) \cdot (-1)^{j-k} L(j,k) = \delta_{nk}.$$

Similarly, composing for example the change of basis from $x^{(n)}$ to x^n with the change of basis from x^n to $x_{(n)}$ gives the change of basis directly from $x^{(n)}$ to $x_{(n)}$

$$L(n,k) = \sum_j \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix},$$

In terms of matrices, if L denotes the matrix with entries $L_{nk} = L(n,k)$ and L^- denotes the matrix with entries $L^-_{nk} = (-1)^{n-k} L(n,k)$, then one is the inverse of the other: $L^- = L^{-1}$ Similarly, composing the matrix of unsigned Stirling numbers of the first kind with the matrix of Stirling numbers of the second kind gives the Lah numbers: $L = |s| \cdot S$ the numbers

$$\begin{Bmatrix} n \\ k \end{Bmatrix}, \begin{bmatrix} n \\ k \end{bmatrix}, L(n,k)$$

can be defined as the number of partitions of n elements into k non-empty unlabelled subsets, each of which is unordered, cyclically ordered, or linearly ordered, respectively. In particular, this implies the following inequalities:

$$\begin{Bmatrix} n \\ k \end{Bmatrix} \leq \begin{bmatrix} n \\ k \end{bmatrix} \leq L(n,k).$$

Symmetric Formulae

Abramowitz and Stegun give the following symmetric formulae that relate the Stirling numbers of the first and second kind.

$$s(n,k) = \sum_{j=0}^{n-k} (-1)^j \binom{n-1+j}{n-k+j} \binom{2n-k}{n-k-j} S(n-k+j, j)$$

and

$$S(n,k) = \sum_{j=0}^{n-k} (-1)^j \binom{n-1+j}{n-k+j} \binom{2n-k}{n-k-j} s(n-k+j, j).$$

Stirling Numbers with Negative Integral Values

The Stirling numbers can be extended to negative integral values, but not all do so in the same way. Regardless of the approach taken, it is worth noting that Stirling numbers of first and second kind are connected by the relations:

$$\begin{bmatrix} n \\ k \end{bmatrix} = \begin{Bmatrix} -k \\ -n \end{Bmatrix} \qquad \text{and} \qquad \begin{Bmatrix} n \\ k \end{Bmatrix} = \begin{bmatrix} -k \\ -n \end{bmatrix}$$

When n and k are nonnegative integers. So we have following table for

NOTES

$$\begin{bmatrix} -n \\ -k \end{bmatrix}$$

n \ k	−1	−2	−3	−4	−5
−1	1	1	1	1	1
−2	0	1	3	7	15
−3	0	0	1	6	25
−4	0	0	0	1	10
−5	0	0	0	0	1

Donald Knuth defined the more general Stirling numbers by extending a recurrence relation to all integers. In this approach, $\begin{bmatrix} n \\ k \end{bmatrix}$ and $\begin{Bmatrix} n \\ k \end{Bmatrix}$

are zero if n is negative and k is nonnegative, or if n is nonnegative and k is negative, and so we have, for any integers n and k,

$$\begin{bmatrix} n \\ k \end{bmatrix} = \begin{Bmatrix} -k \\ -n \end{Bmatrix} \qquad \text{and} \qquad \begin{Bmatrix} n \\ k \end{Bmatrix} = \begin{bmatrix} -k \\ -n \end{bmatrix}.$$

On the other hand, for positive integers n and k, David Branson defined

$$\begin{bmatrix} -n \\ -k \end{bmatrix}, \begin{Bmatrix} -n \\ -k \end{Bmatrix}, \begin{bmatrix} -n \\ k \end{bmatrix}, \text{ and } \begin{Bmatrix} -n \\ k \end{Bmatrix} \text{ (but not } \begin{bmatrix} n \\ -k \end{bmatrix} \text{ or } \begin{Bmatrix} n \\ -k \end{Bmatrix}\text{)}.$$

In this approach, one has the following extension of the recurrence relation of the Stirling numbers of the first kind:

$$\begin{bmatrix} -n \\ k \end{bmatrix} = \frac{(-1)^{n+1}}{n!} \sum_{i=1}^{n} (-1)^{i+1} \frac{\binom{n}{i}}{i^k}.$$

For example,

$$\begin{bmatrix} -5 \\ k \end{bmatrix} = \frac{5 - \frac{10}{2^k} + \frac{10}{3^k} - \frac{5}{4^k} + \frac{1}{5^k}}{120}$$

This leads to the following table of values of $\begin{bmatrix} -n \\ k \end{bmatrix}$.

NOTES

n \ k	0	1	2	3	4
−1	1	1	1	1	1
−2	$\frac{-1}{2}$	$\frac{-3}{4}$	$\frac{-7}{8}$	$\frac{-15}{16}$	$\frac{-31}{32}$
−3	$\frac{1}{6}$	$\frac{11}{36}$	$\frac{85}{216}$	$\frac{575}{1296}$	$\frac{3661}{7776}$
−4	$\frac{-1}{24}$	$\frac{-25}{288}$	$\frac{-415}{3456}$	$\frac{-5845}{41472}$	$\frac{-76111}{497664}$
−5	$\frac{1}{120}$	$\frac{137}{7200}$	$\frac{12019}{432000}$	$\frac{874853}{25920000}$	$\frac{58067611}{1555200000}$

In this case $\sum_{n=-1}^{-\infty} \begin{bmatrix} -n \\ -k \end{bmatrix} = B_k$ where B_k is a Bell number, and so one may define the negative Bell numbers by,

$$\sum_{n=-1}^{-\infty} \begin{bmatrix} -n \\ k \end{bmatrix} = B_{-k}$$

For example, this produces

$$\sum_{n=-1}^{-\infty} \begin{bmatrix} -n \\ 2 \end{bmatrix} = B_{-2} = 0.421773\ldots.$$

1.6 STIRLING NUMBERS OF THE FIRST KIND AND THE SECOND KIND

In mathematics, especially in combinatorics, Stirling numbers of the first kind arise in the study of permutations. In particular, the Stirling numbers of the first kind count permutations according to their number of cycles (counting fixed points as cycles of length one).

The Stirling numbers of the first and second kind can be understood as inverses of one another when viewed as triangular matrices. A triangular matrix is a special kind of square matrix. A square matrix is called lower triangular if all the entries above the main diagonal are zero. Similarly, a square matrix is called upper triangular if all the entries below the main diagonal are zero. In this unit, we will discuss about the Stirling numbers of the first kind.

Definitions

The original definition of Stirling numbers of the first kind was algebraic, they are the coefficients $s(n, k)$ in the expansion of the falling factorial

$$(x)_n = x(x-1)(x-2)\cdots(x-n+1)$$

NOTES

Into powers of the variable x:

$$(x)_n = \sum_{k=0}^{n} s(n,k)x^k,$$

For example,

$(x)_3 = x(x-1)(x-2) = 1 \cdot x^3 - 3 \cdot x^2 + 2 \cdot x$, leading to the values

$s(3,3) = 1$, $s(3,2) = -3$, and $s(3,1) = 2$.

Subsequently, it was discovered that the absolute values $|s(n, k)|$ of these numbers are equal to the number of permutations of certain kinds. These absolute values, which are known as unsigned Stirling numbers of the first kind, are often denoted

$$c(n,k) \text{ or } \begin{bmatrix} n \\ k \end{bmatrix}.$$

They may be defined directly to be the number of permutations of n elements with k disjoint cycles. For example, of the $3! = 6$ permutations of three elements, there is one permutation with three cycles (the identity permutation, given in one-line notation by 123 or in cycle notation by (1) (2) (3)), three permutations with two cycles (132 = (1) (23), 213 = (12) (3), and 321 = (13) (2)) and two permutations with one cycle (321 = (132) and 231 = (123)).

Thus, $\begin{bmatrix} 3 \\ 3 \end{bmatrix} = 1$, $\begin{bmatrix} 3 \\ 2 \end{bmatrix} = 3$ and $\begin{bmatrix} 3 \\ 1 \end{bmatrix} = 2$.

These can be seen to agree with the previous calculation of $s(n.k)$ for $n = 3$.

The unsigned Stirling numbers may also be defined algebraically, as the coefficients of the rising factorial:

$$x^{\overline{n}} = x(x+1)\cdots(x+n-1) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} x^k.$$

The notations used on this page for Stirling numbers are not universal, and may conflict with notations in other sources. (The square bracket notation $\begin{bmatrix} n \\ k \end{bmatrix}$ is also common notation for the Gaussian coefficients.)

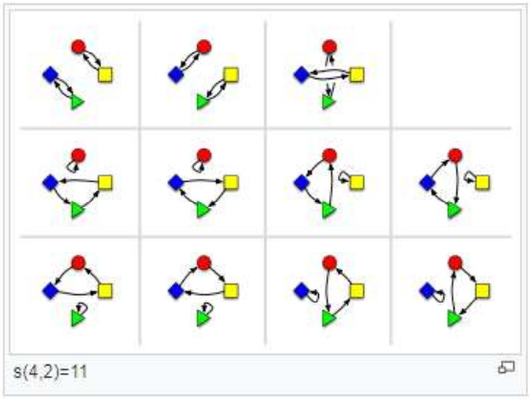
Example 2

The image at right shows that $\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 11$ the symmetric group on 4 objects has 3 permutations of the form (• •) (•,•) (having 2 orbits, each of size 2), and 8 permutations of the form (• • •) (•) (having 1 orbit of size 3 and 1 orbit of size 1).

Signs

The signs of the (signed) Stirling numbers of the first kind are predictable and depend on the parity of $n – k$. In particular,

$$s(n,k) = (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix}.$$



NOTES

Recurrence Relation

The unsigned Stirling numbers of the first kind can be calculated by the recurrence relation,

$$\begin{bmatrix} n+1 \\ k \end{bmatrix} = n \begin{bmatrix} n \\ k \end{bmatrix} + \begin{bmatrix} n \\ k-1 \end{bmatrix}$$

For $k > 0$, with the initial conditions

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1 \quad \text{and} \quad \begin{bmatrix} n \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ n \end{bmatrix} = 0$$

For $n > 0$.

It follows immediately that the (signed) Stirling numbers of the first kind satisfy the recurrence

$$s(n+1,k) = -n \cdot s(n,k) + s(n,k-1).$$

Algebraic Proof

We prove the recurrence relation using the definition of Stirling numbers in terms of rising factorials. Distributing the last term of the product, we have

$$x^{\overline{n+1}} = x(x+1)\cdots(x+n-1)(x+n) = n \cdot x^{\overline{n}} + x \cdot x^{\overline{n}}.$$

The coefficient of x^k on the left-hand side of this equation is,

$$\begin{bmatrix} n+1 \\ k \end{bmatrix}.$$

The coefficient of x^k in $n \cdot x^{\overline{n}}$ is $n \cdot \begin{bmatrix} n \\ k \end{bmatrix}$, while the coefficient of x^k in $x \cdot x^{\overline{n}}$ is $\begin{bmatrix} n \\ k-1 \end{bmatrix}$.

NOTES

Since the two sides are equal as polynomials, the coefficients of x^k on both sides must be equal, and the result follows.

Combinatorial Proof

We prove the recurrence relation using the definition of Stirling numbers in terms of permutations with a given number of cycles (or equivalently, orbits).

Consider forming a permutation of $n + 1$ objects from a permutation of n objects by adding a distinguished object. There are exactly two ways in which this can be accomplished. We could do this by forming a singleton cycle, i.e., leaving the extra object alone. This increases the number of cycles by 1 and so accounts for the $\begin{bmatrix} n \\ k-1 \end{bmatrix}$ term in the recurrence formula. We could also insert the new Object into one of the existing cycles. Consider an arbitrary permutation of n objects with k cycles, and label the objects $a_1, \ldots a_n$, so that the permutation is represented by,

$$\underbrace{(a_1 \ldots a_{j_1})(a_{j_1+1} \ldots a_{j_2}) \ldots (a_{j_{k-1}+1} \ldots a_n)}_{k \text{ cycles}}.$$

To form a new permutation of $n + 1$ objects and k cycles one must insert the new object $n \begin{bmatrix} n \\ k \end{bmatrix}$ into this array. There are n ways to perform this insertion, inserting the new object immediately following any of the n already present. This explains the term of the recurrence relation. These two cases include all possibilities, so the recurrence relation follows.

Table of Values

Below is a triangular array of unsigned values for the Stirling numbers of the first kind, similar in form to Pascal's triangle. These values are easy to generate using the recurrence relation in the previous section.

n \ k	0	1	2	3	4	5	6	7	8	9
0	1									
1	0	1								
2	0	1	1							
3	0	2	3	1						
4	0	6	11	6	1					
5	0	24	50	35	10	1				
6	0	120	274	225	85	15	1			
7	0	720	1764	1624	735	175	21	1		
8	0	5040	13068	13132	6769	1960	322	28	1	
9	0	40320	109584	118124	67284	22449	4536	546	36	1

NOTES

Properties

Simple Identities

Note that although

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1, \text{ we have } \begin{bmatrix} n \\ 0 \end{bmatrix} = 0 \text{ if } n > 0$$

And

$$\begin{bmatrix} 0 \\ k \end{bmatrix} = 0 \text{ if } k > 0 \text{ or more generally } \begin{bmatrix} n \\ k \end{bmatrix} = 0 \text{ if } k > n.$$

also

$$\begin{bmatrix} n \\ 1 \end{bmatrix} = (n-1)!, \quad \begin{bmatrix} n \\ n \end{bmatrix} = 1, \quad \begin{bmatrix} n \\ n-1 \end{bmatrix} = \binom{n}{2},$$

And

$$\begin{bmatrix} n \\ n-2 \end{bmatrix} = \frac{1}{4}(3n-1)\binom{n}{3} \quad \text{and} \quad \begin{bmatrix} n \\ n-3 \end{bmatrix} = \binom{n}{2}\binom{n}{4}.$$

Similar relationships involving the Stirling numbers hold for the Bernoulli polynomials. Many relations for the Stirling numbers shadow similar relations on the binomial coefficients. The study of these 'Shadow Relationships' is termed umbral calculus and culminates in the theory of Sheffer sequences. Generalizations of the Stirling numbers of both kinds to arbitrary complex-valued inputs may be extended through the relations of these triangles to the Stirling convolution polynomials.

Combinatorial Proofs

These identities may be derived by enumerating permutations directly. For example, a permutation of n elements with $n-3$ cycles must have one of the following forms:

- $n-6$ fixed points and three two-cycles
- $n-5$ fixed points, a three-cycle and a two-cycle
- $n-4$ fixed points and a four-cycle.

The three types may be enumerated as follows:

- Choose the six elements that go into the two-cycles, decompose them into two-cycles and take into account that the order of the cycles is not important:

  $$\binom{n}{6}\binom{6}{2,2,2}\frac{1}{6}$$

- Choose the five elements that go into the three-cycle and the two-cycle, choose the elements of the three-cycle and take into account that three elements generate two three-cycles:

NOTES

$$\binom{n}{5}\binom{5}{3} \times 2$$

- Choose the four elements of the four-cycle and take into account that four elements generate six four-cycles:

$$\binom{n}{4} \times 6.$$

Sum the three contributions to obtain,

$$\binom{n}{6}\binom{6}{2,2,2}\frac{1}{6} + \binom{n}{5}\binom{5}{3} \times 2 + \binom{n}{4} \times 6 = \binom{n}{2}\binom{n}{4}.$$

Other Relations

Expansions for Fixed *k*

Since the Stirling numbers are the coefficients of a polynomial with roots 0, 1, ..., $n-1$, one has by Vieta's formulas that,

$$\begin{bmatrix} n \\ n-k \end{bmatrix} = \sum_{0 \le i_1 < i_2 < \ldots < i_k < n} i_1 i_2 \cdots i_k.$$

In other words, the Stirling numbers of the first kind are given by elementary symmetric polynomials evaluated at 0, 1, ..., $n-1$. In this form, the simple identities given above take the form,

$$\begin{bmatrix} n \\ n-1 \end{bmatrix} = \sum_{i=0}^{n-1} i = \binom{n}{2},$$

$$\begin{bmatrix} n \\ n-2 \end{bmatrix} = \sum_{i=0}^{n-1}\sum_{j=0}^{i-1} ij = \frac{3n-1}{4}\binom{n}{3},$$

$$\begin{bmatrix} n \\ n-3 \end{bmatrix} = \sum_{i=0}^{n-1}\sum_{j=0}^{i-1}\sum_{k=0}^{j-1} ijk = \binom{n}{2}\binom{n}{4},$$

And so on.

One may produce alternative forms for the Stirling numbers of the first kind with a similar approach preceded by some algebraic manipulation: since

$$(x+1)(x+2)\cdots(x+n-1) = (n-1)! \cdot (x+1)\left(\frac{x}{2}+1\right)\cdots\left(\frac{x}{n-1}+1\right),$$

It follows from Newton's formulas that one can expand the Stirling numbers of the first kind in terms of generalized harmonic numbers. This yields identities like,

NOTES

$$\begin{bmatrix} n \\ 2 \end{bmatrix} = (n-1)!\, H_{n-1},$$

$$\begin{bmatrix} n \\ 3 \end{bmatrix} = \frac{1}{2}(n-1)! \left[(H_{n-1})^2 - H_{n-1}^{(2)} \right]$$

$$\begin{bmatrix} n \\ 4 \end{bmatrix} = \frac{1}{3!}(n-1)! \left[(H_{n-1})^3 - 3H_{n-1}H_{n-1}^{(2)} + 2H_{n-1}^{(3)} \right],$$

$$H_n = \frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{n}$$

Where H_n is the harmonic number $H_n = \frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{n}$ and $H_n^{(m)}$ is the generalized harmonic number,

$$H_n^{(m)} = \frac{1}{1^m} + \frac{1}{2^m} + \ldots + \frac{1}{n^m}.$$

These relations can be generalized to give,

$$\frac{1}{(n-1)!} \begin{bmatrix} n \\ k+1 \end{bmatrix} = \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^{n-1} \cdots \sum_{i_k=i_{k-1}+1}^{n-1} \frac{1}{i_1 i_2 \cdots i_k} = \frac{w(n,k)}{k!}$$

Where $w(n, m)$ is defined recursively in terms of the generalized harmonic numbers by,

$$w(n,m) = \delta_{m,0} + \sum_{k=0}^{m-1} (1-m)_k H_{n-1}^{(k+1)} w(n, m-1-k).$$

(Here δ is the Kronecker delta function and $(m)_k$ is the Pochhammer symbol.)

For fixed $n \geq 0$ these weighted harmonic number expansions are generated by the generating function,

$$\frac{1}{n!} \begin{bmatrix} n+1 \\ k \end{bmatrix} = [x^k] \exp\left(\sum_{m \geq 1} \frac{(-1)^{m-1} H_n^{(m)}}{m} x^m \right),$$

Where the notation $[x^k]$ means extraction of the coefficient of x^k from the following formal power series.

More generally, sums related to these weighted harmonic number expansions of the Stirling numbers of the first kind can be defined through generalized zeta series transforms of generating functions.

One can also "invert" the relations for these Stirling numbers given in terms of the k- order harmonic numbers to write the integer-order generalized harmonic numbers in terms of weighted sums of terms involving the Stirling numbers of the

NOTES

first kind. For example, when $k = 2, 3$ the second-order and third-order harmonic numbers are given by,

$$(n!)^2 \cdot H_n^{(2)} = \begin{bmatrix} n+1 \\ 2 \end{bmatrix}^2 - 2 \begin{bmatrix} n+1 \\ 1 \end{bmatrix} \begin{bmatrix} n+1 \\ 3 \end{bmatrix}$$

$$(n!)^3 \cdot H_n^{(3)} = \begin{bmatrix} n+1 \\ 2 \end{bmatrix}^3 - 3 \begin{bmatrix} n+1 \\ 1 \end{bmatrix} \begin{bmatrix} n+1 \\ 2 \end{bmatrix} \begin{bmatrix} n+1 \\ 3 \end{bmatrix} + 3 \begin{bmatrix} n+1 \\ 1 \end{bmatrix}^2 \begin{bmatrix} n+1 \\ 4 \end{bmatrix}.$$

More generally, one can invert the Bell polynomial generating function for the Stirling numbers expanded in terms of the *m*-order harmonic numbers to obtain that for integers $m \geq 2$

$$H_n^{(m)} = -m \times [x^m] \log \left(1 + \sum_{k \geq 1} \begin{bmatrix} n+1 \\ k+1 \end{bmatrix} \frac{(-x)^k}{n!} \right).$$

Factorial-Related Sums

For all positive integer *m* and *n*, one has

$$(n+m)! = \sum_{k=0}^{n} \sum_{j=0}^{m} m^{\overline{n-k}} \begin{bmatrix} m \\ j \end{bmatrix} \binom{n}{k} k!,$$

Where $a^{\bar{b}} = a(a+1) \cdots (a+b-1)$ is the rising factorial. This formula is a dual of Spivey's result for the Bell numbers.

Other related formulas involving the falling factorials, Stirling numbers of the first kind, and in some cases Stirling numbers of the second kind include the following:

$$n^{\underline{n-m}} = \sum_k \begin{bmatrix} n+1 \\ k+1 \end{bmatrix} \begin{Bmatrix} k \\ m \end{Bmatrix} (-1)^{m-k}$$

$$\binom{n}{m} (n-1)^{\underline{n-m}} = \sum_k \begin{bmatrix} n \\ k \end{bmatrix} \begin{Bmatrix} k \\ m \end{Bmatrix}$$

$$\binom{n}{m} = \sum_k \begin{Bmatrix} n+1 \\ k+1 \end{Bmatrix} \begin{bmatrix} k \\ m \end{bmatrix} (-1)^{m-k}$$

$$\begin{bmatrix} n+1 \\ m+1 \end{bmatrix} = \sum_{0 \leq k \leq n} \begin{bmatrix} k \\ m \end{bmatrix} n^{\underline{n-k}}.$$

Inversion Relations and the Stirling Transform

For any pair of sequences $\{f_n\}$ and $\{g_n\}$, related by a finite sum Stirling number formula given by,

$$g_n = \sum_{k=1}^{n} \begin{Bmatrix} n \\ k \end{Bmatrix} f_k,$$

NOTES

For all integers $n \geq 0$, we have a corresponding inversion formula for f_n given by,

$$f_n = \sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix} (-1)^{n-k} g_k.$$

These inversion relations between the two sequences translate into functional equations between the sequences exponential generating functions given by the Stirling (generating function) transform as,

$$\widehat{G}(z) = \widehat{F}\left(e^z - 1\right)$$

and

$$\widehat{F}(z) = \widehat{G}\left(\log(1+z)\right).$$

The differential operators $D = d/dx$ and $\vartheta = zD$ are related by the following formulas for all integers $n \geq 0$:

$$\frac{1}{k!}\frac{d^k}{dx^k} f(x) = \sum_{n=k}^{\infty} \frac{s(n,k)}{n!} \Delta^n f(x)$$

$$\frac{1}{k!}\Delta^k f(x) = \sum_{n=k}^{\infty} \frac{S(n,k)}{n!} \frac{d^n}{dx^n} f(x).$$

Congruences

The following congruence identity may be proved via a generating function-based approach:

$$\begin{bmatrix} n \\ m \end{bmatrix} \equiv \binom{\lfloor n/2 \rfloor}{m - \lceil n/2 \rceil} = [x^m]\left(x^{\lceil n/2 \rceil}(x+1)^{\lfloor n/2 \rfloor}\right) \pmod 2,$$

More recent results providing Jacobi-type J-fractions that generate the single factorial function and generalized factorial-related products lead to other new congruence results for the Stirling numbers of the first kind. For example, working modulo 2 we can prove that,

$$\begin{bmatrix} n \\ 1 \end{bmatrix} \equiv \frac{2^n}{4}[n \geq 2] + [n = 1] \pmod 2$$

$$\begin{bmatrix} n \\ 2 \end{bmatrix} \equiv \frac{3 \cdot 2^n}{16}(n-1)[n \geq 3] + [n = 2] \pmod 2$$

$$\begin{bmatrix} n \\ 3 \end{bmatrix} \equiv 2^{n-7}(9n-20)(n-1)[n \geq 4] + [n = 3] \pmod 2$$

$$\begin{bmatrix} n \\ 4 \end{bmatrix} \equiv 2^{n-9}(3n-10)(3n-7)(n-1)[n \geq 5] + [n = 4] \pmod 2$$

$$\begin{bmatrix} n \\ 5 \end{bmatrix} \equiv 2^{n-13}(27n^3 - 279n^2 + 934n - 1008)(n-1)[n \geq 6] + [n = 5] \pmod 2$$

$$\begin{bmatrix} n \\ 6 \end{bmatrix} \equiv \frac{2^{n-15}}{5}(9n^2 - 71n + 120)(3n-14)(3n-11)(n-1)[n \geq 7] + [n = 6] \pmod 2,$$

NOTES

and working modulo 3 we can similarly prove that,

$$\begin{bmatrix} n \\ 1 \end{bmatrix} \equiv \sum_{j=\pm 1} \frac{1}{36}\left(9 - 5j\sqrt{3}\right) \times \left(3 + j\sqrt{3}\right)^n [n \geq 2] + [n = 1] \pmod 3$$

$$\begin{bmatrix} n \\ 2 \end{bmatrix} \equiv \sum_{j=\pm 1} \frac{1}{216}\left((44n - 41) - (25n - 24) \cdot j\sqrt{3}\right) \times \left(3 + j\sqrt{3}\right)^n [n \geq 3] + [n = 2] \pmod 3$$

$$\begin{bmatrix} n \\ 3 \end{bmatrix} \equiv \sum_{j=\pm 1} \frac{1}{15552}\left((1299n^2 - 3837n + 2412) - (745n^2 - 2217n + 1418) \cdot j\sqrt{3}\right) \times \left(3 + j\sqrt{3}\right)^n [n \geq 4] + [n = 3] \pmod 3$$

$$\begin{bmatrix} n \\ 4 \end{bmatrix} \equiv \sum_{j=\pm 1} \frac{1}{179936}\left((6409n^3 - 383778n^2 + 70901n - 37092) - (3690n^3 - 22374n^2 + 41088n - 21708) \cdot j\sqrt{3}\right) \times \left(3 + j\sqrt{3}\right)^n [n \geq 5] + [n = 4] \pmod 3$$

More generally, for fixed integers $h \geq 3$ if we define the ordered roots,

$$(\omega_{h,i})_{i=1}^{h-1} := \left\{ \omega_j : \sum_{i=0}^{h-1} \binom{h-1}{i} \frac{h!}{(i+1)!} (-\omega_j)^i = 0,\ 1 \leq j < h \right\},$$

Then we may expand congruences for these Stirling numbers defined as the coefficients,

$$\begin{bmatrix} n \\ m \end{bmatrix} = [R^m] R(R+1) \cdots (R+n-1),$$

in the following form where the functions, $p_{h,i}^{[m]}(n)$ denote fixed polynomials of degree *m* in *n* for each *h*, *m* and *i*:

$$\begin{bmatrix} n \\ m \end{bmatrix} = \left(\sum_{i=0}^{h-1} p_{h,i}^{[m]}(n) \times \omega_{h,i}^n \right) [n > m] + [n = m] \pmod h,$$

The *r*-order harmonic numbers and for the generalized factorial products, $p_n(\alpha, R) := R(R+\alpha) \cdots (R+(n-1)\alpha)$. In the previous examples, the notation denotes Iverson's convention.

Check Your Progress

1. What is meant by basic combinatorial numbers?
2. How an unordered sample dealing with combination without repetition?
3. Define the permutation with repetitions.
4. How do you define permutation when all the things are not different?
5. Define Stirling number.
6. What do you understand by Stirling numbers of the first kind?

1.7 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The number of combinations of *n* different things taken *r* at a time is

$${}^nC_r = \frac{n(n-1)(n-2)...(n-r+1)}{1, 2, 3...r}.$$

2. In any arrangement or sampling, if order does not matter then it is a combination. Here also we have two different cases, one in which repetition is allowed and another in which repetition is not allowed. We can think of a case of repetition being allowed is to think of coins in your pocket. Your pocket has coins of 1, 1, 1, 2, 5, 5… like that. The case when repetition is not allowed can be thought of as a sequence of number 1,2,3 which is considered one sample and here ordering is not important. If orders are taken then there are six number of arrangements like, 123, 132, 231, 213, 312, and 321. So unordered sample is like a set. It is a set of three elements 1,2,3. Here we will discuss those where repetition is not allowed and so we are dealing with combination without repetition.
3. The number of permutations of n different things taken r at a time when the things can be repeated any number of times is n^r.

   The total number of permutations of n things taken 1, 2, 3, . . ., r at a time when the things may be repeated any number of times is $n + n^2 + \ldots + n^r = n. \dfrac{n^r - 1}{n - 1}$
4. The number of permutations of n things taken together when all the things are not different is given by

   $$\frac{n!}{p!q!r!\ldots}$$

   Here, among the n things, p things of them are of one kind; q of them are of second kind, and so on.

   $(p + q + r + \ldots\ldots = n)$
5. Stirling numbers arise in a variety of analytic and combinatorial problems. They are named after James Stirling, who introduced them in the 18th century. Two different sets of numbers bear this name: the Stirling numbers of the first kind and the Stirling numbers of the second kind. Additionally, Lah numbers are sometimes referred to as Stirling numbers of the third kind. Each kind is detailed in its respective article, this one serving as a description of relations between them.
6. The original definition of Stirling numbers of the first kind was algebraic, they are the coefficients $s(n, k)$ in the expansion of the falling factorial

   $(x)_n = x(x-1)\ (x-2)\ \ldots\ (x-n+1)$

   Into powers of the variable x:

   $$(x)_n = \sum_{k=0}^{n} s(n,k)x^k,$$

For example, $(x)_3 = x(x-1)(x-2) = 1 \cdot x^3 - 3 \cdot x^2 + 2 \cdot x$,

Leading to the values $s(3,3) = 1$, $s(3,2) = -3$, and $s(3,1) = 2$.

NOTES

1.8 SUMMARY

- The number of combinations of *n* different things taken *r* at a time is

$${}^{n}C_{r} = \frac{n(n-1)(n-2)...(n-r+1)}{1, 2, 3...r}.$$

- In any arrangement or sampling, if order does not matter then it is a combination.
- The number of permutations of *n* different things taken *r* at a time when the things can be repeated any number of times is n^r.
- The total number of permutations of *n* things taken 1, 2, 3, . . ., *r* at a time when the things may be repeated any number of times is $n + n^2 + \ldots + n^r = n. \frac{n^r - 1}{n-1}$
- The number of permutations of *n* things taken together when all the things are not different is given by

$$\frac{n!}{p!q!r!...}$$

- Stirling numbers arise in a variety of analytic and combinatorial problems. they are named after James Stirling, who introduced them in the 18th century.
- Two different sets of numbers bear this name: the Stirling numbers of the first kind and the Stirling numbers of the second kind. Additionally, Lah numbers are sometimes referred to as Stirling numbers of the third kind. Each kind is detailed in its respective article, this one serving as a description of relations between them.
- A common property of all three kinds is that they describe coefficients relating three different sequences of polynomials that frequently arise in combinatorics. Moreover, all three can be defined as the number of partitions of n elements into *k* non-empty subsets, with different ways of counting orderings within each subset.
- The notation of brackets and braces, in analogy to binomial coefficients, was introduced in 1935 by Jovan Karamata and promoted later by Donald Knuth.

- The Lah numbers $L(n,k)=\binom{n-1}{k-1}\frac{n!}{k!}$ are sometimes called Stirling numbers of the third kind. By convention, $L(0, 0) = 1$ and $L(n, k) = 0$, if $(n > k)$ or $k = 0 < n$
- Stirling numbers of the first kind arise in the study of permutations. In particular, the Stirling numbers of the first kind count permutations according to their number of cycles (counting fixed points as cycles of length one).
- The unsigned Stirling numbers of the first kind can be calculated by the recurrence relation

$$\begin{bmatrix} n+1 \\ k \end{bmatrix} = n\begin{bmatrix} n \\ k \end{bmatrix} + \begin{bmatrix} n \\ k-1 \end{bmatrix}.$$

- The Stirling numbers of the first kind are given by elementary symmetric polynomials evaluated at 0, 1, ..., $n-1$.
- Sums related to these weighted harmonic number expansions of the Stirling numbers of the first kind can be defined through generalized zeta series transforms of generating functions.
- For any pair of sequences $\{f_n\}$ and $\{g_n\}$, related by a finite sum Stirling number formula given by $g_n = \sum_{k=1}^{n} \begin{Bmatrix} n \\ k \end{Bmatrix} f_k$,

1.9 KEY WORDS

- **Combinations:** The number of combinations of *n* different things taken *r* at a time is

$${}^{n}C_{r} = \frac{n(n-1)(n-2)...(n-r+1)}{1, 2, 3...r}.$$

- **Permutations with repetitions:** The number of permutations of *n* different things taken *r* at a time when the things can be repeated any number of times is *nr*.
- **Stirling number:** A common property of all three kinds is that they describe coefficients relating three different sequences of polynomials that frequently arise in combinatorics. Moreover, all three can be defined as the number of partitions of *n* elements into *k* non-empty subsets, with different ways of counting orderings within each subset.
- **Inverse matrices:** The Stirling numbers of the first and second kinds can be considered inverses of one another

NOTES

$$\sum_{j\geq 0} s(n,j)S(j,k) = \sum_{j\geq 0}(-1)^{n-j}\begin{bmatrix} n \\ j \end{bmatrix}\begin{Bmatrix} j \\ k \end{Bmatrix} = \delta_{nk}$$

and

$$\sum_{j\geq 0} S(n,j)s(j,k) = \sum_{j\geq 0}(-1)^{j-k}\begin{Bmatrix} n \\ j \end{Bmatrix}\begin{bmatrix} j \\ k \end{bmatrix} = \delta_{nk},$$

Where λ_{nk} is the Kronecker delta. These two relationships may be understood to be matrix inverse relationships.

- **Lah numbers:** The Lah number $L(n,k) = \binom{n-1}{k-1}\frac{n!}{k!}$ are sometimes called Stirling numbers of the third kind. By convention, $L(0, 0) = 1$ and $L(n, k) = 0$ if $(n > k)$ or $k = 0 < n$
- **Recurrence relation:** The unsigned Stirling numbers of the first kind can be calculated by the recurrence relation

$$\begin{bmatrix} n+1 \\ k \end{bmatrix} = n\begin{bmatrix} n \\ k \end{bmatrix} + \begin{bmatrix} n \\ k-1 \end{bmatrix}$$

1.10 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What do you understand by basic combinatorial numbers?
2. Define unordered samples with/without repetition.
3. What is meant by permutation with repetitions?
4. Define the term Stirling number.
5. What is the meaning of Stirling numbers of the first kind?

Long-Answer Questions

1. Elaborate the concept of basic combinatorial numbers.
2. Describe the unordered samples with/without repetition.
3. Discuss the permutations with repetitions.

4. Write short notes on each of the following:
   (a) Stirling number
   (b) Inverse matrices
5. Discuss the Stirling numbers of first kind with the help of example.

NOTES

1.11 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

UNIT 2 RECURRENCE FORMULA

NOTES

Structure

2.0 Introduction
2.1 Objectives
2.2 Recursion
2.3 Recursion and Iteration
2.4 Closed Form Expression
2.5 Recurrence Relations Obtained from Solutions
2.6 Recurrence Formula for S_n^m
2.7 Recurrence Formula for P_n^m
2.8 Answers to Check Your Progress Questions
2.9 Summary
2.10 Key Words
2.11 Self Assessment Questions and Exercises
2.12 Further Readings

2.0 INTRODUCTION

In mathematics, Recursion occurs when a thing is defined in terms of itself or of its type. Recursion is used in a variety of disciplines ranging from linguistics to logic. The most common application of recursion is in mathematics and computer science, where a function being defined is applied within its own definition. While this apparently defines an infinite number of instances (function values), it is often done in such a way that no infinite loop or infinite chain of references can occur.

In recursion and iteration there is another way to evaluate a function from its recursive definition. Instead of successively reducing the computation to evaluate of the function at small integers, we can start with the basis and successively apply the recursive definition to find the values of the function at successive larger integers. Such a procedure is called iterative procedure.

A recurrence relation is an equation that recursively defines a sequence or multidimensional array of values, once one or more initial terms are given; each further term of the sequence or array is defined as a function of the preceding terms. The term difference equation sometimes (and for the purposes of this article) refers to a specific type of recurrence relation. However, "difference equation" is frequently used to refer to any recurrence relation.

In this unit, you will study the recursion, recursion and iteration, closed form expression, recurrence relations obtained from solutions, recurrence formula for S_n^m and recurrence formula for P_n^m.

2.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Describe the basics concepts of recursion
- Discuss about the process of recursion and iteration
- Explain the closed form expression
- Analysis the recurrence relations obtained from solutions
- Discuss the recurrence formula for S_n^m
- Introduce the basic concepts of recurrence formula for P_n^m

2.2 RECURSION

The numbers in the sequence 0, 1, 2, 3, 5, 8, 13, 21,... in which each new term is the sum of the previous two terms are called the Fibonacci numbers. If we denote the $(n+1)$th Fibonacci number by f_n, we have

$$f_n = f_{n-1} + f_{n-2} \text{ for } n \geq 2$$

and $f_0 = 0$ and $f_1 = 1$. This is called a recursive definition, in which each element of the sequence is defined in terms of previous numbers in the sequence.

We can define a function with the set of non-negative integers as its domain by

(a) Specifying the value of the function at zero, and

(b) Giving a rule for finding its value at an integer from its values at smaller integers.

Such a function is called a recursively defined function or such a definition is called recursive definition. For example, consider the sequence of powers of 3 given by

$$a_n = 3^n \text{ for } n = 0,1,2...$$

This sequence can also be defined by giving the first term of the sequence, namely $a_0 = 1$ and a rule for finding a term of the sequence from the previous one, namely

$$a_{n+1} = 3a_n \text{ for } n = 0,1,2,....$$

Example 2.1: Find a recursive definition of binomial coefficients.

Solution: Denote the binomial coefficient ${}_nC_K$ by $C(n,k)$. Then the recursive definition for $C(n,k)$ where $n \geq 0$, $k \geq 0$ and $n \geq k$ is given by

$$C(n,0) = 1;$$

$$C(n,n) = 1; \text{ and}$$

$$C(n,k) = C(n-1, k) + C(n-1, k-1) \text{ for } n > k > 0.$$

NOTES

Example 2.2: Suppose that f is defined recursively by

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

Find $f(1), f(2), f(3)$ and $f(4)$.

Solution: From the recursive definition, it follows that

$$f(1) = 3f(0) + 2 = 3.2 + 2 = 8$$
$$f(2) = 3f(1) + 2 = 3.8 + 2 = 26$$
$$f(3) = 3f(2) + 2 = 3.26 + 2 = 80$$
$$f(4) = 3f(3) + 2 = 3.80 + 2 = 242$$

Example 2.3: Give a recursive definition of $f(n) = n!$.

Solution: Since $0! = 1$ and $(n+1)! = (n+1)n!$ the desired rule is

$$f(0) = 1; \text{ and}$$
$$f(n+1) = (n+1)f(n)$$

Example 2.4: Give a recursive definition of a^n where a is a non zero real number and n is non-negative integer.

Solution: It is easy to see that

$$a^0 = 1; \text{ and } \quad a^{n+1} = a.a^n \text{ for } n \geq 0$$

uniquely define a^n for all non-negative integers n.

Example 2.5: Give a recursive definition for polynomial expression.

Solution: An expression of the form

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$$

where a's are constants ($a_0 \neq 0$) and $n \geq 0$, is called polynomial in x of degree n.

Let S be a set of coefficients. Then recursive definition for polynomial is

(a) A zeroth degree (constant) polynomial is an element of S.

(b) For $n \geq 1$, an nth degree a polynomial expression is an expression of the form $q(x)x+a$ where $q(x)$ is an $(n-1)$th degree polynomial and $a \in S$.

Example 2.6: Using recursive definition for polynomial expression, prove that $f(n) = 3n^3 - 8n^2 + 2n + 4$ is a third-degree polynomial expression.

Solution: Give $f(n) = 3n^3 - 8n^2 + 2n + 4$

$$= (3n^2 - 8n + 2)\, n + 4 = ((3n - 8)\, n - 8)\, n + 4$$
$$= (((3)\, n - 8)\, n + 2)n + 4.$$

Now, 3 is a zeroth degree polynomial

$\Rightarrow (3)n - 8$ is a first degree polynomial

$\Rightarrow ((3)n - 8)\, n + 2$ is a second degree polynomial

$\Rightarrow f(n) = (((3)n - 8)\, n + 2)\, n + 4$ is a third degree polynomial.

Note: The final expression for

$$f(n) = (((3)n - 8)\, n + 2)\, n + 4$$

is called its telescoping form. If we use it to calculate $f(6)$, we need only three multiplications and three additions or subtractions. This is called Horner's method for evaluating a polynomial expression.

Example 2.7: Write $p(x) = x^3 - 6x^2 + 11x - 6$ in telescoping form.

Solution: The telescoping form of

$$p(x) = x^3 - 6x^2 + 11x - 6$$
$$= (((1)x - 6)\, x + 11)\, x - 6$$

Example 2.8: Write $g(x) = -x^5 + 3x^4 + 3x^3 + 2x^2 + x$ in telescoping form.

Solution: The telescoping form of

$$g(x) = x^5 + 3x^4 + 3x^3 + 2x^2 + x$$
$$= (((((-1)\, x + 3)\, x + 3)\, x + 2)\, x + 1)\, x + 0.$$

2.3 RECURSION AND ITERATION

There is another way to evaluate a function from its recursive definition. Instead of successively reducing the computation to evaluate of the function at small integers, we can start with the basis and successively apply the recursive definition to find the values of the function at successive larger integers. Such a procedure is called iterative procedure.

To find $n!$ using an iterative procedure, we start with 1, the value of the factorial function at 0, and multiply successively by each positive integer less than or equal to n.

Note: Often an iterative approach for the evaluation of a recursively defined sequence requires much less computation than a procedure using recursion. But it is sometimes preferable to use a recursive procedure even if it is less efficient than the iterative procedure.

2.4 CLOSED FORM EXPRESSION

Let $f(x_1, x_2, \ldots, x_n)$ be an algebraic expression involving variables. $x_1, x_2, \ldots, x_n$ which are allowed to take values from some predetermined set. f is a closed form expression if there exists a number F such that evaluation of f with any allowed values $x_1, x_2 \ldots, x_n$ will take no more than F operations.

For example, (i) A closed form expression for

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}$$

NOTES

(ii) A closed form expression for

$$1^2 + 2^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{2}.$$

NOTES

Sequence of Integers

A sequence of integers is a function from the natural numbers into integers. That is, if f is a sequence of integers, then $f: N \to Z$ is a function. We use the notation f_n or $f(n)$ to denote the image of any natural n. We call f_n an nth term of the sequence.

Note: A sequence is often called a discrete function.

Example 2.9: Prove by induction that $f(k) = 3k + 2$, $k \geq 0$ is a closed form expression of the sequence f defined recursively by $f(0) = 2$ and $f(k) = f(k-1) + 3$ for $k \geq 1$.

Solution:

1. Basis of induction: If $k = 0$, then $f(0) = 3(0) + 2 = 2$, as defined.
2. Induction step: Now assume that for some $k > 0$,

$$f(k) = 3k + 2 \qquad (1)$$

we have,

$$f(k+1) = f(k) + 3 \text{ by the recursive definition}$$
$$= 3k + 2 + 3 = 3k + 3 + 2$$
$$= 3(k + 1) + 2$$

Therefore, by the principle of mathematical induction the result follows. Hence $f(k) = 3k + 2$, $k \geq 0$ is a closed form expression for the sequence f.

Example 2.10: Define the sequence of numbers f by $f_0 = 100$ and $f_n = 1.08\, f_{n-1}$ for, $n \geq 1$. Prove by induction that $f_n = 100\,(1.08)^n$, $n \geq 0$ is a closed form of expression for the sequence f.

Solution:

1. Basis of induction: If $k = 0$, then $f_0 = 100(1.08)^0 = 100$, as defined.
2. Induction step: Now assume that for some $k > 0$

$$f_k = 100(1.08)^k$$

Then

$$f_{(k+1)} = (1.08) f_k \text{ by the recursive definition}$$
$$= (1.08)100^k$$
$$= 100(1.08)^{k+1} \qquad \text{by Equation (1)}$$

Therefore, by the principle of mathematical induction the result follows. Hence $f_n = 100(1.08)^n$, $n \geq 0$ is a closed form expression for the sequence f.

Notes:

1. The sequence $f: N \to Z$ defined by $f(n) = n^3 - n$, is a sequence of integers.

2. The sequence $f: N \to Z$ defined by $f(0) = 1$ and $f(n) = f(n-1) + 2$ for $n \geq 1$ is a sequence of integers.
3. The codomain of a sequence can be any set. We use the notation $\{f_n\}$ to describe the sequence. We describe sequences by listing the terms of the sequence in the order of increasing subscripts.

For example,

(i) Consider the sequence $f: N \to S$ defined by $f(n) = f_n = \frac{1}{n+1}$. The list of the terms of this sequence $f_0, f_1, f_2, \ldots$ begins with
$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$

(ii) Consider the sequence $f: N \to Z$ defined by $f_n = 4^n$. The list of the terms of the sequence $f_0, f_1, f_2, \ldots$ begins with
1, 4, 16, 64,...

2.5 RECURRENCE RELATIONS OBTAINED FROM SOLUTIONS

Before giving an algorithm for solving a recurrence relations, we will examine a few recurrence relations that arise from certain closed form expressions. The procedure is illustrated by the following examples.

Example 2.11: Form the recurrence relation given $f_n = 3.5^n, n \geq 0$.

Solution: If $n \geq 1$, then

$$f_n = 3.5^n = 3.5.5^{n-1}$$
$$= 5.3.5^{n-1} = 5.f_{n-1}.$$

So the recurrence relation is $f_n - 5f_{n-1} = 0$ with $f_0 = 3$.

Example 2.12: Find the recurrence relation, satisfying $y_n = A(3)^n + B(-2)^n$.

Solution: Given $y_n = A(3)^n + B(-2)^n$

Therefore, $y_{n+1} = A(3)^{n+1} + B(-2)^{n+1} = 3A(3)^n - 2B(-2)^n$

and $y_{n+2} = A(3)^{n+2} + B(-2)^{n+2} = 9A(3)^n + 4B(-2)^n$

Eliminating A and B from these equation, we get

$$\begin{vmatrix} y_n & 1 & 1 \\ y_{n+1} & 3 & -2 \\ y_{n+2} & 9 & 4 \end{vmatrix} = 0$$

(or) $y_{n+2} - y_{n+1} - 6y_n = 0$ which is the required recurrence relation.

Example 2.13: Find the recurrence relation satifying $y_n = A(3)^n + B(-4)^n$.

Solution: Given $y_{n+1} = A(3)^n + B(-4)^n$

Therefore, $y_{n+1} = 3A(3)^n - 4B(-4)^n$

NOTES

and $y_{n+2} = 9A(3)^n + 16B(-4)^n$. Eliminating A and B from these equation, we get

$$\begin{vmatrix} y_n & 1 & 1 \\ y_{n+1} & 3 & -4 \\ y_{n+2} & 9 & 16 \end{vmatrix} = 0$$

(or) $y_{n+2} + y_{n+1} - 12y_n = 0$ which is the required recurrence relation.

Example 2.14: Find the recurrence relation satisfing $y_n = (A+Bn)4^n$

Solution: Given $y_n = (A + Bn)4^n$

$= A4^n + nB4^n$

Therefore, $y_{n+1} = 4A4^n + 4(n+1)B4^n$

and $y_{n+2} = 16A4^n + 16(n+2)B4^n$.

Eliminating A and B from these equation, we get

$$\begin{vmatrix} y_n & 1 & n \\ y_{n+1} & 4 & 4(n+1) \\ y_{n+2} & 16 & 16(n+2) \end{vmatrix} = 0$$

(or) $y_{n+2} - 8y_{n+1} - 16y_n = 0$, which is the required recurrence relation.

2.6 RECURRENCE FORMULA FOR S_n^m

Consider a linear homogeneous recurrence relations of degree k with constant coefficients

$$f_n = a_1 f_{n-1} + a_2 f_{n-2} + \ldots + a_k f_{n-k}$$

Where $a_1, a_2, \ldots, a_k$ are real numbers and $ak \neq 0$. The basic approach for solving linear homogeneous recurrence relations is to look for solutions of the form $f_n = r^n$, where r is a constant. Note that $f_n = r^n$ is solution of the recurrence relation $f_n = a_1 f_{n-1} + a_2 f_{n-2} + \ldots + a_k f_{n-k}$ if and only if

$$r^n = C_1 r^{n-1} + C_2 r^{n-2} + \ldots + C_k r^{n-k}$$

when both sides of this equation are divided by r^{n-k} and the right-hand side is substracted from the left, we obtain

$$r^k - C_1 r^{k-1} - C_2 r^{k-2} - \ldots - C_{k-1} T - C_k = 0$$

consequently, the sequence $\{f_n\}$ with $f_n = r^n$ is a solution if and only if r is a solution of this last equation,

Characteristic Equation: The characteristic equation of the homogeneous k^{th} order linear relation

$$f_n + a_1 f_{n-1} + a_2 f_{n-2} + \ldots + a_k f_{n-k} = 0$$

is the k^{th} degree polynomial equation

$$r^k + C_1 r^{k-1} + C_2 r^{k-2} + \cdots + C_{k-1} r + C_k = 0$$

The solutions of this equation are called the characteristic roots of the recurrence relation.

Example 2.15: What is the characteristic equation of

$$Q(k) + 2Q(k-1) - 3Q(k-2) - 6Q(k-4) = 0$$

Solution: The characteristic equation of

$$Q(k) + 2Q(k-1) - 3Q(k-2) - 6Q(k-4) = 0$$

is $r^4 + 2r^3 - 3r^2 - 6 = 0$.

Note that the absence of a $Q(k-3)$ term means that there is no $r^{4-3} = r$ term in the characteristic equation.

Example 2.16: What is the characteristic equation of $T(k) - 7T(k-2) + 6T(k-3) = 0$

The characteristic equation is $r^3 - 7r + 6 = 0$ and 1,2 and -3 are the characteristic roots.

Algorithm for Solving *k*th Order Homogeneous Linear Recurrence Relation.

Step 1. If $f_n + a_1 f_{n-1} + a_2 f_{n-2} + ... + a_k f_{n-k} = 0$ is a given recurrence relation, then write its characteristic equation.

That is, $r^k + C_1 r^{k-1} + C_2 r^{k-2} + ... + C_{k-1} r + C_k = 0$

Step 2. Find all the characteristic roots of this equation.

Step 3. *Case* (i) If there are k distinct roots, say $c_1, c_2, ..., c_k$ then the general solution of the recurrence relation is

$$f_n = A_1 c_1^{\,k} + A_2 c_r^{\,k} + ... + A_k c^k{}_k$$

Case (ii) Suppose that c_1 is a root of multiplicity m. Then the corresponding solution is

$$f_n = (A_1 r^{m-1} + A_2 r^{m-2} + ... + A_{m-2} r^2 + A_{m-1}{}^T{}_1 + A_m) c^T{}_1 + ... \ .$$

Step 4. Use the boundary conditions to determine the constants $A_1, A_2, ..., A_k$.

Example 2.17: Solve the Fibonacci sequence $\{f_n\}$ defined by $f_n = f_{n-1} + f_{n-2}$ for $n \geq 2$ with the initial conditions $f_0 = 1$ and $f_1 = 1$.

Solution: The first step is to form the characteristic equation corresponding to the given difference equation. In this case it is $r^2 - r - 1 = 0$

Solving, we get $c_1 = \frac{1+\sqrt{5}}{2}$, and $c_2 = \frac{1-\sqrt{5}}{2}$ the characteristic roots. It follows that the general solution is

$$f_n = A_1 c_1^{\,m} + A_2 c_2^{\,m}$$

where A_1 and A_2 are constants. Since $f_0 = 1$ and $f_1 = 1$, we get

$$0 = A_1 + A_2$$

and

$$1 = A_1 c_1 + A_2 c_2$$

NOTES

$$1 = A_1\left(\frac{1+\sqrt{5}}{2}\right) + A_2\left(\frac{1-\sqrt{5}}{2}\right)$$

NOTES

On solving, we get

$$A_1 = \frac{1}{\sqrt{5}} \quad \text{and} \quad A_2 = \frac{-1}{\sqrt{5}}$$

Hence the solution is

$$f_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

Example 2.18: If the recurrence relation is $u_{n+1} - 2u_n = 0$, find the closed form expression (solution) for u_n.

Solution: The characteristic equation of the given recurrence relation is $r-2=0$, i.e., $r=2$. Therefore, the general solution is $u_n = A2^n$. Hence $u_n = u_0 \cdot 2^n$ is the closed form expression, where the value of u_0 is the initial condition.

Example 2.19: Find $f(n)$ if $f(n) = 7f(n-1) - 10f(n-2)$, given that $f(0) = 4$ and $f(1) = 17$.

Solution: The characteristic equation of the given recurrence relation is $r^2 - 7r + 10 = 0$

It's characteristic roots are $r = 2, 5$. So the general solution of the recurrence relation is $f(n) = A_1(2)^n A_2(5)^n$.

Since $f(0) = 4$, $4 = A_1 + A_2$

Again, $f(1) = 17$ implies $17 = 2A_1 + 5A_2$

Solving, we get $A_1 = 1$ and $A_2 = 3$. Therefore, $f(n) = (2)^n + 3(5)^n$

Example 2.20: Find $T(k)$ if $T(k) - 7T(k-2) + 6T(k-3) = 0$, where $T(0) = 8$, $T(1) = 6$ and $T(2) = 22$.

Solution: The characteristic equation is $r^3 - 7r + 6 = 0$

Its roots are 1,2 and –3. Therefore, the general solution is $T(k) = A_1(1)^k + A_2(2)^k + A_3(-3)^k$.

Now

$$T(0) = 8 \Rightarrow A_1 + A_2 + A_3 = 8.$$
$$T(1) = 6 \Rightarrow A_1 + 2A_2 - 3A_3 = 6.$$
$$T(2) = 22 \Rightarrow A_1 + 4A_2 + 9A_3 = 22.$$

Solving, we get $A_1 = 5$, $A_2 = 2$ and $A_3 = 1$.

Hence,

$$T(k) = 5 + 2(2)^k + 1(-3)^k$$
$$= 5 + 2^{k+1} + (-3)^k$$

Example 2.21: Solve $f_k - 8f_{k-1} + 16f_{k-2} = 0$ where $f_2 = 16$ and $f_3 = 80$.

Solution: The characteristic equation is $r^2 - 8r + 16 = 0$ (or) $(r-4)^2 = 0$.

So $r = 4$ is a double characteristic root. Therefore, the general solution is $f_k = (A_1 + A_2 k)4^k$.

Now, $f_2 = 16 \Rightarrow (A_1 + 2A_2)16 = 16$

and $f_3 = 80 \Rightarrow (A_1 + 3A_2)64 = 80$

Solving, we get $A_1 = \frac{1}{2}$ and $A_2 = \frac{1}{4}$.

Hence the solution is

$$f_k = \left(\frac{1}{2} + \frac{1}{4}k\right)4k$$
$$= (2 + k)4^{k-1}$$

Example 2.22: Find a solution to the recurrence relation $C_n = -3C_{n-1} - 3C_{n-2} - C_{n-3}$ for $n \geq 3$ with initial conditions $C_0 = 1$, $C_1 = -2$ and $C_2 = 1$.

Solution: The characteristic equation is $r^3 + 3r^2 + 3r + 1 = 0$ (or) $(r+1)^3 = 0$.

So $r = -1$ is a characteristic root of multiplicity 3. Therefore, the general solution is

$$C_n = (A_1 + A_2 n + A_3 n^2)(-1)^n$$

Now, $C_0 = 1 \Rightarrow A_1 = 1$

$C_1 = -2 \Rightarrow -(A_1 + A_2 + A_3) = -2$

$C_2 = 1 \Rightarrow A_1 + 2A_2 + 4A_3 = 1$

Solving, we get $A_1 = 1$, $A_2 = 2$, $A_3 = 1$. Hence the solution is $C_n = (1 + 2n - n^2)(-1)^n$.

2.7 RECURRENCE FORMULA FOR P_n^m

The solution of a linear non homogeneous recurrence relation with constant coefficient is the sum of the two parts, the homogeneous solution, which satisfies the recurrence relation when the right hand side of the equation is set 1to 0, and the particular solution, which satisfies the difference equation with $f(n)$ on the right hand side.

There is no general procedure for determining the particular solution of a difference equation. However, in simple cases, this solution can be obtained by the method of inspection. To determine the particular solution we use the following rules.

***Rule* 1**. When $f(n)$ is of the form of a polynomial of degree m in n

$$k_0 + k_1 n + k_2 n^2 + \ldots + k_{m-1} n^{m-1} + k_m n^m$$

the corresponding particular solution will be of the form

$$Q_0 + Q_1 n + Q_2 n^2 + ... + Q_{m-1} n^{m-1} + Q_m n^m$$

***Rule* 2**. When $f(n)$ is of the form

$$(k_0 + k_1 n + k_2 n^2 + ... + k_{m+1} n^{m-1} + k_m n^m) a^n$$

NOTES

the corresponding particular solution is of the form

$$(Q_0 + Q_1 n + Q_2 n^2 + ... + Q_{m-1} n^{m-1} + Q_m n^m) a^n$$

if a is not a characteristic root of the recurrence relation.

Example 2.23: Solve $S(k) - S(k-1) - 6S(k-2) = -30$, where $S(0) = 20$, $S(1) = -5$.

Solution: The associated homogeneous relation is $S(k) - S(k-1) - 6S(k-2) = 0$.

The characteristic equation is $r^2 - r - 6 = 0$.

Its characteristic roots are $r = -2, 3$. So the homogeneous solution is $A_1(-2)^k + A_2(3)^k$.

Since the right hand side of $S(k) - S(k-1) - 6S(k-2) = -30$ (1) is a constant, by Rule 1, the particular solution will also be a constant, say Q into (1), we obtain

$$Q - Q - 6Q = -30 \Rightarrow -6Q = -30$$

$$\Rightarrow Q = 5$$

Therefore, the general solution is $S(k) = A_1(-2)^k + A_2(3)^k + 5$

Using the initial conditions, $S(0) = 20$, $S(1) = -5$

$$20 = A_1 + A_2 + 5$$

$$-5 = -2A_1 + 3A_2 + 5$$

which yield, $A_1 = 11$ and $A_2 = 4$. Hence the complete solution is $S(k) = 11(-2)^k + 4(3)^k + 5$.

***Rule* 3**. If a is a characteristic root of multiplicity $r-1$, when $f(n)$ is of the form

$$(k_0 + k_1 n + k_2 n^2 + ... + k_{m-1} n^{m-1} + k_m n^m) a^n$$

the corresponding particular solution is of the form

$$n^{r-1}(Q_0 + Q_1 n + Q_2 n^2 + ... + Q_{m-1} n^{m-1} + Q_m n^m) a^n.$$

Note: The general solution of the recurrence relation is the sum of the homogeneous solution and particular solution. If no initial conditions are given, then you have finished. If m initial conditions are given, obtain m linear equations in m unknowns and solve the system, if possible, to get a complete solution.

Example 2.24: Solve the recurrence relation $f_n - 5f_{n-1} + 6f_{n-2} = 1$.

Solution: The associated homogeneous relation is $f_n - 5f_{n-1} + 6f_{n-2} = 0$.

The characteristic equation is $r^2 - 5r + 6 = 0$

Its characteristic roots are $r = 2, 3$ so the homogeneous solution is

$$A_1(2)^n + A_2(3)^n \quad (1)$$

Since the right hand side $f(n) = 1$ is a constant, by Rule 1, the particular solution will also be a constant, say Q. Substituting Q into Equation (1), we obtain

$$Q - 5Q + 6Q = 1$$

This implies that $Q = \frac{1}{2}$. Therefore, the complete solution

$$f_n = A_1(2)^n + A_2(3)^n + \frac{1}{2}.$$

Example 2.25: Find the particular solution of the recurrence relation

$$f(n) + 5f(n-1) + 6f(n-2) = 3n^2 - 2n + 1 \quad (1)$$

Solution: By Rule 1, the particular solution is of the form

$$Q_0 + Q_1 n + Q_2 n^2 \quad (2)$$

Substituting Equation (2) into (1), we obtain

$$(Q_0 + Q_1 n + Q_2 n^2) + 5(Q_0 + Q_1 (n-1) + Q_2 (n-1)^2) + 6(Q_0 + Q_1(n-2) + Q_2(n-2)^2) = 3n^2 - 2n + 1$$

which simplifies to

$$(12Q_0 - 17Q_1 + 29Q_2) + (12Q_1 - 34Q_2)n + 12Q_2 n^2 = 3n^2 - 2n + 1 \quad (3)$$

Comparing the two sides of Equation (3), we obtain

$$12Q_2 = 3; \quad 12Q_1 - 34Q_2 = -2; \quad 12Q_0 - 17Q_1 + 29Q_2 = 1$$

which yields $Q_2 = \frac{1}{4}, Q_1 = \frac{13}{24}, Q_0 = \frac{71}{288}$. Therefore, the particular solution is

$\frac{71}{288} + \frac{13}{24}n + \frac{1}{4}n^2$.

Example 2.26: Solve $a_r + 5a_{r-1} = 9$ with initial condition $a_0 = 6$.

Solution: The associated homogeneous recurrence relation is $a_r + 5a_{r-1} = 0$.

It's characteristic equation is $r + 5 = 0$.

Therefore, $r = -5$. So the homogeneous solution is $A_1(-5)^r$. Since the right hand side of the given relation is a constant, the particular solution will also be a constant Q. Substituting in the recurrence relation, we get

$$Q + 5Q = 9$$

So $Q = \frac{3}{2}$. Therefore the general solution is $a^r = A_1(-5)^r + \frac{3}{2}$. Using the initial condition $a_0 = 6$, we get

$$6 = A_1 + \frac{3}{2} \Rightarrow A_1 = \frac{9}{2}$$

Hence the complete solution is

$$a_r = \frac{9}{2}(-5)^r + \frac{3}{2}$$

NOTES

Example 2.27: Solve the recurrence relation $f(n) - 7f(n-1) + 10f(n-2) = 6 + 8n$ with $f(0) = 1$ and $f(1) = 2$. (1)

Solution: The homogeneous solution is $A_1(2)^n + A(2)^n + A_2(5)^n$. By Rule 1, the particular solution is of the form $Q_0 + Q_1{}^n$.

Substituting in Equation (1), we get

$$(Q_0 + Q_1{}^n) - 7(Q_0 + Q_1(n-1)) + 10(Q_0 + Q_1(n-2)) = 6 + 8n$$

Comparing the two sides, we obtain

$$4Q_0 - 13Q_1 = 6 \quad \text{and} \quad 4Q_1 = 8$$

which yields, $Q_0 = 8$, $Q_1 = 2$. Therefore, the particular solution is $8 + 2n$ and the general solution is

$$f(n) = A_1(2)^n + A_2(5)^n + 8 + 2n$$

Now the initial conditions,

$$f(0) = 1 \Rightarrow A_1 + A_2 + 8 = 1$$

and

$$f(1) = 2 \Rightarrow 2A_1 + 5A_2 + 10 = 2$$

Solving, we get $A_1 = -9$ and $A_2 = 2$. Hence the complete solution is

$$f(n) = -9(2)^n + 2(5)^n + 8 + 2n$$

Example 2.28: Find the particular solution of the recurrence relation

$$a_n + 5a_{n-1} + 6a_{n-2} = 42(4)^n. \quad (1)$$

Solution: Now $r^2 + r + 6 = 0$, $r = -2, -3$ are characteristic roots. Since 4 is not a characteristic root, by Rule 2, we assume that the general form of the particular solution is $Q.(4)^n$. Substituting in Equation (1), we obtain

$$Q.(4)^n + 5Q.(4)^{n-1} + 6Q.(4)^{n-2} = 42(4)^n$$

$$\Rightarrow Q.4^{n-2}[16 + 20 + 6] = 42(4)^n \Rightarrow Q.4^{n-2}(42) = 42(4)^n$$

$$\Rightarrow Q = 4^2 = 16$$

Therefore, the particular solution is $16(4)^n = 4^{n+2}$.

Example 2.29: Find the particular solution of the recurrence relation

$$f_n + f_{n-1} = 3n2^n \quad (1)$$

Solution: The characteristic equation is $r + 1 = 0$. Therefore $r = -1$ is a characteristic root. Since 2 is not a characteristic root, by Rule 2, the general form of the particular solution is $(Q_0 + Q_1 n)2^n$. Substituting in Equation (1), we obtain

$$(Q_0 + Q_1 n)2^n + (Q_0 + Q_1(n-1))2^{n-1} = 3n2^n$$

which simplifies to

$$Q_0 2^n + Q_1 n2^n + \frac{1}{2} Q_0 2^n + \frac{1}{2} Q_1 n2^n - \frac{1}{2} Q_1 n2^n = 3n2^n$$

$$\Rightarrow (\frac{3}{2} Q_0 - \frac{1}{2} Q_1)2^n + \frac{3}{2} Q_1 n2^n = 3n2^n$$

Comparing the two sides, we obtain

$$\frac{3}{2} Q_0 - \frac{1}{2} Q_1 = 0$$

$$\frac{3}{2} Q_1 = 3$$

Thus, $Q_0 = \frac{2}{3}$ and $Q_1 = 2$ and the particular solution is $(\frac{2}{3} + 2n)2^n$.

Example 2. 30: Find the particular solution of the recurrence relation

$$f(n) - 2f(n-1) = 3.2^n \quad (1)$$

Solution: The characteristic equation is $r - 2 = 0$. Since $r = 2$ is the characteristic root of multiplicity 1, by Rule (3), the general form of the particular solution is $Q_n 2^n$. Substituting in Equation (1), we obtain

$$Qn2^n - 2Q.\ (n-1)2^{n-1} = 3.2^n$$

$$\Rightarrow Q \cdot 2^n = 3 \cdot 2^n$$

$$\Rightarrow Q = 3$$

Thus the particular solution is $3n2^n$.

Example 2.31: Find the general solution of

$$f(n) - 3f(n-1) - 4f(n-2) = 4^n \quad (1)$$

Solution: The associated homogeneous relation is

$$f(n) - 3f(n-1) - 4f(n-2) = 0.$$

Its characteristic equation is $r^2 - 3r - 4 = 0$.

Solving, we get $r = -1, 4$ the characteristic roots. Therefore, the homogeneous solution is $A_1(-1)^n + A_2(4)^n$. Since 4 is a characteristic root, by Rule 3, we assume that the general form of the particular solution is $Qn4^n$.

Substituting in (1), we get

$$Qn4^n - 3Q(n-1)4^{n-1} - 4Q(n-2)4^{n-2} = 4^n$$

$$\Rightarrow Qn4^n - 3Qn4^{n-1} + 3Q4^{n-1} - 4Qn4^{n-2} + 8Q4^{n-2} = 4^n$$

$$\Rightarrow (16Qn - 12Qn + 12Q - 4Qn + 8Q)4^{n-2} = (16)4^{n-2}$$

$$\Rightarrow 20Q = 16 \Rightarrow Q = \frac{4}{5}$$

Therefore, the particular solution is $\frac{4}{5} n4^n$. Hence, the general solution of the recurrence relation of

$$f(n) = A_1(-1)^n + A_2(4)^n + \frac{4}{5} n4^n$$

NOTES

Note: What if the characteristic equation gives rise to complex roots? Here, our methods are still valid, but the method for expressing the solutions of the recurrence relations is different. Since an understanding of these representations require some background in complex numbers, we suggest that an interested reader refer to a more advanced treatment of recurrence relations.

Check Your Progress

1. Define Fibonacci number.
2. What is meant by iteration procedure?
3. What do you understand by discrete function?
4. What is recurrence relation for the sequence $\{f_n\}$?
5. What is the solution of a linear non homogeneous recurrence relation?
6. Define the characteristic roots of the recurrence relation.
7. How do we define the recurrence formula for P_n^m?

2.8 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The numbers in the sequence 0, 1, 2, 3, 5, 8, 13, 21,... in which each new term is the sum of the previous two terms are called the Fibonacci numbers.
2. There is another way to evaluate a function from its recursive definition. Instead of successively reducing the computation to evaluate of the function at small integers, we can start with the basis and successively apply the recursive definition to find the values of the function at successive larger integers. Such a procedure is called iterative procedure.
3. A sequence of integers is a function from the natural numbers into integers. That is, if f is a sequence of integers, then $f: N \rightarrow Z$ is a function. We use the notation f_n or $f(n)$ to denote the image of any natural n. We call f_n an nth term of the sequence. A sequence is often called a discrete function.
4. A recurrence relation for the sequence $\{fn\}$ is a formula that express f_n in terms of one or more of the previous terms of the sequence, namely, f_0, f_1,...,f_{n-1}, for all integers n with $n \geq n_0$, when n_0 is a non-negative integer. A sequence is called a solution of a recurrence relation if its terms satisfy the recurrence relation.
5. The solution of a linear non homogeneous recurrence relation with constant coefficient is the sum of the two parts, the homogeneous solution, which satisfies the recurrence relation when the right-hand side of the equation is set 1 to 0, and the particular solution, which satisfies the difference equation with $f(n)$ on the right-hand side.

6. The characteristic equation of the homogeneous k^{th} order linear relation $f_n+a_1f_{n-1}+a_2f_{n-2}+...+a_kf_{n-k} = 0$ is the k^{th} degree polynomial equation $r^k + C_1r^{k-1} + C_2r^{k-2} + \cdots + C_{k-1}r + C_k = 0$

   The solutions of this equation are called the characteristic roots of the recurrence relation.

7. The solution of a linear non homogeneous recurrence relation with constant coefficient is the sum of the two parts, the homogeneous solution, which satisfies the recurrence relation when the right-hand side of the equation is set 1 to 0, and the particular solution, which satisfies the difference equation with $f(n)$ on the right-hand side.

2.9 SUMMARY

- The numbers in the sequence 0, 1, 2, 3, 5, 8, 13, 21,... in which each new term is the sum of the previous two terms are called the Fibonacci numbers.
- There is another way to evaluate a function from its recursive definition. Instead of successively reducing the computation to evaluate of the function at small integers, we can start with the basis and successively apply the recursive definition to find the values of the function at successive larger integers. Such a procedure is called iterative procedure.
- Often an iterative approach for the evaluation of a recursively defined sequence requires much less computation than a procedure using recursion. But it is sometimes preferable to use a recursive procedure even if it is less efficient than the iterative procedure.
- A sequence of integers is a function from the natural numbers into integers. That is, if f is a sequence of integers, then $f: N \rightarrow Z$ is a function. We use the notation f_n or $f(n)$ to denote the image of any natural n. We call f_n an nth term of the sequence. A sequence is often called a discrete function.
- The codomain of a sequence can be any set. We use the notation $\{f_n\}$ to describe the sequence. We describe sequences by listing the terms of the sequence in the order of increasing subscripts.
- Consider a linear homogeneous recurrence relations of degree k with constant coefficients

$$f_n = a_1f_{n-1} + a_2f_{n-2}+...+a_kf_{n-k}$$

- The characteristic equation of the homogeneous k^{th} order linear relation

$$f_n+a_1f_{n-1}+a_2f_{n-2}+...+a_kf_{n-k} = 0$$

is the k^{th} degree polynomial equation

$$r^k + C_1r^{k-1} + C_2r^{k-2} + \cdots + C_{k-1}r + C_k = 0$$

NOTES

The solutions of this equation are called the characteristic roots of the recurrence relation.

- The solution of a linear non homogeneous recurrence relation with constant coefficient is the sum of the two parts, the homogeneous solution, which satisfies the recurrence relation when the right-hand side of the equation is set 1to 0, and the particular solution, which satisfies the difference equation with $f(n)$ on the right-hand side.
- When $f(n)$ is of the form of a polynomial of degree m in n

$$k_0 + k_1 n + k_2 n^2 + \ldots + k_{m-1} n^{m-1} + k_m n^m$$

the corresponding particular solution will be of the form

$$Q_0 + Q_1 n + Q_2 n^2 + \ldots + Q_{m-1} n^{m-1} + Q_m n^m$$

- When $f(n)$ is of the form

$$(k_0 + k_1 n + k_2 n^2 + \ldots + k_{m+1} n^{m-1} + k_m n^m) a^n$$

the corresponding particular solution is of the form

$$(Q_0 + Q_1 n + Q_2 n^2 + \ldots + Q_{m-1} n^{m-1} + Q_m n^m) a^n$$

if a is not a characteristic root of the recurrence relation.

- If a is a characteristic root of multiplicity $r - 1$, when $f(n)$ is of the form $(k_0 + k_1 n + k_2 n^2 + \ldots + k_{m-1} n^{m-1} + k_m n^m) a^n$ the corresponding particular solution is of the form

$$n^{r-1}(Q_0 + Q_1 n + Q_2 n^2 + \ldots + Q_{m-1} n^{m-1} + Q_m n^m) a^n.$$

- The general solution of the recurrence relation is the sum of the homogeneous solution and particular solution. If no initial conditions are given, then you have finished. If m initial conditions are given, obtain m linear equations in m unknowns and solve the system, if possible, to get a complete solution.

2.10 KEY WORDS

- **Fibonacci numbers:** The numbers in the sequence 0, 1, 2, 3, 5, 6, 13, 21,… in which each new term is the sum of the previous two terms are called the Fibonacci numbers.
- **Iterative procedure:** There is another way to evaluate a function from its recursive definition. Instead of successively reducing the computation to evaluate of the function at small integers, we can start with the basis and successively apply the recursive definition to find the values of the function at successive larger integers. Such a procedure is called iterative procedure.
- **Discrete function:** A sequence of integers is a function from the natural numbers into integers. That is, if f is a sequence of integers, then $f: N \to Z$ is a function. We use the notation f_n or $f(n)$ to denote the image of any natural n. We call f_n an nth term of the sequence.

- **Recurrence relation:** The general solution of the recurrence relation is the sum of the homogeneous solution and particular solution. If no initial conditions are given, then you have finished. If *m* initial conditions are given, obtain *m* linear equations in *m* unknowns and solve the system, if possible, to get a complete solution.

2.11 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define the term recursive.
2. How do we define recursion and iteration?
3. What do you understand by sequence of integers?
4. State the recurrence relations obtained from solutions.
5. What is the recurrence formula for S_n^m?
6. Define recurrence formula for P_n^m.

Long-Answer Questions

1. Explain briefly about the sequence and function of recursion.
2. Describe the concept of recursion and iteration.
3. Discuss the closed form expression and also explain the discrete function in sequence of integers.
4. Elaborate the concept of recurrence relations obtained from solutions.
5. Discuss the recurrence formula for S_n^m and also give the characteristic equation.
6. Illustrate on the algorithm for solving *k*th order homogeneous linear recurrence relation.
7. What is the recurrence for P_n^m and also give the general solution of the recurrence relation?

2.12 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

NOTES

UNIT 3 PATTERNS OF DISTRIBUTIONS

NOTES

Structure

3.0 Introduction
3.1 Objectives
3.2 Theoretical Distributions: Discrete Distribution
  3.2.1 Binomial Distribution
  3.2.2 Poisson Distribution
  3.2.3 Limiting Form of Binomial and Poisson Fitting
3.3 Theoretical Distributions: Continuous Distribution
  3.3.1 Normal Distribution
3.4 Moment Generating Functions and Fitting of Data
3.5 Calculation of Mean and Variance from Moment Generating Functions
3.6 Answers to Check Your Progress Questions
3.7 Summary
3.8 Key Words
3.9 Self Assessment Questions and Exercises
3.10 Further Readings

3.0 INTRODUCTION

In mathematics a random variable is a function that associates a unique numerical value with every outcome of an experiment. The value of the random variable will vary from trial to trial as the experiment is repeated. There are two types of random variables – discrete and continuous. The probability distribution of a discrete random variable is a list of probabilities associated with each of its possible values. It is also sometimes called the probability function or probability mass function. The probability density function of a continuous random variable is a function which can be integrated to obtain the probability that the random variable takes a value in a given interval. Binomial distribution is used in finite sampling problems where each observation is one of two possible outcomes ('successes or 'failure'). Poisson distribution is used for modeling rates of occurrence. The term 'normal distribution' refers to a particular way in which observations will tend to pile up around a particular value rather than be spread evenly across a range of values. It is generally most applicable to continuous data and is intrinsically associated with parametric statistics. Graphically, the normal distribution is best described by a bell-shaped curve. This curve is described in terms of the point at which its height is maximum, i.e., its mean and its width or standard deviation. In this unit, you will also learn about the Moment Generating Function (MGF), which generates the moments for the probability distribution of a random variable.

In this unit, you will study the theoretical distributions: discrete distribution, theoretical distributions: continuous distribution, moment generating functions and fitting of data and calculation of mean and variance from moment generating functions.

3.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Understand the basic concept of probability distribution
- Explain the types of probability distribution
- Describe the binomial distribution based on Bernoulli process
- Describe Poisson distribution
- Analyse Poisson distribution as an approximation of binomial distribution
- Explain the basic theory, characteristics and family of normal distributions
- Measure the area under the normal curve
- Explain the moment generating function
- Calculate mean and variance from moment generating functions

3.2 THEORETICAL DISTRIBUTIONS: DISCRETE DISTRIBUTION

When a random variable x takes discrete values $x_1, x_2,...., x_n$ with probabilities $p_1, p_2,...,p_n$, we have a discrete probability distribution of X.

The function $p(x)$ for which $X = x_1, x_2,..., x_n$ takes values $p_1, p_2,....,p_n$, is the probability function of X.

The variable is discrete because it does not assume all values. Its properties are:

$$p(x_i) = \text{Probability that } X \text{ assumes the value } x.$$
$$= \text{Prob } (x = x_i) = p_i$$
$$p(x) \geq 0, \Sigma p(x) = 1$$

For example, four coins are tossed and the number of heads X noted. X can take value 0, 1, 2, 3, 4 heads.

$$p(X=0) = \left(\frac{1}{2}\right)^4 = \frac{1}{16}$$

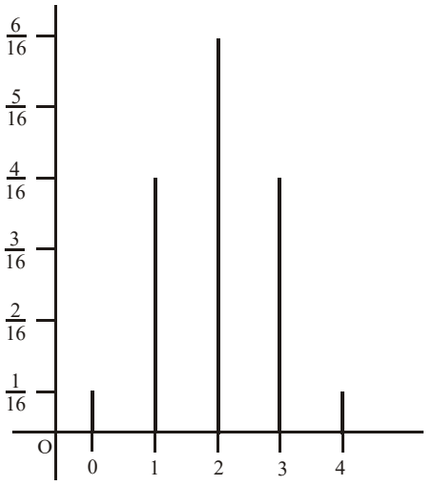
$$p(X=1) = {}^4C_1\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^3 = \frac{4}{16}$$

$$p(X=2) = {}^4C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 = \frac{6}{16}$$

$$p(X=3) = {}^4C_3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right) = \frac{4}{16}$$

$$p(X=4) = {}^4C_4\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^0 = \frac{1}{16}$$

NOTES



$$\sum_{x=2}^{4} p(x) = \frac{1}{16} + \frac{4}{16} + \frac{6}{16} + \frac{4}{16} + \frac{1}{16} = 1$$

This is a discrete probability distribution.

Example 3.1: If a discrete variable X has the following probability function, then, find (*i*) a (*ii*) $p(X \leq 3)$ (*iii*) $p(X \geq 3)$

Solution:

x_1	$p(x_i)$
0	0
1	a
2	$2a$
3	$2a^2$
4	$4a^2$
5	$2a$

Since $\Sigma p(x) = 1$, $0 + a + 2a + 2a^2 + 4a^2 + 2a = 1$

$\therefore$ $6a^2 + 5a - 1 = 0$, so that $(6a - 1)(a + 1) = 0$

$$a = \frac{1}{6} \text{ or } a = -1 \text{ (not admissible)}$$

For $a = \frac{1}{6}$, $p(X \leq 3) = 0 + a + 2a + 2a^2 = 2a^2 + 3a = \frac{5}{9}$

$$p(X \geq 3) = 4a^2 + 2a = \frac{4}{9}$$

Discrete Distributions

There are several discrete distributions. Some other discrete distributions are described below.

3.2.1 Binomial Distribution

Binomial distribution (or the Binomial probability distribution) is a widely used probability distribution concerned with a discrete random variable and as such is

NOTES

an example of a discrete probability distribution. The binomial distribution describes discrete data resulting from what is often called as the Bernoulli process. The tossing of a fair coin a fixed number of times is a Bernoulli process and the outcome of such tosses can be represented by the binomial distribution. The name of Swiss mathematician Jacob Bernoulli is associated with this distribution. This distribution applies in situations where there are repeated trials of any experiment for which only one of two mutually exclusive outcomes (often denoted as 'success' and 'failure') can result on each trial.

Bernoulli Process

Binomial distribution is considered appropriate in a Bernoulli process which has the following characteristics:

(a) ***Dichotomy***. This means that each trial has only two mutually exclusive possible outcomes, e.g., 'Success' or 'failure', 'Yes' or 'No', 'Heads' or 'Tails' and the like.

(b) ***Stability.*** This means that the probability of the outcome of any trial is known (or given) and remains *fixed* over time, i.e., remains the same for all the trials.

(c) ***Independence.*** This means that the trials are statistically independent, i.e., to say the happening of an outcome or the event in any particular trial is independent of its happening in any other trial or trials.

Probability Function of Binomial Distribution

The random variable, say X, in the Binomial distribution is the number of 'successes' in n trials. The probability function of the binomial distribution is written as under:

$$f(X = r) = {}^{n}C_{r} p^{r} q^{n-r}$$

$$r = 0, 1, 2 \ldots n$$

Where, n = Numbers of trials.

p = Probability of success in a single trial.

$q = (1 - p)$ = Probability of 'failure' in a single trial.

r = Number of successes in 'n' trials.

Parameters of Binomial Distribution

This distribution depends upon the values of p and n which in fact are its parameters. Knowledge of p truly defines the probability of X since n is known by definition of the problem. The probability of the happening of exactly r events in n trials can be found out using the above stated binomial function.

The value of p also determines the general appearance of the binomial distribution, if shown graphically. In this context the usual generalizations are:

(a) When p is small (say 0.1), the binomial distribution is skewed to the right, i.e., the graph takes the form shown in Figure 3.1.

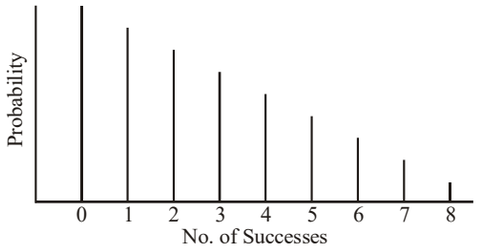


Fig. 3.1 *p is small*

(b) When p is equal to 0.5, the binomial distribution is symmetrical and the graph takes the form as shown in Figure 3.2.

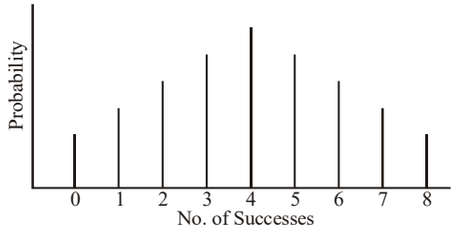


Fig. 3.2 *p = 0.5*

(c) When p is larger than 0.5, the binomial distribution is skewed to the left and the graph takes the form as shown in Figure 3.3.

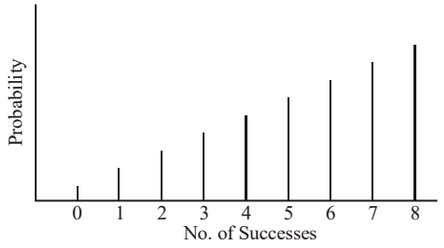


Fig. 3.3 *p > 0.5*

But if 'p' stays constant and 'n' increases, then as 'n' increases the vertical lines become not only numerous but also tend to bunch up together to form a bell shape, i.e., the binomial distribution tends to become symmetrical and the graph takes the shape as shown in Figure 3.4.

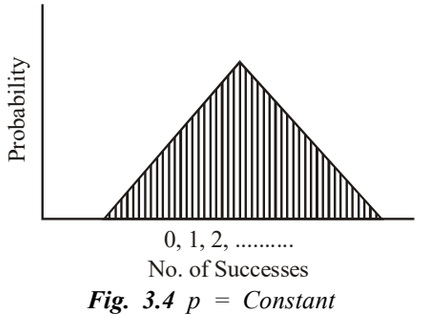


Fig. 3.4 *p = Constant*

Important Measures of Binomial Distribution

The expected value of random variable [i.e., $E(X)$] or mean of random variable (i.e., $\bar{X}$) of the binomial distribution is equal to $n.p$ and the variance of random variable is equal to $n.\ p.\ q$ or $n.\ p.\ (1 - p)$. Accordingly the standard deviation

NOTES

of binomial distribution is equal to $\sqrt{n.p.q}$. The other important measures relating to binomial distribution are as under:

$$\text{Skewness} = \frac{1-2p}{\sqrt{n.p.q}}$$

$$\text{Kurtosis} = 3 + \frac{1-6p+6q^2}{n.p.q}$$

When to Use Binomial Distribution

The use of binomial distribution is most appropriate in situations fulfilling the conditions outlined above. Two such situations, for example, can be described as follows.

(a) When we have to find the probability of 6 heads in 10 throws of a fair coin.

(b) When we have to find the probability that 3 out of 10 items produced by a machine, which produces 8 per cent defective items on an average, will be defective.

Example 3.2: A fair coin is thrown 10 times. The random variable X is the number of head(s) coming upwards. Using the binomial probability function, find the probabilities of all possible values which X can take and then verify that binomial distribution has a mean: $\overline{X} = n.p.$ and variance: $\sigma^2 = n.p.q$

Solution: Since the coin is fair and so, when thrown, can come either with head upward or tail upward. Hence, p (head) $= \frac{1}{2}$ and q (no head) $= \frac{1}{2}$. The required probability function is,

$$f(X = r) = {}^{n}C_r\, p^r q^{n-r}$$

$$r = 0, 1, 2 \ldots 10$$

The following table of binomial probability distribution is constructed using this function.

X_i (*Number of Heads*)	*Probability* pr_i	$X_i\, pr_i$	$(X_i - \overline{X})$	$(X_i - \overline{X})^2$	$(X_i - \overline{X})^2.p_i$
0	${}^{10}C_0\, p^0\, q^{10} = 1/1024$	0/1024	–5	25	25/1024
1	${}^{10}C_1\, p^1\, q^9 = 10/1024$	10/1024	–4	16	160/1024
2	${}^{10}C_2\, p^2\, q^8 = 45/1024$	90/1024	–3	9	405/1024
3	${}^{10}C_3\, p^3\, q^7 = 120/1024$	360/1024	–2	4	480/1024
4	${}^{10}C_4\, p^4\, q^6 = 210/1024$	840/1024	–1	1	210/1024
5	${}^{10}C_5\, p^5\, q^5 = 252/1024$	1260/1024	0	0	0/1024
6	${}^{10}C_6\, p^6\, q^4 = 210/1024$	1260/1024	1	1	210/1024
7	${}^{10}C_7\, p^7\, q^3 = 120/1024$	840/1024	2	4	480/1024
8	${}^{10}C_8\, p^8\, q^2 = 45/1024$	360/1024	3	9	405/1024
9	${}^{10}C_9\, p^9\, q^1 = 10/1024$	90/1024	4	16	160/1024
10	${}^{10}C_{10}\, p^{10}\, q^0 = 1/1024$	10/1024	5	25	25/1024
		$\Sigma\overline{X} = 5120/1024$			Variance = $\sigma^2 =$
		$\overline{X} = 5$			$\Sigma(X_i - \overline{X})^2.pr_i =$
					2560/1024 = 2.5

The mean of the binomial distribution[1] is given by $n.\ p. = 10 \times \frac{1}{2} = 5$ and the variance of this distribution is equal to $n.\ p.\ q. = 10 \times \frac{1}{2} \times \frac{1}{2} = 2.5$

These values are exactly the same as we have found them in the above table.

Hence, these values stand verified with the calculated values of the two measures as shown in the table.

NOTES

3.2.2 Poisson Distribution

Poisson distribution is also a discrete probability distribution with which is associated the name of a Frenchman, Simeon Denis Poisson who developed this distribution. This distribution is frequently used in context of Operations Research and for this reason has a great significance for management people. This distribution plays an important role in *Queuing* theory, Inventory control problems and also in Risk models.

Unlike binomial distribution, Poisson distribution cannot be deducted on purely theoretical grounds based on the conditions of the experiment. In fact, it must be based on experience, i.e., on the empirical results of past experiments relating to the problem under study. Poisson distribution is appropriate specially when probability of happening of an event is very small (so that q or $(1-p)$ is almost equal to unity) and n is very large such that the average of series (viz., $n.\ p.$) is a finite number. Experience has shown that this distribution is good for calculating the probabilities associated with X occurrences in a given time period or specified area.

The random variable of interest in Poisson distribution is number of occurrences of a given event during a given interval (interval may be time, distance, area, etc.). We use capital X to represent the discrete random variable and lower case x to represent a specific value that capital X can take. The probability function of this distribution is generally written as under:

$$f(X_i = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$x = 0,\ 1,\ 2\ldots$$

Where, λ = Average number of occurrences per specified interval.[2] In other words, it is the mean of the distribution.

e = 2.7183 being the basis of natural logarithms.

x = Number of occurrences of a given event.

Poisson Process

The distribution applies in case of Poisson process which has following characteristics.

- Concerning a given random variable, the mean relating to a given interval can be estimated on the basis of past data concerning the variable under study.

NOTES

- If we divide the given interval into very very small intervals we will find:
  (a) The probability that exactly one event will happen during the very very small interval is a very small number and is constant for every other very small interval.
  (b) The probability that two or more events will happen within a very small interval is so small that we can assign it a zero value.
  (c) The event that happens in a given very small interval is independent, when the very small interval falls during a given interval.
  (d) The number of events in any small interval is not dependent on the number of events in any other small interval.

Parameter and Important Measures of Poisson Distribution

Poisson distribution depends upon the value of λ, the average number of occurrences per specified interval which is its only parameter. The probability of exactly x occurrences can be found out using Poisson probability function stated above.[3] The expected value or the mean of Poisson random variable is λ and its variance is also λ.[4] The standard deviation of Poisson distribution is, $\sqrt{\lambda}$.

Underlying the Poisson model is the assumption that if there are on the average λ occurrences per interval t, then there are on the average k λ occurrences per interval kt. For example, if the number of arrivals at a service counted in a given hour, has a Poisson distribution with λ = 4, then y, the number of arrivals at a service counter in a given 6 hour day, has the Poisson distribution λ = 24, i.e., 6×4.

When to Use Poisson Distribution

The use of Poisson distribution is resorted to those cases when we do not know the value of 'n' or when 'n' can not be estimated with any degree of accuracy. In fact, in certain cases it does not make any sense in asking the value of 'n'. For example, the goals scored by one team in a football match are given, it cannot be stated how many goals could not be scored. Similarly, if one watches carefully one may find out how many times the lightning flashed but it is not possible to state how many times it did not flash. It is in such cases we use Poisson distribution. The number of death per day in a district in one year due to a disease, the number of scooters passing through a road per minute during a certain part of the day for a few months, the number of printing mistakes per page in a book containing many pages, are a few other examples where Poisson probability distribution is generally used.

Example 3.3: Suppose that a manufactured product has 2 defects per unit of product inspected. Use Poisson distribution and calculate the probabilities of finding a product without any defect, with 3 defects and with four defects.

Solution: If the product has 2 defects per unit of product inspected. Hence, λ = 2.

Poisson probability function is as follows:

$$f(X_i = x) = -\frac{\lambda^x . e^{-\lambda}}{x!}$$

$$x = 0, 1, 2, \ldots$$

Using the above probability function, we find the required probabilities as under:

$$P(\text{without any defects, i.e., } x = 0) = \frac{2^0 . e^{-2}}{0}$$

$$= \frac{1.(0.13534)}{1} = 0.13534$$

$$P(\text{with 3 defects, i.e., } x = 3) = \frac{2^3 . e^{-2}}{3} = \frac{2 \times 2 \times 2(0.13534)}{3 \times 2 \times 1}$$

$$= \frac{0.54136}{3} = 0.18045$$

$$P(\text{with 4 defects, i.e., } x = 4) = \frac{2^4 . e^{-2}}{4} = \frac{2 \times 2 \times 2 \times 2(0.13534)}{4 \times 3 \times 2 \times 1}$$

$$= \frac{0.27068}{3} = 0.09023$$

Example 3.4: How would you use a Poisson distribution to find approximately the probability of exactly 5 successes in 100 trials the probability of success in each trial being $p = 0.1$?

Solution: In the question we have been given,

$n = 100$ and $p = 0.1$

$\therefore \quad \lambda = n.p = 100 \times 0.1 = 10$

To find the required probability, we can use Poisson probability function as an approximation to Binomial probability function as shown below:

$$f(X_i = x) = \frac{\lambda^x . e^{-\lambda}}{x!} = \frac{(n.p)^x . e^{-(n.p)}}{x!}$$

or $$P(5)^7 = \frac{10^5 . e^{-10}}{5} = \frac{(100000)(0.00005)}{5 \times 4 \times 3 \times 2 \times 1} = \frac{5.00000}{5 \times 4 \times 3 \times 2 \times 1}$$

$$= \frac{1}{24} = 0.042$$

3.2.3 Limiting Form of Binomial and Poisson Fitting

Fitting a Binomial Distribution

When a binomial distribution is to be fitted to the given data, then the following procedure is adopted:

(a) Determine the values of 'p' and 'q' keeping in view that $\bar{X} = n.\ p.$ and $q = (1 - p)$.

(b) Find the probabilities for all possible values of the given random variable applying the binomial probability function, viz.,

$$f(X_i = r) = {}^nC_r p^r q^{n-r}$$
$$r = 0,\ 1,\ 2,\ldots n$$

(c) Work out the expected frequencies for all values of random variable by multiplying N (the total frequency) with the corresponding probability.

(d) The expected frequencies, so calculated, constitute the fitted binomial distribution to the given data.

NOTES

Fitting a Poisson Distribution

When a Poisson distribution is to be fitted to the given data, then the following procedure is adopted:

(a) Determine the value of λ, the mean of the distribution.

(b) Find the probabilities for all possible values of the given random variable using the Poisson probability function, viz.,

$$f\left(X_i = x\right) = \frac{\lambda^x . e^{-\lambda}}{x}$$
$$x = 0,\ 1,\ 2,\ldots$$

(c) Work out the expected frequencies as follows:

$$n.p.(X_i = x)$$

(d) The result of Case (c) above is the fitted Poisson distribution to the given data.

Poisson Distribution as an Approximation of Binomial Distribution

Under certain circumstances Poisson distribution can be considered as a reasonable approximation of Binomial distribution and can be used accordingly. The circumstances which permit all this, are when 'n' is large approaching to infinity and p is small approaching to zero (n = Number of Trials, p = Probability of 'Success'). Statisticians usually take the meaning of large n, for this purpose, when $n \geq 20$ and by small 'p' they mean when $p \leq 0.05$. In cases where these two conditions are fulfilled, we can use mean of the binomial distribution (viz., $n.p.$) in place of the mean of Poisson distribution (viz., λ) so that the probability function of Poisson distribution becomes as stated below:

$$f\left(X_i = x\right) = \frac{\left(n.p\right)^x . e^{-(np)}}{x}$$

We can explain Poisson distribution as an approximation of the Binomial distribution with the help of following example.

Example 3.5 Given is the following information:

(a) There are 20 machines in a certain factory, i.e., n = 20.

(b) The probability of machine going out of order during any day is 0.02.

What is the probability that exactly 3 machines will be out of order on the same day? Calculate the required probability using both Binomial and Poissons

Distributions and state whether Poisson distribution is a good approximation of the Binomial distribution in this case.

Solution: Probability, as per *Poisson probability function* (using *n.p* in place of λ)

NOTES

(since $n \geq 20$ and $p \leq 0.05$)

$$f(X_i = x) = \frac{(n.p)^x .e^{-np}}{x!}$$

Where, x means number of machines becoming out of order on the same day.

$$P(X_i = 3) = \frac{(20 \times 0.02)^3 .e^{-(20 \times 0.02)}}{3}$$

$$= \frac{(0.4)^3 .(0.67032)}{3 \times 2 \times 1} = \frac{(0.064)(0.67032)}{6}$$

$$= 0.00715$$

Probability, as per *Binomial probability function,*

$$f(X_i = r) = {}^nC_r \, p^r q^{n-r}$$

Where, $n = 20$, $r = 3$, $p = 0.02$ and hence $q = 0.98$

$$\therefore \quad f(X_i = 3) = {}^{20}C_3 (0.02)^3 \ (0.98)^{17}$$

$$= 0.00650$$

The difference between the probability of 3 machines becoming out of order on the same day calculated using probability function and binomial probability function is just 0.00065. The difference being very very small, we can state that in the given case Poisson distribution appears to be a good approximation of Binomial distribution.

Example 3.6: How would you use a Poisson distribution to find approximately the probability of exactly 5 successes in 100 trials the probability of success in each trial being $p = 0.1$?

Solution: In the question we have been given,

$$n = 100 \text{ and } p = 0.1$$

$$\therefore \quad \lambda = n.p = 100 \times 0.1 = 10$$

To find the required probability, we can use Poisson probability function as an approximation to Binomial probability function, as shown below:

$$f(X_i = x) = \frac{\lambda^x .e^{-\lambda}}{x!} = \frac{(n.p)^x .e^{-(n.p)}}{x!}$$

or, $$P(5)^7 = \frac{10^5 .e^{-10}}{5} = \frac{(100000)(0.00005)}{5 \times 4 \times 3 \times 2 \times 1} = \frac{5.00000}{5 \times 4 \times 3 \times 2 \times 1}$$

$$= \frac{1}{24} = 0.042$$

NOTES

Check Your Progress

1. Explain the Bernoulli process.
2. Write the probability function of binomial distribution.
3. Explain the different parameters of binomial distributions.
4. Explain the important measures of binomial distribution.
5. Under what circumstances will you use binomial distribution?
6. What is Poisson distribution?
7. Where and when will you use Poisson distribution?
8. Under what circumstances, is a Poisson distribution considered as an approximation of binomial distribution?

3.3 THEORETICAL DISTRIBUTIONS: CONTINUOUS DISTRIBUTION

When a random variate can take any value in the given interval $a \le x \le b$, it is a continuous variate and its distribution is a *Continuous Probability Distribution.*

Theoretical distributions are often continuous. They are useful in practice because they are convenient to handle mathematically. They can serve as good approximations to discrete distributions.

The range of the variate may be finite or infinite.

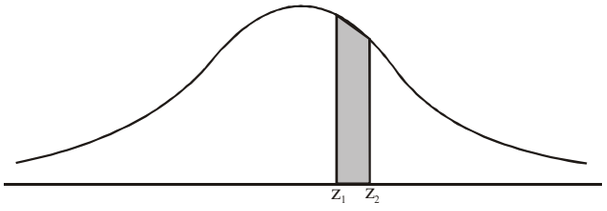
A continuous random variable can take all values in a given interval.

A continuous probability distribution is represented by a smooth curve.

The total area under the curve for a probability distribution is necessarily unity. The curve is always above the x axis because the area under the curve for any interval represents probability and probabilities cannot be negative.

If X is a continous variable, the probability of X falling in an interval with end points z_1, z_2 may be written $p(z_1 \le X \le z_2)$.

This probability corresponds to the shaded area under the curve.



A function is a probability density function if,

$\int_{-\infty}^{\infty} p(x)dx = 1, p(x) \geq 0, -\infty < x < \infty,$ i.e., the area under the curve $p(x)$ is 1 and the probability of x lying between two values a, b, i.e., $p(a < x < b)$ is positive. The most prominent example of a continuous probability function is the normal distribution.

NOTES

Cumulative Probability Function (CPF)

The Cumulative Probability Function (CPF) shows the probability that x takes a value less than or equal to, say, z and corresponds to the area under the curve up to z:

$$p(x \leq z) = \int_{-\infty}^{z} p(x)dx$$

This is denoted by $F(x)$.

3.3.1 Normal Distribution

Among all the probability distributions the normal probability distribution is by far the most important and frequently used continuous probability distribution. This is so because this distribution well fits in many types of problems. This distribution is of special significance in inferential statistics since it describes probabilistically the link between a statistic and a parameter (i.e., between the sample results and the population from which the sample is drawn). The name of Karl Gauss, eighteenth century mathematician-astronomer, is associated with this distribution and in honour of his contribution, this distribution is often known as the Gaussian distribution.

The normal distribution can be theoretically derived as the limiting form of many discrete distributions. For instance, if in the binomial expansion of $(p + q)^n$, the value of 'n' is infinity and $p = q = \frac{1}{2}$, then a perfectly smooth symmetrical curve would be obtained. Even if the values of p and q are not equal but if the value of the exponent 'n' happens to be very very large, we get a curve of normal probability smooth and symmetrical. Such curves are called normal probability curves (or at times known as normal curves of error) and such curves represent the normal distributions.[5]

The probability function in case of normal probability distribution[6] is given as:

$$f(x) = \frac{1}{\sigma.\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Where, μ = The mean of the distribution.

σ^2 = Variance of the distribution.

The normal distribution is thus defined by two parameters, viz., μ and σ^2. This distribution can be represented graphically as shown in Figure 3.5.

NOTES

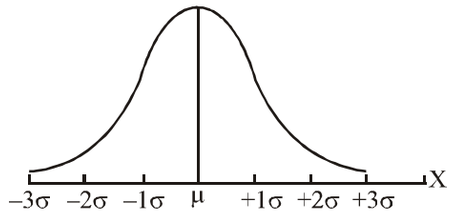
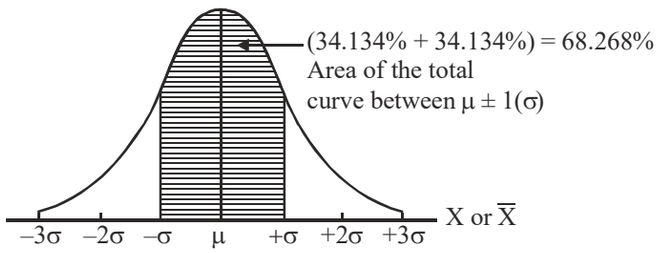


Fig. 3.5 *Curve Representing Normal Distribution*

Characteristics of Normal Distribution

The characteristics of the normal distribution or that of normal curve are, as given below:

1. It is symmetric distribution.[7]
2. The mean μ defines where the peak of the curve occurs. In other words, the ordinate at the mean is the highest ordinate. The height of the ordinate at a distance of one standard deviation from mean is 60.653% of the height of the mean ordinate and similarly the height of other ordinates at various standard deviations (σ_s) from mean happens to be a fixed relationship with the height of the mean ordinate.
3. The curve is asymptotic to the base line which means that it continues to approach but never touches the horizontal axis.
4. The variance (σ^2) defines the spread of the curve.
5. Area enclosed between mean ordinate and an ordinate at a distance of one standard deviation from the mean is always 34.134% of the total area of the curve. It means that the area enclosed between two ordinates at one sigma (S.D.) distance from the mean on either side would always be 68.268% of the total area. This can be shown as follows:



Similarly, the other area relationships are as follows:

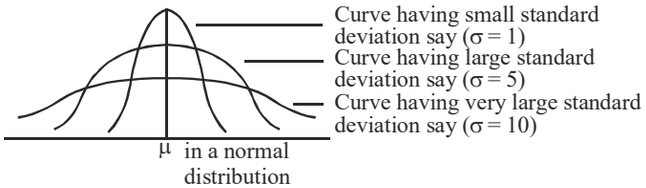
Between		*Area Covered to Total Area of the Normal Curve*[8]
μ ± 1	S.D.	68.27%
μ ± 2	S.D.	95.45%
μ ± 3	S.D.	99.73%
μ ± 1.96	S.D.	95%
μ ± 2.578	S.D.	99%
μ ± 0.6745	S.D.	50%

6. The normal distribution has only one mode since the curve has a single peak. In other words, it is always a unimodal distribution.
7. The maximum ordinate divides the graph of normal curve into two equal parts.
8. In addition to all the above stated characteristics the curve has the following properties:
   (a) $\mu = \bar{x}$ (b) $\mu_2 = \sigma^2 =$ Variance
   (c) $\mu_4 = 3\sigma^4$ (d) Moment Coefficient of Kurtosis = 3

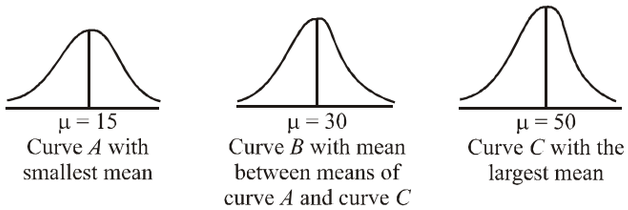
Family of Normal Distributions

We can have several normal probability distributions but each particular normal distribution is being defined by its two parameters viz., the mean (μ) and the standard deviation (σ). There is, thus, not a single normal curve but rather a family of normal curves. We can exhibit some of these as under:

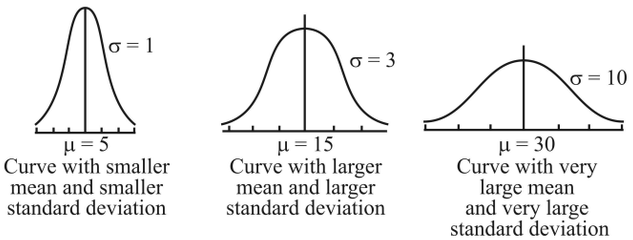
Normal curves with identical means but different standard deviations:



Normal curves with identical standard deviation but each with different means:



Normal curves each with different standard deviations and different means:



How to Measure the Area under the Normal Curve?

We have stated above some of the area relationships involving certain intervals of standard deviations (plus and minus) from the means that are true in case of a normal curve. But what should be done in all other cases? We can make use of the statistical tables constructed by mathematicians for the purpose. Using these tables we can find the area (or probability, taking the entire area of the curve as equal to 1) that the normally distributed random variable will lie within certain distances from the mean. These distances are defined in terms of standard deviations. While

NOTES

using the tables showing the area under the normal curve we talk in terms of standard variate (symbolically Z) which really means standard deviations without units of measurement and this 'Z' is worked out as under:

$$Z = \frac{X - \mu}{\sigma}$$

Where, Z = The standard variate (or number of standard deviations from X to the mean of the distribution).

X = Value of the random variable under consideration.

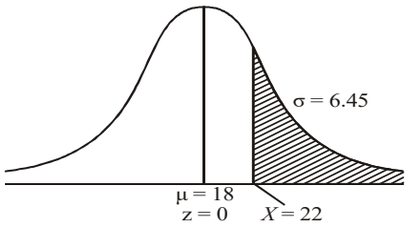
μ = Mean of the distribution of the random variable.

σ = Standard deviation of the distribution.

The table showing the area under the normal curve (often termed as the standard normal probability distribution table) is organized in terms of standard variate (or Z) values. It gives the values for only half the area under the normal curve, beginning with $Z = 0$ at the mean. Since the normal distribution is perfectly symmetrical the values true for one half of the curve are also true for the other half. We now illustrate the use of such a table for working out certain problems.

Example 3.7: A banker claims that the life of a regular saving account opened with his bank averages 18 months with a standard deviation of 6.45 months. Answer the following: (a) What is the probability that there will still be money in 22 months in a savings account opened with the said bank by a depositor? (b) What is the probability that the account will have been closed before two years?

Solution: (a) For finding the required probability we are interested in the area of the portion of the normal curve as shaded and shown below:

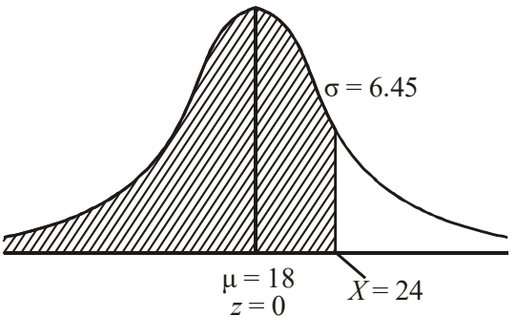


Let us calculate Z as under:

$$Z = \frac{X - \mu}{\sigma} = \frac{22 - 18}{6.45} = 0.62$$

The value from the table showing the area under the normal curve for $Z = 0.62$ is 0.2324. This means that the area of the curve between $\mu = 18$ and $X = 22$ is 0.2324. Hence, the area of the shaded portion of the curve is (0.5) – (0.2324) = 0.2676 since the area of the entire right hand portion of the curve always happens to be 0.5. Thus the probability that there will still be money in 22 months in a savings account is 0.2676.

(b) For finding the required probability we are interested in the area of the portionof the normal curve as shaded and shown in figure:



For the purpose we calculate,

$$Z = \frac{24 - 18}{6.45} = 0.93$$

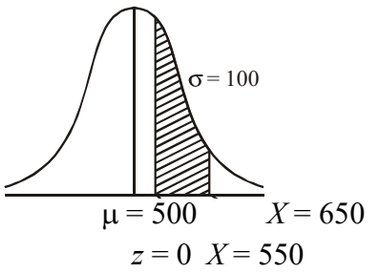
The value from the concerning table, when $Z = 0.93$, is 0.3238 which refers to the area of the curve between $\mu = 18$ and $X = 24$. The area of the entire left hand portion of the curve is 0.5 as usual.

Hence, the area of the shaded portion is (0.5) + (0.3238) = 0.8238 which is the required probability that the account will have been closed before two years, i.e., before 24 months.

Example 3.8: Regarding a certain normal distribution concerning the income of the individuals we are given that mean=500 rupees and standard deviation =100 rupees. Find the probability that an individual selected at random will belong to income group,

(a) ₹ 550 to ₹ 650 (b) ₹ 420 to 570

Solution: (a) For finding the required probability we are interested in the area of the portion of the normal curve as shaded and shown below:



For finding the area of the curve between $X = 550$ to 650, let us do the following calculations:

$$Z = \frac{550 - 500}{100} = \frac{50}{100} = 0.50$$

Corresponding to which the area between $\mu = 500$ and $X = 550$ in the curve as per table is equal to 0.1915 and,

$$Z = \frac{650 - 500}{100} = \frac{150}{100} = 1.5$$

NOTES

Corresponding to which, the area between $\mu = 500$ and $X = 650$ in the curve, as per table, is equal to 0.4332.

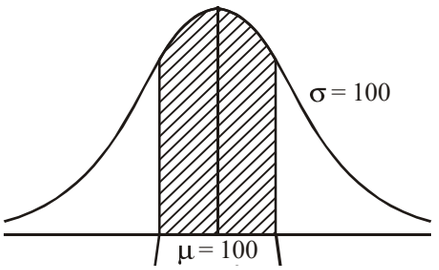
Hence, the area of the curve that lies between $X = 550$ and $X = 650$ is,

$$(0.4332) - (0.1915) = 0.2417$$

This is the required probability that an individual selected at random will belong to income group of ₹ 550 to ₹ 650.

(b) For finding the required probability we are interested in the area of the portion of the normal curve as shaded and shown below:

To find the area of the shaded portion we make the following calculations:



$$Z = \frac{570 - 500}{100} = 0.70$$

Corresponding to which the area between $\mu = 500$ and $X = 570$ in the curve as per table is equal to 0.2580.

and $$Z = \frac{420 - 500}{100} = -0.80$$

Corresponding to which the area between $\mu = 500$ and $X = 420$ in the curve as per table is equal to 0.2881.

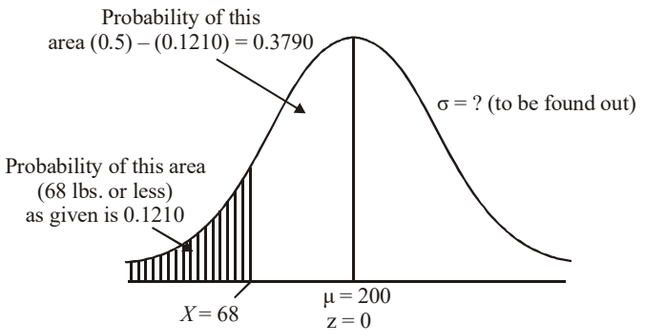
Hence, the required area in the curve between $X = 420$ and $X = 570$ is,

$$(0.2580) + (0.2881) = 0.5461$$

This is the required probability that an individual selected at random will belong to income group of ₹ 420 to ₹ 570.

Example 3.9: A certain company manufactures $1\frac{1''}{2}$ all-purpose rope made from imported hemp. The manager of the company knows that the average load-bearing capacity of the rope is 200 lbs. Assuming that normal distribution applies, find the standard deviation of load-bearing capacity for the $1\frac{1''}{2}$ rope if it is given that the rope has a 0.1210 probability of breaking with 68 lbs. or less pull.

Solution: Given information can be depicted in a normal curve as shown below:



If the probability of the area falling within μ = 200 and X = 68 is 0.3790 as stated above, the corresponding value of Z as per the table[9] showing the area of the normal curve is – 1.17 (minus sign indicates that we are in the left portion of the curve)

Now to find σ, we can write,

$$Z = \frac{X - \mu}{\sigma}$$

or $$-1.17 = \frac{68 - 200}{\sigma}$$

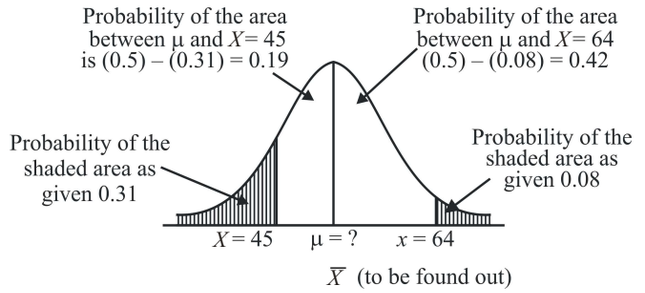
or $$-1.17\sigma = -132$$

or $$\sigma = 112.8 \text{ lbs. approx.}$$

Thus, the required standard deviation is 112.8 lbs. approximately.

Example 3.10: In a normal distribution, 31 per cent items are below 45 and 8 per cent are above 64. Find the $\bar{X}$ and σ of this distribution.

Solution: We can depict the given information in a normal curve as shown below:



If the probability of the area falling within μ and X = 45 is 0.19 as stated above, the corresponding value of Z from the table showing the area of the normal curve is – 0.50. Since, we are in the left portion of the curve, we can express this as under,

$$-0.50 = \frac{45 - \mu}{\sigma} \quad (1)$$

Similarly, if the probability of the area falling within μ and X = 64 is 0.42, as stated above, the corresponding value of Z from the area table is, +1.41. Since, we are in the right portion of the curve we can express this as under,

$$1.41 = \frac{64 - \mu}{\sigma} \quad (2)$$

NOTES

If we solve Equations (1) and (2) above to obtain the value of μ or $\bar{X}$, we have,

$$-0.5\,\sigma = 45 - \mu \quad (3)$$

$$1.41\,\sigma = 64 - \mu \quad (4)$$

NOTES

By subtracting the Equation (4) from Equation (3) we have,

$$-1.91\,\sigma = -19$$

$$\therefore \quad \sigma = 10$$

Putting $\sigma = 10$ in Equation (3) we have,

$$-5 = 45 - \mu$$

$$\therefore \quad \mu = 50$$

Hence, $\bar{X}$ (or μ)=50 and σ =10 for the concerning normal distribution.

3.4 MOMENT GENERATING FUNCTIONS AND FITTING OF DATA

A function that generates moments of a random variable is known as the Moment Generating Function (MGF) of the random variable. A moment generating function may or may not exist but if it exists it is unique.

For the random variable X, its moment generating function $M_X(t)$ or $\psi(t)$ is defined as follows:

$\psi(t)$ or $M_X(t) = E(e^{tx})$

$$= \begin{cases} \sum e^{tx_i} f_i \text{ when } X \text{ is discrete and } f_i = P(X = x_i) \\ \int_{-\infty}^{\infty} e^{tx} f(x)\,dx \text{ when } X \text{ is continuous and } f(x) \text{ is } p.d.f. \text{ of } X. \end{cases}$$

Note that $E(e^{tX})$ is a function of t. The rth raw moment of X is the coefficient of $\frac{t^r}{r!}$ in the power series expansion of $M_X(t)$. That is, if

$$M_X(t) = a_0 + a_1 t + a_2 t^2 + \ldots \text{ to } \infty,$$

Then, $(r!)a_2 = \mu'_r$.

Hence, $$\mu'_r = \left[\frac{d^r}{dt^r} M(t)\right]_{t=0} = M_X^{(r)}(0)$$

Notes: (*i*) We have $M_X(t) = E(e^{tX})$

Differentiating $M_X(t)$ with respect to t, k times, we get

$$M_X^k(t) = E(X^k e^{tX})$$

$\therefore M_X^k(0) = E(X^k) = \mu'_k$ or α_k, provided the moment exists. Thus, α_k is obtained by differentiating $M_X(t)$ k times and putting $t = 0$.

(*ii*) We assume that $M_X(t)$ or $\psi(t)$ can be expanded as a power series in t by Maclaurin's theorem. Then we have,

$$\psi(t) = \psi(0) + t\,\psi'(0) + \frac{t^2}{\underline{|2}}\psi''(0) + \ldots + \frac{t^k}{\underline{|k}}\psi^k(0) + \ldots$$

$$= \sum_{k=0}^{\infty} \frac{t^k}{\lfloor k} \psi^k(0) \quad \text{where } \psi^0(0) = \psi(0) = E(e^0) = E(1) = 1$$

$$= \sum_{k=0}^{\infty} \frac{t^k}{\lfloor k} \alpha_k$$

Thus, it is proved that the kth order moment about origin α_k if it exists, is equal to the coefficient of $\frac{t^k}{\lfloor k}$ in the expansion of $\psi(t)$ as a power series in t for any positive integer k.

This is why $\psi(t)$ or $M_X(t)$ is called moment generating function.

(*iii*) If a and b are constants, then

(*a*) $M_{X+a}(t) = E(e^{(X+a)t}) = e^{at}\, E(e^{Xt}) = e^{at}\, M_X(t)$

(*b*) $M_{aX}(t) = E(e^{aXt}) = M_X(at)$

(*c*) $M_{\frac{X+a}{b}}(t) = E\left(e^{\left(\frac{X+a}{b}\right)t}\right) = e^{\frac{at}{b}}\, E\left(e^{\frac{Xt}{b}}\right) = e^{\frac{at}{b}}\, M_X\left(\frac{t}{b}\right)$

Example 3.11: Find the moment generating function of the random variable X whose probability function is,

$$f(x) = \begin{cases} \frac{1}{2} & \text{if } x = -1 \\ \frac{1}{2} & \text{if } x = 1 \\ 0 & \text{elsewhere} \end{cases}$$

Hence, obtain the first three raw and central moments of X.

Solution: We have

$$M_X(t) = E(e^{tX}) = \frac{1}{2} e^{-t} + \frac{1}{2} e^{t} = \frac{1}{2}(e^{t} + e^{-t}) = \cosh t$$

Now, for the raw moments, we expand $M_X(t)$ in an infinite power series as follows:

$$M_X(t) = 1 + \frac{t^2}{2!} + \frac{t^4}{4!} + \ldots \text{ to } \infty$$

Hence, $\mu'_1 = 0,\ \mu'_2 = 1,\ \mu'_3 = 0,\ \mu'_4 = 1.$

[Note $M'_X(0) = 0,\ M''_X(0) = 1,\ M^{iii}_X(0) = 0,\ M^{iv}_X(0) = 1$]

$\therefore$ $\mu_1 = 0,\ \mu_2 = \mu'_2 - \mu'^2_1 = 1 - 0^2 = 1,\ \mu_3 = 0,\ \mu_4 = 6$

Example 3.12: Find the moment generating function of the random variable whose probability function is,

$$f(x) = \begin{cases} \frac{1}{2} & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

Find the central moments also.

Solution: We see

$$M_X(t) = E(e^{tx})$$

NOTES

$$= \frac{1}{2}\int_0^2 e^{tx}dx = \frac{1}{2}\left[\frac{e^{tx}}{t}\right]_0^2 = \frac{1}{2t}\left(e^{2t}-1\right)$$

NOTES

Writing in a power series, we get $M_X(t) = 1 + \frac{2t}{2!} + \frac{4t^2}{3!} + \frac{8t^3}{4!} + \ldots \text{ to } \infty$

or, $$M_X(t) = 1 + \frac{t}{1!} + \frac{4}{3}\frac{t^2}{2!} + \frac{2t^3}{3!} + \ldots \text{ to } \infty$$

Now, $$\mu'_1 = 1$$

$$\mu'_2 = \frac{4}{3}$$

$$\mu'_3 = 2$$

$$\mu'_4 = \frac{16}{5}$$

$\therefore$ $$\mu_1 = 0,$$

$$\mu_2 = \frac{4}{3} - 1^2 = \frac{1}{3},$$

$$\mu_3 = 2 - 3\left(\frac{4}{3}\right)(1) + 2\,(1)^3 = 0$$

$$\mu_4 = \frac{16}{5} - 4\,(2)\,(1) + 6\left(\frac{4}{3}\right)(1)^2 - 3(1)^4 = \frac{1}{5}$$

Definition: The characteristic function $\phi_X(t)$ of a random variable X is defined as:

$$\phi_X(t) = E(e^{itX})$$

Note that the characteristic function of a random variable always exists.

A characteristic function of a random variable also gives the moments of the random variable In fact, the coefficient of $\frac{(it)^r}{r!}$, is the rth raw moment.

Example 3.13: Obtain the characteristic function of the Laplace variate whose probability function is:

$$f(x) = \frac{1}{2\lambda}e^{-\frac{|x-\mu|}{\lambda}}, -\infty < x < \infty \;\; \lambda > 0.$$

Solution: We have

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx}\,\frac{1}{2\lambda}e^{-\frac{|x-\mu|}{x}}\,dx$$

$$= \frac{1}{2\lambda}\int_{-\infty}^{\mu} e^{itx+\frac{(x-\mu)}{\lambda}}\,dx + \frac{1}{2\lambda}\int_{\mu}^{\infty} e^{itx-\frac{x-\mu}{\lambda}}\,dx$$

$$= \frac{1}{2\lambda}e^{-\frac{\mu}{\lambda}}\int_{-\infty}^{\mu} e^{\left(it-\frac{1}{\lambda}\right)x}\,dx + \frac{1}{2\lambda}e^{+\frac{\mu}{\lambda}}\int_{\mu}^{\infty} e^{\left(it+\frac{1}{\lambda}\right)x}\,dx$$

$$= \frac{1}{2\lambda}e^{-\frac{\mu}{\lambda}}\left[\frac{e^{\left(it-\frac{1}{\lambda}\right)x}}{it-\frac{1}{\lambda}}\right]_{-\infty}^{\mu} + \frac{1}{2\lambda}e^{+\frac{\mu}{\lambda}}\left[\frac{e^{\left(it+\frac{1}{\lambda}\right)x}}{it+\frac{1}{\lambda}}\right]_{\mu}^{\infty}$$

NOTES

$$= \frac{1}{2\lambda} e^{-\frac{\mu}{\lambda}} \frac{e^{\left(it-\frac{1}{\lambda}\right)\mu}}{it-\frac{1}{\lambda}} - \frac{1}{2\lambda} e^{+\frac{\mu}{\lambda}} \frac{e^{\left(it+\frac{1}{\lambda}\right)\mu}}{it+\frac{1}{\lambda}}$$

$$= \frac{1}{2}\left(\frac{e^{it\mu}}{1+it\lambda} - \frac{e^{it\mu}}{it\lambda - 1}\right)$$

$$= \frac{e^{it\mu}}{1+t^2\lambda^2}$$

Probability Differential: Let X be a continuous random variable. If $\delta x > 0$, then $P(x < X \le x + \delta x) = F(x + \delta x) - F(x)$.

$= \delta x\, F'(x + \theta\, \delta x)$ where $0 < \theta < 1$

[by Lagrange mean value theorem of differential calculus]

So, $\dfrac{\mathrm{p}(x < X \le x + \delta x)}{\delta x} = F'(x + \theta\, \delta x)$

$\therefore \lim\limits_{\delta x \to 0} \dfrac{p(x < X \le x + \delta x)}{\delta x} = \lim\limits_{\delta x \to 0} F'(x + \theta\, \delta x) = F'(x)$ if x be a point of continuity of $F'(x) = f(x)$ where $f(x)$ is the probability density function of X.

Thus, we get

$$f(x) = \lim_{\delta x \to 0} \frac{\mathrm{p}(x < X \le x + \delta x)}{\delta x}$$

$$= \lim_{dx \to 0} \frac{p(x < X \le x + dx)}{dx}$$

[$\because \delta x$ = differential of $x = dx$ for the independent variable x]

Henceforth, we shall write $f(x)\, dx$ for $P(x < X \le x + dx)$ which will actually mean $\lim\limits_{dx \to 0} \dfrac{p(x < X \le x + dx)}{dx} = f(x)$.

The expression $P(x < X \le x + dx)$ will always be used in the above limiting sense and so there will be no ambiguity throughout our discussion. The expression $P(x < X \le x + dx)$ which is taken to be equal to $f(x)\, dx$ is called the *probability differential* for the continuous random variable X.

3.5 CALCULATION OF MEAN AND VARIANCE FROM MOMENT GENERATING FUNCTIONS

1. **Binomial Distribution:** Let X be a binomial (n, p) variate, then the M.G.F. of X is given by:

$$\psi(t) = (q + pe^t)^n \text{ where } q = 1 - p$$

$$\therefore \quad \psi'(t) = n(q + pe^t)^{n-1}\, pe^t$$

$$\psi''(t) = n(n-1)\,(q + pe^t)^{n-2}\,(pe^t)^2 + np\,(q + pe^t)^{n-1}\, e^t$$

NOTES

$$\therefore \quad \psi'(0) = n(q+p)^{n-1}p = np = E(X) = m[\because \ (p+q) = 1]$$

and $$\alpha_2 = \psi''(0) = n(n-1)(q+p)^{n-2}p^2 + np$$
$$= n(n-1)p^2 + np$$

$$\therefore \quad \text{var}(X) = \alpha_2 - m^2 = n(n-1)p^2 + np - n^2p^2$$
$$= n^2p^2 - np^2 + np - n^2p^2 = np - np^2 = np(1-p)$$
$$= npq$$

2. **Poisson Distribution:** Let X be a Poisson λ variate, then the M.G.F. of X is given by:

$$\psi(t) = e^{\lambda(e^t - 1)}$$

$$\therefore \quad \psi'(t) = e^{\lambda(e^t-1)} \cdot \lambda e^t$$
$$\psi''(t) = (\lambda e^t)^2 e^{\lambda(e^t-1)} + \lambda e^t e^{\lambda(e^t-1)}$$

$$\therefore \quad \psi'(0) = E(X) = m = \lambda$$

and $$\alpha_2 = \psi''(0) = \lambda^2 + \lambda$$

$$\therefore \quad \text{var}(X) = \alpha_2 - m^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

3. **Uniform Distribution:** Let X be a uniform random variable in $[a, b]$; then the M.G.F. of X is given by:

$$\psi(t) = \frac{1}{b-a}\left[\frac{e^{bt} - e^{at}}{t}\right]$$

$$= \frac{1}{t(b-a)}\left[\left(1 + bt + \frac{b^2t^2}{\underline{|2}} + \frac{b^3t^3}{\underline{|3}} + \ldots\right) - \left(1 + at + \frac{a^2t^2}{\underline{|2}} + \frac{a^2t^3}{\underline{|3}} + \ldots\right)\right]$$

$$= \frac{1}{t(b-a)}\left[(b-a)t + (b^2 - a^2)\frac{t^2}{2} + (b^3 - a^3)\frac{t^3}{6} + (b^4 - a^4)\frac{t^4}{24} + \ldots\right]$$

$$= 1 + (b+a)\frac{t}{2} + \frac{b^2 + a^2 + ab}{6}t^2 + \frac{(b^2 + a^2)(b+a)}{24}t^3 + \ldots$$

$$\therefore \quad \psi'(t) = \frac{b+a}{2} + \frac{(b^2 + a^2 + ab)}{3}t + \frac{(b^2 + a^2)(b+a)}{8}t^2 + \ldots.$$

$$\therefore \quad \psi''(t) = \frac{b^2 + a^2 + ab}{3} + \frac{(b^2 + a^2)(b+a)}{4}t + \ldots.$$

$$\therefore \quad \psi'(0) = E(X) = m = \frac{b+a}{2}$$

and $$\alpha_2 = \psi''(0) = \frac{b^2 + a^2 + ab}{3}$$

$$\therefore \quad \text{var}(X) = \alpha_2 - m^2 = \frac{b^2 + a^2 + ab}{3} - \left(\frac{b+a}{2}\right)^2$$
$$= \frac{4b^2 + 4a^2 + 4ab - 3b^2 - 3a^2 - 6ab}{12}$$
$$= \frac{b^2 + a^2 - 2ab}{12} = \frac{(b-a)^2}{12}$$

NOTES

4. **Normal Distribution:** Let X be a normal (m, σ) variate, then the M.G.F. of X is given by:

$$\psi(t) = e^{tm + \frac{1}{2}t^2\sigma^2}$$

$$\therefore \quad \psi'(t) = e^{tm + \frac{1}{2}t^2\sigma^2}(m + 2t\sigma^2)$$

$$\psi''(t) = (m + t\sigma^2)^2 \; e^{tm + \frac{1}{2}t^2\sigma^2} + \sigma^2 \; e^{tm + \frac{1}{2}t^2\sigma^2}$$

$$\therefore \quad \psi'(0) = E(X) = \text{Mean} = m$$

and $\quad \psi''(0) = \alpha_2 = m^2 + \sigma^2$

$$\therefore \quad \text{var}(X) = \alpha_2 - m^2 = m^2 + \sigma^2 - m^2 = \sigma^2$$

5. **Exponential Distribution:** Let X be a random variable of the exponential distribution, then the M.G.F. of X is given by:

$$\psi(t) = \frac{\lambda}{\lambda - t} = \frac{1}{1 - \frac{t}{\lambda}} = \left(1 - \frac{t}{\lambda}\right)^{-1}$$

$$= 1 + \frac{t}{\lambda} + \frac{t^2}{\lambda^2} + \frac{t^3}{\lambda^3} + \ldots.$$

$$\therefore \quad \psi'(t) = \frac{1}{\lambda} + \frac{2t}{\lambda^2} + \frac{3t^2}{\lambda^3} + \ldots$$

$$\psi''(t) = \frac{2}{\lambda^2} + \frac{6t}{\lambda^3} + \ldots$$

$$\therefore \quad \psi'(0) = \frac{1}{\lambda} = m = \text{Mean}$$

and $\quad \psi''(0) = \alpha_2 = \frac{2}{\lambda^2}$

$$\therefore \quad \text{var}(X) = \alpha_2 - m^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Example 3.14: The probability mass function of a random variable X is $f_i = P(X = i) = 2^{-i}$ where $i = 1, 2, 3, \ldots$ Find the M.G.F. of X and hence find the mean and variance of X.

Solution: Now $\psi(t)$ = M.G.F. of $X = E(e^{tx}) = \sum_{i=1}^{\infty} e^{ti} f_i$

$$= \sum_{i=1}^{\infty} e^{ti} 2^{-i} = \sum_{i=1}^{\infty} \left(\frac{e^t}{2}\right)^i = \frac{e^t}{2} \sum_{i=1}^{\infty} \left(\frac{e^t}{2}\right)^{i-1}$$

$$= \frac{e^t}{2} \sum_{k=0}^{\infty} \left(\frac{e^t}{2}\right)^k \qquad \text{where } i - 1 = k$$

$$= \frac{e^t}{2} \; [1 + r + r^2 + r^3 + \ldots] \text{ where } r = \frac{e^t}{2}$$

NOTES

$$= \frac{e^t}{2}\left[\frac{1}{1-r}\right] = \frac{e^t}{2}\left[\frac{1}{1-e^t/2}\right] = \frac{e^t}{2-e^t}$$

$$\therefore \psi'(t) = \frac{e^t(2-e^t) - e^t(-e^t)}{(2-e^t)^2} = \frac{e^t[2-e^t+e^t]}{(2-e^t)^2} = \frac{2e^t}{(2-e^t)^2}$$

$$\psi''(t) = \frac{(2e^t)^2\ 2e^t - 2e^t\,2(2-e^t)\,(-e^t)}{(2-e^t)^4} = \frac{2\,[e^t\ (2-e^t) + 2e^t e^t]}{(2-e^t)^3}$$

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

$\therefore$ $\psi'(0) = m = \text{Mean} = E(X) = 2$

and $\psi''(0) = \alpha_2 = \frac{2[1(2+1)]}{(2-1)^3} = 6$

$\therefore$ $\text{Var}(X) = \alpha_2 - m^2 = 6 - 4 = 2$

Example 3.15: Find the M.G.F. of the following continuous distribution with probability density function:

$$f(x) = \begin{cases} \frac{1}{2}x^2e^{-x} & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

and hence find the mean and variance.

Solution: The M.G.F. of this distribution is given by:

$$\psi(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx = \int_0^{\infty} e^{tx}\,\frac{1}{2}x^2e^{-x}dx$$

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

$$= \int_0^{\infty} e^{-z}\,\frac{z^2}{2(1-t)^2}\,\frac{dz}{(1-t)} \quad \text{where } (1-t)\,x = z,\ dx = \frac{dz}{1-t}$$

$$= \frac{1}{2(1-t)^3}\int_0^{\infty} e^{-z}z^2\,dz$$

$$= \frac{1}{2(1-t)^3}\left[\frac{z^2\ e^{-z}}{-1} - 2ze^{-z} - 2e^{-z}\right]_0^{\infty}$$

$$= \frac{2}{2(1-t)^3} = \frac{1}{(1-t)^3}$$

$\therefore$ $\psi'(t) = \frac{3}{(1-t)^4}$, $\psi''(t) = \frac{12}{(1-t)^5}$

$\therefore$ $\psi'(0) = 3 = E(X) = m = \text{Mean}$

and $\psi'(0) = \alpha_2 = 12$

$\text{var}(X) = \alpha_2 - m^2 = 12 - 9 = 3$

Check Your Progress

9. What is a normal distribution?
10. Explain any four characteristics of normal distribution.
11. How can we measure the area under the curve?
12. What is moment generating function?

NOTES

3.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Bernoulli process or Binomial distribution is considered appropriate and has the following characteristics;
   (a) Dichotomy: This means that each trial has only two mutually exclusive possible outcomes. For example, success or failure, yes or no, heads or tail, etc.
   (b) Stability: This means that the probability of the outcome of any trial is known and remains fixed over time, i.e., remains the same for all the trials.
   (c) Independence: This means that the trials are statistically independent, i.e., to say the happening of an outcome or the event in any particular trial is independent of its happening in any other trial or trials.
2. The probability function of binomial distribution is written as under:

$$f(X=r) = {}^{n}C_{r}\, p^{r} q^{n-r}$$

$$r = 0,1,2,...n$$

Where, n = Numbers of trials.
p = Probability of success in a single trial.
$q = (1-p)$ = Probability of failure in a single trial.
r = Number of successes in n trials.

3. The parameters of binomial distribution are p and n, where p specifies the probability of success in a single trial and n specifies the number of trials.
4. The important measures of binomial distribution are:

$$\text{Skewness} = \frac{1-2p}{\sqrt{n.p.q}}$$

$$\text{Kurtosis} = 3 + \frac{1-6p+6q^{2}}{n.p.q}$$

5. We need to use binomial distribution under the following circumstances:
   (a) When we have to find the probability of heads in 10 throws of a fair coin.
   (b) When we have to find the probability that 3 out of 10 items produced by a machine, which produces 8% defective items on an average, will be defective.

NOTES

6. Poisson distribution is a discrete probability distribution that is frequently used in the context of Operations Research. Unlike binomial distribution, Poisson distribution cannot be deduced on purely theoretical grounds based on the conditions of the experiment. In fact, it must be based on the experience, i.e., on the empirical results of past experiments relating to the problem under study.
7. Poisson distribution is used when probability of happening of an event is very small and n is very large such that the average of series is a finite number. This distribution is good for calculating the probabilities associated with X occurrences in a given time period or specified area.
8. When n is large approaching to infinity and p is small approaching to zero, Poisson distribution is considered as an approximation of binomial distribution.
9. Normal distribution is the most important and frequently used continuous probability distribution among all the probability distributions. This is so because this distribution well fits in many types of problems. This distribution is of special significance in inferential statistics since it describes probabilistically the link between a statistic and a parameter.
10. The characteristic of a normal distribution are:
    (a) It is symmetric distribution.
    (b) The curve is asymptotic to the base line, which means that it continues to approach but never touches the horizontal axis.
    (c) The variance (σ^2) defines the spread of the curve.
    (d) The normal distribution has only one mode since the curve has a single peak. In other words, it is always a unimodal distribution.
11. For measuring the area under a curve, we make use of the statistical tables constructed by mathematicians. Using these tables, we can find the area that the normally distributed random variable will lie within certain distances from the mean. These distances are defined in terms of standard deviations. While using the tables showing the area under the normal curve, it is considered in terms of standard variate, which means standard deviations without units of measurement and it is calculated as:

$$Z = \frac{X - \mu}{\sigma}$$

    Where, Z = The standard variate or number of standard deviations from X to the mean of the distribution.
    X = Value of the random variable under consideration.
    μ = Mean of the distribution of the random variable.
    σ = Standard deviation of the distribution.
12. A function that generates moments of a random variable is known as the moment generating function of the random variable. A moment generating function may or may not exist but if it exists it is unique.

3.7 SUMMARY

NOTES

- When a random variable x takes discrete values $x_1, x_2,...., x_n$ with probabilities $p_1, p_2,..., p_n$, we have a discrete probability distribution of X.
- The function $p(x)$ for which $X = x_1, x_2,..., x_n$ takes values $p_1, p_2,...., p_n$, is the probability function of X.
- Binomial distribution (or the Binomial probability distribution) is a widely used probability distribution concerned with a discrete random variable and as such is an example of a discrete probability distribution. The binomial distribution describes discrete data resulting from what is often called as the Bernoulli process.
- The tossing of a fair coin a fixed number of times is a Bernoulli process and the outcome of such tosses can be represented by the binomial distribution. The name of Swiss mathematician Jacob Bernoulli is associated with this distribution. This distribution applies in situations where there are repeated trials of any experiment for which only one of two mutually exclusive outcomes (often denoted as 'success' and 'failure') can result on each trial.
- The random variable, say X, in the Binomial distribution is the number of 'successes' in n trials. The probability function of the binomial distribution is written as under:

$$f(X = r) = nC_r p^r q^{n-r}$$

$$r = 0, 1, 2 \ldots n$$

- The expected value of random variable [i.e., $E(X)$] or mean of random variable (i.e., $\bar{X}$) of the binomial distribution is equal to $n.p$ and the variance of random variable is equal to $n.\ p.\ q$ or $n.\ p.\ (1 - p)$. Accordingly the standard deviation of binomial distribution is equal to $\sqrt{n.p.q}$.
- Poisson distribution is also a discrete probability distribution with which is associated the name of a Frenchman, Simeon Denis Poisson who developed this distribution. This distribution plays an important role in Queuing theory, Inventory control problems and also in Risk models.
- Poisson distribution must be based on experience, i.e., on the empirical results of past experiments relating to the problem under study. Poisson distribution is appropriate specially when probability of happening of an event is very small (so that q or $(1-p)$ is almost equal to unity) and n is very large such that the average of series (viz., $n.\ p.$) is a finite number.
- The random variable of interest in Poisson distribution is number of occurrences of a given event during a given interval (interval may be time, distance, area, etc.). We use capital X to represent the discrete random variable and lower case x to represent a specific value that capital X can take.
- Poisson distribution depends upon the value of λ, the average number of occurrences per specified interval which is its only parameter. The probability

of exactly x occurrences can be found out using Poisson probability function. The expected value or the mean of Poisson random variable is λ and its variance is also λ. The standard deviation of Poisson distribution is, $\sqrt{\lambda}$.

NOTES

- Under certain circumstances Poisson distribution can be considered as a reasonable approximation of Binomial distribution and can be used accordingly. The circumstances which permit all this, are when 'n' is large approaching to infinity and p is small approaching to zero (n =Number of trials, p = Probability of 'success').
- When a random variate can take any value in the given interval $a \leq x \leq b$, it is a continuous variate and its distribution is a Continuous Probability Distribution.
- Theoretical distributions are often continuous. They are useful in practice because they are convenient to handle mathematically. They can serve as good approximations to discrete distributions.
- A continuous random variable can take all values in a given interval. A continuous probability distribution is represented by a smooth curve.
- The total area under the curve for a probability distribution is necessarily unity. The curve is always above the x axis because the area under the curve for any interval represents probability and probabilities cannot be negative.
- Among all the probability distributions the normal probability distribution is by far the most important and frequently used continuous probability distribution. This is so because this distribution well fits in many types of problems. This distribution is of special significance in inferential statistics since it describes probabilistically the link between a statistic and a parameter (i.e., between the sample results and the population from which the sample is drawn).
- The name of Karl Gauss, eighteenth century mathematician-astronomer, is associated with normal distribution and in honour of his contribution, this distribution is often known as the Gaussian distribution.
- The normal distribution can be theoretically derived as the limiting form of many discrete distributions. The normal distribution is defined by two parameters, viz., μ and σ^2.
- The mean μ defines where the peak of the curve occurs. In other words, the ordinate at the mean is the highest ordinate. The variance (σ^2) defines the spread of the curve.
- A function that generates moments of a random variable is known as the moment generating function of the random variable. A moment generating function may or may not exist but if it exists it is unique.
- The characteristic function $\phi_X(t)$ of a random variable X is defined as $\phi_X(t) = E(e^{itX})$.

3.8 KEY WORDS

- **Binomial distribution:** It is also called as Bernoulli process and is used to describe discrete random variable.
- **Poisson distribution:** It is used to describe the empirical results of past experiments relating to the problem and plays important role in queuing theory, inventory control problems and risk models.
- **Normal distribution:** It is referred as most important and frequently used continuous probability distribution as it fits well in many types of problems.
- **Moment generating function:** A function that generates moments of a random variable is known as the moment generating function of the random variable, it may or may not exist but if it exists it is unique.

NOTES

3.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define probability distribution and probability functions.
2. Describe binomial distribution and its measures.
3. How a binomial distribution can be fitted to a given data?
4. Describe Poisson distribution and its important measures.
5. Poisson distribution can be an approximation of binomial distribution. Explain.
6. When is the Poisson distribution used?
7. Define normal distribution.
8. Explain any six characteristics of normal distribution.
9. Write the formula for measuring the area under the curve.
10. Explain the circumstances when the normal probability distribution can be used.
11. For the random variable X, define its moment generating function.
12. State the characteristic function $\phi_X(t)$ of a random variable X.

Long-Answer Questions

1. Given is the following probability distribution:

X_i	$pr(X_i)$
0	1/8
1	2/8
2	3/8
3	2/8

NOTES

Calculate the expected value of X_i, its variance, and standard deviation.

2. A coin is tossed 3 times. Let X be the number of runs in the sequence of outcomes: first toss, second toss, third toss. Find the probability distribution of X. What values of X are most probable?
3. (a) Explain the meaning of Bernoulli process pointing out its main characteristics.
   (b) Give a few examples narrating some situations wherein binomial *pr*: distribution can be used.
4. State the distinctive features of the Binomial, Poisson and Normal probability distributions. When does a Binomial distribution tend to become a Normal and a Poisson distribution? Explain.
5. Explain the circumstances when the following probability distributions are used:
   (a) Binomial distribution
   (b) Poisson distribution
   (c) Normal distribution
6. Certain articles were produced of which 0.5 per cent are defective, are packed in cartons, each containing 130 articles. When proportion of cartons are free from defective articles? What proportion of cartons contain 2 or more defective?
   (Given $e^{-0.5}$=0.6065).
7. The following mistakes per page were observed in a book:

No. of Mistakes Per Page	*No. of Times the Mistake Occurred*
0	211
1	90
2	19
3	5
4	0
Total	345

Fit a Poisson distribution to the data given above and test the goodness of fit.

8. In a distribution exactly normal, 7 per cent of the items are under 35 and 89 per cent are under 63. What are the mean and standard deviation of the distribution?
9. Assume the mean height of soldiers to be 68.22 inches with a variance of 10.8 inches. How many soldiers in a regiment of 1000 would you expect to be over six feet tall?

10. Fit a normal distribution to the following data:

Height in inches	*Frequency*
60–62	5
63–65	18
66–68	42
69–71	27
72–74	8

11. Show that the function $f(x)$ given by:

$$f(x) = \begin{cases} x, & 0 < x < 1 \\ k - x, & 1 < x < 2 \\ 0, & \text{elsewhere} \end{cases}$$

is a probability density function for a suitable value of the constant k.

Calculate the probability that the random variable lies between $\frac{1}{2}$ and $\frac{3}{2}$.

12. A random variable X is defined as follows:

$P(X = 1) = p$ and $P(X = 0) = 1 - p$ where $0 < p < 1$.

Find the mean, variance and the central moments μ_2, μ_3, μ_4 of the distribution.

3.10 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

Footnotes

1. The value of the binomial probability function for various values of n and p are also available in tables (known as binomial tables) which can be used for the purpose to ease calculation work. The tables are of considerable help particularly when n is large. (See Appendix)

NOTES

2. For Binomial distribution we had stated that mean = *n. p.*
   ∴ Mean for Poisson distribution (or λ) = *n. p.*
   ∴ $p = \lambda/n$
   Hence, mean = $\frac{\lambda}{n}.n$
3. There are tables which gives the $e^{-\lambda}$ values. These tables also give the $e^{-\lambda}\ \frac{\lambda^x}{x}$ values for $x = 0, 1, 2, \ldots$ for a given λ and thus facilitate the calculation work.
4. Variance of the Binomial distribution is *n. p. q.* and the variance of Poisson distribution is λ. Therefore, λ = *n. p. q.* Since *q* is almost equal to unity and as pointed out earlier *n. p.*= λ in Poisson distribution. Hence, variance of Poisson distribution is also λ.
5. Quite often mathematicians use the normal approximation of the binomial distribution whenever '*n*' is equal to or greater than 30 and *np* and *nq* each are greater than 5.
6. Equation of the normal curve in its simplest form is,

$$y = y_0.e^{\left(x^2/2\sigma^2\right)}$$

   Where, y = The computed height of an ordinate at a distance of X from the mean.
   y_0 = The height of the maximum ordinate at the mean. It is a constant in the equation and is worked out as under:

$$y_0 = \frac{N_i}{\sigma\sqrt{2\pi}}$$

   Where, N = Total number of items in the sample
   i = Class interval
   π = 3.1416
   ∴ = $\sqrt{2\pi} = \sqrt{6.2832} = 2.5066$
   and e = 2.71828, base of natural logarithms
   σ = Standard deviation
   x = Any given value of the dependent variable expressed as a deviation from the mean.
7. A symmetric distribution is one which has no skewness. As such it has the following statistical properties:
   (a) Mean=Mode=Median (i.e., *X*=*Z*=*M*)
   (b) (Upper Quantile – Median)=(Median – Lower Quantile) (i.e., $Q_3 - M = M - Q_1$)
   (c) Mean Deviation=0.7979(Standard Deviation)
   (d) $\frac{Q_3 - Q_1}{2} = 0.6745$ (Standard Deviation)
8. This also means that in a normal distribution, the probability of area lying between various limits are as follows:

Limits	Probability of area lying within the stated limits
μ ± 1 S.D.	0.6827
μ ± 2 S.D.	0.9545
μ ± 3 S.D.	0.9973 (This means that almost all cases lie within μ ± 3 S.D. limits)

9. The table is to be read in the reverse order for finding *Z* value (See Appendix).

UNIT 4 GENERATING FUNCTIONS

NOTES

Structure

4.0 Introduction
4.1 Objectives
4.2 Generating Functions
4.3 The Algebra of Formal Power Series
4.4 Generating Functions for Permutations
4.5 Generating Functions for Partitions
4.6 Answers to Check Your Progress Questions
4.7 Summary
4.8 Key Words
4.9 Self Assessment Questions and Exercises
4.10 Further Readings

4.0 INTRODUCTION

In mathematics, a generating function is a way of encoding an infinite sequence of numbers (a_n) by treating them as the coefficients of a formal power series. This series is called the generating function of the sequence. Unlike an ordinary series, the formal power series is not required to converge: in fact, the generating function is not actually regarded as a function, and the "variable" remains an indeterminate. Generating functions were first introduced by Abraham de Moivre in 1730, in order to solve the general linear recurrence problem. One can generalize to formal power series in more than one indeterminate, to encode information about infinite multi-dimensional arrays of numbers.

A formal power series is a generalization of a polynomial, where the number of terms is allowed to be infinite, with no requirements of convergence. Thus, the series may no longer represent a function of its variable, merely a formal sequence of coefficients, in contrast to a power series, which defines a function by taking numerical values for the variable within a radius of convergence. In a formal power series, the powers of the variable are used only as position-holders for the coefficients, so that the coefficient of x^5 is the fifth term in the sequence. In combinatorics, the method of generating functions uses formal power series to represent numerical sequences and multisets, for instance allowing concise expressions for recursively defined sequences regardless of whether the recursion can be explicitly solved. More generally, formal power series can include series with any finite (or countable) number of variables, and with coefficients in an arbitrary ring.

In this unit, you will study the generating functions, the algebra of formal series, generating functions for permutations and generating functions for partitions.

NOTES

4.1 OBJECTIVES

After going through this unit, you will be able to:

- Explain the definition of generating function
- Discuss about the algebra of formal power series
- Understand the basic concept of generating functions for permutations
- Introduce the basic concept of generating functions for partitions

4.2 GENERATING FUNCTIONS

In mathematics, a generating function is a formal power series in one indeterminate, whose coefficients encode information about a sequence of numbers a_n, i.e., indexed by the natural numbers. Generating functions were originally established by Abraham de Moivre in 1730 for solving the general linear recurrence problem. There are various types of generating functions, such as ordinary generating functions, exponential generating functions, Lambert series, Bell series and Dirichlet series. The specifications of generating function, if any, depends upon the characteristic feature of the sequence and the aspects of the problem being addressed.

Generating functions are often expressed in closed form rather than as a series by some expression involving operations defined for formal power series. These expressions in terms of the indeterminate x may involve arithmetic operations, differentiation with respect to x and composition with (i.e., substitution into) other generating functions. Since these operations are also defined for functions, the result seems like a function of x. For 'generating function', the closed form expression can often be interpreted as a function that can be evaluated at (sufficiently small) concrete values of x and which has the formal power series as its Taylor series. Generating functions are not functions in the formal sense of a mapping from a domain to a codomain; in fact they are sometimes more correctly called generating series.

4.3 THE ALGEBRA OF FORMAL POWER SERIES

A formal power series is a generalization of a polynomial, where the number of terms is allowed to be infinite, with no requirements of convergence. Thus, the series may no longer represent a function of its variable, merely a formal sequence of coefficients, in contrast to a power series, which defines a function by taking numerical values for the variable within a radius of convergence. In a formal power series, the powers of the variable are used only as position-holders for the coefficients, so that the coefficient of x^5 is the fifth term in the sequence. In

NOTES

combinatorics, the method of generating functions uses formal power series to represent numerical sequences and multisets, for instance allowing concise expressions for recursively defined sequences regardless of whether the recursion can be explicitly solved. More generally, formal power series can include series with any finite (or countable) number of variables, and with coefficients in an arbitrary ring.

In algebraic geometry and commutative algebra, rings of formal power series are especially tractable topologically complete local rings, allowing calculus-like arguments within a purely algebraic framework. They are analogous in many ways to p-adic numbers. Formal power series can be created from Taylor polynomials using formal moduli.

Introduction

A formal power series can be loosely thought of as an object that is like a polynomial, but with infinitely many terms. Alternatively, for those familiar with power series (or Taylor series), one may think of a formal power series as a power series in which we ignore questions of convergence by not assuming that the variable X denotes any numerical value (not even an unknown value). For example, consider the series

$$A = 1 - 3X + 5X^2 - 7X^3 + 9X^4 - 11X^5 + \cdots.$$

If we studied this as a power series, its properties would include, for example, that its radius of convergence is 1. However, as a formal power series, we may ignore this completely; all that is relevant is the sequence of coefficients [1, –3, 5, –7, 9, –11, ...]. In other words, a formal power series is an object that just records a sequence of coefficients. It is perfectly acceptable to consider a formal power series with the factorials [1, 1, 2, 6, 24, 120, 720, 5040, ...] as coefficients, even though the corresponding power series diverges for any nonzero value of X.

Arithmetic on formal power series is carried out by simply pretending that the series are polynomials. For example, if

$$B = 2X + 4X^3 + 6X^5 + \cdots,$$

then we add A and B term by term:

$$A + B = 1 - X + 5X^2 - 3X^3 + 9X^4 - 5X^5 + \cdots.$$

We can multiply formal power series, again just by treating them as polynomials (see in particular Cauchy product):

$$AB = 2X - 6X^2 + 14X^3 - 26X^4 + 44X^5 + \cdots.$$

Notice that each coefficient in the product AB only depends on a finite number of coefficients of A and B. For example, the X^5 term is given by

$$44X^5 = (1 \times 6X^5) + (5X^2 \times 4X^3) + (9X^4 \times 2X).$$

NOTES

For this reason, one may multiply formal power series without worrying about the usual questions of absolute, conditional and uniform convergence which arise in dealing with power series in the setting of analysis.

Once we have defined multiplication for formal power series, we can define multiplicative inverses as follows. The multiplicative inverse of a formal power series A is a formal power series C such that $AC = 1$, provided that such a formal power series exists. It turns out that if A has a multiplicative inverse, it is unique, and we denote it by A–1. Now we can define division of formal power series by defining B/A to be the product BA–1, provided that the inverse of A exists. For example, one can use the definition of multiplication above to verify the familiar formula

$$\frac{1}{1+X} = 1 - X + X^2 - X^3 + X^4 - X^5 + \cdots.$$

An important operation on formal power series is coefficient extraction. In its most basic form, the coefficient extraction operator $[X^n]$ applied to a formal power series in one variable extracts the coefficient of the nth power of the variable, so that $[X^2]A = 5$ and $[X^5]A = -11$. Other examples include

$$\begin{aligned}
\left[X^3\right](B) &= 4, \\
\left[X^2\right](X + 3X^2Y^3 + 10Y^6) &= 3Y^3, \\
\left[X^2Y^3\right](X + 3X^2Y^3 + 10Y^6) &= 3, \\
[X^n]\left(\frac{1}{1+X}\right) &= (-1)^n, \\
[X^n]\left(\frac{X}{(1-X)^2}\right) &= n.
\end{aligned}$$

Similarly, many other operations that are carried out on polynomials can be extended to the formal power series setting, as explained below.

The Ring of Formal Power Series

If one considers the set of all formal power series in X with coefficients in a commutative ring R, the elements of this set collectively constitute another ring which is written $R[[X]]$, and called the ring of formal power series in the variable X over R.

Definition of the Formal Power Series Ring

One can characterize $R[[X]]$ abstractly as the completion of the polynomial ring $R[X]$ equipped with a particular metric. This automatically gives $R[[X]]$ the structure of a topological ring (and even of a complete metric space). But the general construction of a completion of a metric space is more involved than what is needed here, and would make formal power series seem more complicated than they are. It is possible to describe $R[[X]]$ more explicitly, and define the ring structure and topological structure separately, as follows.

Ring Structure

NOTES

As a set, $R[[X]]$ can be constructed as the set R^N of all infinite sequences of elements of R, indexed by the natural numbers (taken to include 0). Designating a sequence whose term at index n is a_n by (a_n), One defines addition of two such sequences by

$$(a_n)_{n\in\mathbb{N}} + (b_n)_{n\in\mathbb{N}} = (a_n + b_n)_{n\in\mathbb{N}}$$

and multiplication by

$$(a_n)_{n\in\mathbb{N}} \times (b_n)_{n\in\mathbb{N}} = \left(\sum_{k=0}^{n} a_k b_{n-k}\right)_{n\in\mathbb{N}}.$$

This type of product is called the Cauchy product of the two sequences of coefficients, and is a sort of discrete convolution. With these operations, R^N becomes a commutative ring with zero element (0,0,0, …) becomes a commutative ring with zero element (1,0,0, …).

The product is in fact the same one used to define the product of polynomials in one indeterminate, which suggests using a similar notation. One embeds R into $R[[X]]$ by sending any (constant) $a \in R$ to the sequence (a,0,0,…) and designates the sequence (0,1,0,0,…) by X; then using the above definitions every sequence with only finitely many nonzero terms can be expressed in terms of these special elements as

$$(a_0, a_1, a_2, \ldots, a_n, 0, 0, \ldots) = a_0 + a_1 X + \cdots + a_n X^n = \sum_{i=0}^{n} a_i X^i;$$

these are precisely the polynomials in X. Given this, it is quite natural and convenient to designate a general sequence $(a_n)_{n\in\mathbb{N}}$ by the formal expression

$$\sum_{i\in\mathbb{N}} a_i X^i$$

even though the latter is not an expression formed by the operations of addition and multiplication defined above (from which only finite sums can be constructed). This notational convention allows reformulation of the above definitions as

$$\left(\sum_{i\in\mathbb{N}} a_i X^i\right) + \left(\sum_{i\in\mathbb{N}} b_i X^i\right) = \sum_{i\in\mathbb{N}} (a_i + b_i) X^i$$

and

$$\left(\sum_{i\in\mathbb{N}} a_i X^i\right) \times \left(\sum_{i\in\mathbb{N}} b_i X^i\right) = \sum_{n\in\mathbb{N}} \left(\sum_{k=0}^{n} a_k b_{n-k}\right) X^n.$$

which is quite convenient, but one must be aware of the distinction between formal summation (a mere convention) and actual addition.

NOTES

Topological Structure

Having stipulated conventionally that

$$(a_0, a_1, a_2, a_3, \ldots) = \sum_{i=0}^{\infty} a_i X^i,$$

one would like to interpret the right-hand side as a well-defined infinite summation. To that end, a notion of convergence in R^N is defined and a topology on R^N is constructed. There are several equivalent ways to define the desired topology.

- We may give R^N the product topology, where each copy of R is given the discrete topology.
- We may give R^N we may give $I = X$ is the ideal generated by is the ideal generated by term a_0 is zero.
- The desired topology could also be derived from the following metric.

  The distance between distinct sequences $(a_n), (b_n) \in R^{\mathbb{N}}$ is defined to be $d((a_n), (b_n)) = 2^{-k}$

  where k is the smallest natural number such that $a_k \neq b_k$); the distance between two equal sequences is of course zero.

Informally, two sequences $\{a_n\}$ and $\{b_n\}$ become closer and closer if and only if more and more of their terms agree exactly. Formally, the sequence of partial sums of some infinite summation converges if for every fixed power of X the coefficient stabilizes: there is a point beyond which all further partial sums have the same coefficient. This is clearly the case for the right hand side of (1) regardless of the values a_n since inclusion of the term for $i = n$ gives the last (and in fact only) change to the coefficient of X^n. It is also obvious that the limit of the sequence of partial sums is equal to the left-hand side. This topological structure, together with the ring operations described above, form a topological ring. This is called the ring of formal power series over R is denoted by $R[[X]]$. The topology has the useful property that an infinite summation converges if and only if the sequence of its terms converges to 0, which just means that any fixed power of X occurs in only finitely many terms.

The topological structure allows much more flexible usage of infinite summations. For instance, the rule for multiplication can be restated simply as

$$\left(\sum_{i \in \mathbb{N}} a_i X^i\right) \times \left(\sum_{i \in \mathbb{N}} b_i X^i\right) = \sum_{i,j \in \mathbb{N}} a_i b_j X^{i+j}.$$

since only finitely many terms on the right affect any fixed X^N Infinite products are also defined by the topological structure; it can be seen that an infinite product converges if and only if the sequence of its factors converges to 1.

Operations on Formal Power Series

One can perform algebraic operations on power series to generate new power series. Besides the ring structure operations defined above, we have the following.

NOTES

Power Series Raised to Powers

For any natural number n we have

$$\left(\sum_{k=0}^{\infty} a_k X^k\right)^n = \sum_{m=0}^{\infty} c_m X^m$$

Were

$c_0 = a_0^n,$

$$c_m = \frac{1}{m a_0} \sum_{k=1}^{m} (kn - m + k) a_k c_{m-k}, \quad m \geq 1.$$

(This formula can only be used if m and a_0 are invertible in the ring of coefficients.)

In the case of formal power series with complex coefficients, the complex powers are well defined at least for series f with constant term equal to 1. In this case, f^{α} (This formula can only be used if m and a0 are invertible in the ring of coefficients.) $f^{\alpha} = \exp(\alpha \log(f))$. In the case of formal power series with complex coefficients, the complex powers are well defined at least for series f with constant term $f^{\alpha} g^{\alpha} = (fg)^{\alpha}$ equal to 1. In this case, or as the solution of the differential equation $f(f^{\alpha})' = \alpha f^{\alpha} f'$ with constant term 1, the three definitions being equivalent. The rules of calculus easily follow.

4.4 GENERATING FUNCTIONS FOR PERMUTATIONS

In this section, we will discuss about ordinary and exponential generating functions.

Ordinary Generating Function

The *ordinary generating function* of a sequence a_n is:

$$G(a_n; x) = \sum_{n=0}^{\infty} a_n x^n.$$

When the term *generating function* is used without qualification, it is usually taken to mean an ordinary generating function.

If a_n is the probability mass function of a discrete random variable, then its ordinary generating function is called a probability generating function. The ordinary generating function can be generalized to arrays with multiple indices. For example,

NOTES

the ordinary generating function of a two-dimensional array $a_{m,n}$ (where n and m are natural numbers) is given as:

$$G(a_{m,n};x,y)=\sum_{m,n=0}^{\infty}a_{m,n}x^m y^n$$

Ordinary generating functions for the sequence of square numbers $a_n = n^2$ are given as:

$$G(n^2;x)=\sum_{n=0}^{\infty}n^2x^n=\frac{x(x+1)}{(1-x)^3}$$

Polynomials are a unique case of ordinary generating functions, corresponding to finite sequences or equivalently sequences that disappear after a certain point. These are significant in that many finite sequences can usefully be interpreted as generating functions, such as the Poincaré polynomial, and others.

A key generating function is the constant sequence 1, 1, 1, 1, 1, 1, 1, 1, 1, ..., whose ordinary generating function is of the form:

$$\sum_{n=0}^{\infty}x^n=\frac{1}{1-x}.$$

The left-hand side of the above equation is the Maclaurin series expansion of the right-hand side. Alternatively, the right-hand side expression can be justified by multiplying the power series on the left by $1 - x$, and checking that the result is the constant power series 1, i.e., all coefficients except the one of x^0 disappear. Moreover there can be no other power series with this property. The left-hand side therefore designates the multiplicative inverse of $1 - x$ in the ring of power series.

Expressions for the ordinary generating function of other sequences are easily derived from this one. For example, the substitution $x \to ax$ gives the generating function for the geometric sequence $1,a,a^2,a^3,...$ for any constant a:

$$\sum_{n=0}^{\infty}(ax)^n=\frac{1}{1-ax}.$$

The equality also follows directly from the fact that the left-hand side is the Maclaurin series expansion of the right-hand side. In particular,

$$\sum_{n=0}^{\infty}(-1)^n x^n=\frac{1}{1+ax}.$$

Regular 'gaps' in the sequence can be introduced by replacing x by some power of x, so for example for the sequence 1, 0, 1, 0, 1, 0, 1, 0, ... one gets the generating function of the form:

$$\sum_{n=0}^{\infty} x^{2n} = \frac{1}{1-x^2}.$$

By squaring the initial generating function or by finding the derivative of both sides with respect to x and making a change of running variable $n \to n-1$, we observe that the coefficients form the sequence 1, 2, 3, 4, 5, ..., and so on. Thus we have:

NOTES

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

The third power has as coefficients the triangular numbers 1, 3, 6, 10, 15, 21, ... whose term n is the Binomial coefficient $\binom{n+2}{2}$, so that:

$$\sum_{n=0}^{\infty} \binom{n+2}{2} x^n = \frac{1}{(1-x)^3}.$$

More generally, for any positive integer k, it is true that:

$$\sum_{n=0}^{\infty} \binom{n \mid k}{2} x^n = \frac{1}{(1-x)^{k+1}}.$$

Note that, since

$$2\binom{n+2}{2} - 3\binom{n+1}{1} + \binom{n}{0} = 2\frac{(n+1)(n+2)}{2} - 3(n+1) + 1 = n^2.$$

The ordinary generating function for the sequence 0, 1, 4, 9, 16, ..., of square numbers is obtained by linear combination of Binomial coefficient generating sequences as follows:

$$G\left(n^2; x\right) \sum_{n=0}^{\infty} n^2 x^n = \frac{2}{(1-x)^3} - \frac{3}{(1-x)^2} + \frac{1}{1-x} = \frac{x(x+1)}{(1-x)^3}.$$

Rational Functions: The ordinary generating function of a sequence can be expressed as a rational function (the ratio of two polynomials) if and only if the sequence is a linear recursive sequence with constant coefficients; this generalizes the examples above. Going in the reverse direction, every sequence generated by a fraction of polynomials satisfies a linear recurrence with constant coefficients; these coefficients are identical to the coefficients of the fraction denominator polynomial.

Multiplication Yields Convolution: Multiplication of ordinary generating functions yields a discrete convolution (the Cauchy product) of the sequences. For example, the sequence of cumulative sums $\left(a_0, a_0 + a_1, a_0 + a_1 + a_2, \ldots\right)$ of a sequence with ordinary generating function G(a_n; x) has the generating function

$G(a_n;x)\frac{1}{1-x}$ because 1/(1–*x*) is the ordinary generating function for the sequence (1, 1, ...).

NOTES

Relation to Discrete-Time Fourier Transform: When the series converges absolutely, $G(a_n;e^{-iw}) = \sum_{n=0}^{\infty} a_n e^{-iwn}$ is the discrete-time Fourier transform of the sequence a_0, a_1, ...

Asymptotic Growth of a Sequence: In calculus, often the growth rate of the coefficients of a power series can be used to deduce a radius of convergence for the power series. The reverse can also hold; often the radius of convergence for a generating function can be used to deduce the asymptotic growth of the underlying sequence.

For example, if an ordinary generating function $G(a_n; x)$ that has a finite radius of convergence of *r* can be written as:

$$G(a_n;x) = \frac{A(x) + B(x)(1-x/r)^{-\beta}}{x^{\alpha}}$$

Where *A*(*x*) and *B*(*x*) are functions that are analytic to a radius of convergence greater than *r* (or are entire), and where $B(r) \neq 0$ then:

$$a_n \sim \frac{B(r)}{r^{\alpha}\Gamma(\beta)} n^{\beta-1}(1/r)^n \sim \frac{B(r)}{r^{\alpha}} \binom{n+\beta-1}{n \quad \beta-1} (1/r)^n,$$

This is obtained by using the Gamma function or a Binomial coefficient. Instead, if *G* is an exponential generating function then it is $a_n/n!$ that grows according to these asymptotic formulae.

Asymptotic Growth of the Sequence of Squares: As derived above, the ordinary generating function for the sequence of squares is $\frac{x(x+1)}{(1-x)^3}$. With $r = 1$, $\alpha = 0$, $\beta = 3$, $A(x) = 0$, and $B(x) = x(x+1)$, we can verify that the squares grow as expected, like the squares:

$$a_n \sim \frac{B(r)}{r^{\alpha}\Gamma(\beta)} n^{\beta-1}(1/r)^n = \frac{1(1+1)}{1^0\Gamma(3)} n^{3-1}(1/1)^n = n^2.$$

Asymptotic Growth of the Catalan Numbers: The ordinary generating function for the Catalan numbers is $\frac{1-\sqrt{1-4x}}{2x}$. With $r = 1/4$, $\alpha = 1$, $\alpha = -1/2$, $A(x) = 1/2$, and $B(x) = -1/2$, we can conclude that, for the Catalan numbers:

$$a_n \sim \frac{B(r)}{r^{\alpha}\Gamma(\beta)} n^{\beta-1}(1/r)^n = \frac{-1/2}{(1/4)^1\Gamma(-1/2)} n^{-1/2-1}\left(\frac{1}{1/4}\right)^n = \frac{n^{-3/2}4^n}{\sqrt{\pi}}.$$

NOTES

Bivariate and Multivariate Generating Functions: The generating functions can be defined in several variables for arrays with several indices. These are called multivariate generating functions or sometimes super generating functions. For two variables, these are often called bivariate generating functions.

For example, since $(1+x)^n$ is the ordinary generating function for Binomial coefficients for a fixed n, one may ask for a bivariate generating function that generates the Binomial coefficients $\binom{n}{k}$ for all k and n. To do this, consider $(1+x)^n$ as itself a series, in n, and find the generating function in y that has these as coefficients. Since the generating function for a^n is $1/(1-ay)$, the generating function for the Binomial coefficients is:

$$\sum_{n,k}\binom{n}{k}x^k y^n = \frac{1}{1-(1+x)y} = \frac{1}{1-y-xy}.$$

Example 4.1: Consider the expressions that define a recurrence relation. The sequence is 1, 1, 5, 13, 41, 121, 365... Find a formula for the (n+1)th number in the sequence.

$$x_n = 2x_{n-1} + 3x_{n-2}; \text{ for } n \geq 2$$
$$x_0 = 1$$
$$x_1 = 1$$

Solution: Let $G(z)$ be generating function of the sequence, meaning the coefficient of each power (in ascending order) is the corresponding number in the sequence. So the generating functions is of the form:

$$G(z) = 1 + z + 5z^2 + 13z^3 + 41z^4 + 121z^5 + \ldots$$

Now, by a series of algebraic manipulations, we can find the closed form of the generating function and from that the formula for each coefficient

$$\begin{aligned}
G(z) &= x_0 + x_1 z + x_2 z^2 + x_3 z^3 + x_4 z^2 + x_5 z^5 + \cdots \\
2z \times G(z) &= 2x_0 z + 2x_1 z^2 + \cdots \\
3z^2 \times G(z) &= 3x_0 z^2 + \cdots \\
G(z) - 2zG(z) - 3z^2 G(z) &= x_0 + (x_1 - 2x_0)z + \\
&\quad (x_0 - (2x_1 - 3x_0)z^2 + \\
&\quad (x_3 - (2x_2 - 3x_1)z^3 + \cdots
\end{aligned}$$

NOTES

By definition $x_n - 2x_{n-1} - 3x_{n-2} = 0$

$$(1-2z-3z^2) \times G(z) = x_0 + (x_1 - 2x_0)z$$

$$G(z) = \frac{1-z}{1-2z-3z^2}$$

$$G(z) = \frac{1-z}{(1-3z)(1+z)}$$

By the method of partial fractions we get:

$$G(z) = \frac{1}{2} \times \frac{1}{1-3z} + \frac{1}{2} \times \frac{1}{1+z}$$

Each part of the sum is in a recognizable closed form. We can conclude that:

$$x_n = \frac{1}{2} \times 3^n + \frac{1}{2} \times (-1)^n$$

Exponential Generating Function

The *exponential generating function* of a sequence a_n is:

$$\mathrm{EG}(a_n; x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}.$$

Exponential generating functions are generally more convenient than ordinary generating functions for combinatorial enumeration problems that involve labelled objects.

Exponential generating functions for the sequence of square numbers $a_n = n^2$ are given as:

$$\mathrm{EG}(n^2; x) = \sum_{n=0}^{\infty} \frac{n^2 x^n}{n!} = x(x+1)e^x$$

Typically, the exponential generating function is similar to the ordinary generating function, but with some variations. If $f(x)$ is the exponential generating function for the sequence $\{a_n\}$, we write

$$f(x) \xleftrightarrow{\text{egf}} \{a_n\},$$

This means that the function $f(x)$ has power series representation of the form:

$$f(x) = \sum_{n \geq 0} a_n \frac{x^n}{n!}.$$

As already discussed that the function $\frac{1}{1-x}$ is the ordinary generating function for the sequence that is constantly 1; but for what sequence is it the exponential generating function?

Note that for all n, we have:

$$\frac{1}{1-x} = \sum_{n\geq 0} x^n = \sum_{n\geq 0} n!\frac{x^n}{n!}.$$

If we set $\{a_n\} = \{n!\}$, we have that:

$$\frac{1}{1-x} = \sum_{n\geq 0} n!\frac{x^n}{n!},$$

Consequently we can write:

$$\frac{1}{1-x} \xleftrightarrow{\text{egf}} \{n!\}_{n\geq 0}.$$

Thus what is the exponential generating function of the sequence that is constantly 1? Recall that e^x has the power series representation:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n\geq 0} \frac{x^n}{n!},$$

So we have:

$$e^x \xleftrightarrow{\text{egf}} \{1\}_{n\geq 0}.$$

This specifies that such a generating function is termed 'Exponential Generating Function'.

4.5 GENERATING FUNCTIONS FOR PARTITIONS

In number theory and combinatorics, a partition of a positive integer n, also called an integer partition, is a way of writing n as a sum of positive integers. Two sums that differ only in the order of their summands are considered the same partition. If order matters, the sum becomes a composition. For example, 4 can be partitioned in five distinct ways as follows:

4

3 + 1

2 + 2

2 + 1 + 1

1 + 1 + 1 + 1

The order dependent composition 1 + 3 is the same partition as 3 + 1, while 1 + 2 + 1 and 1 + 1 + 2 are the same partition as 2 + 1 + 1.

Similarly, the seven partitions of 5 are:

NOTES

- 5
- 4 + 1
- 3 + 2
- 3 + 1 + 1
- 2 + 2 + 1
- 2 + 1 + 1 + 1
- 1 + 1 + 1 + 1 + 1

In some sources partitions are treated as the sequence of summands, rather than as an expression with plus signs. For example, the partition 2 + 2 + 1 might instead be written as the tuple (2, 2, 1) or in the even more compact form $(2^2, 1)$ where the superscript indicates the number of repetitions of a term.

A summand in a partition is also called a part. The number of partitions of n is given by the partition function $p(n)$. So $p(4) = 5$. The notation $\lambda\ n$ means that λ is a partition of n. Partition a positive integer n into positive summands and seeking the number of such partitions, without regard to order.

For example, $p(1)=1$: 1

$p(2)=2$: $2=1+1$

$p(3)=3$: $3=2+1=1+1+1$

$p(4)=5$: $4=3+1=2+2=2+1+1=1+1+1+1$

$p(5)=7$: $5=4+1=3+2=3+1+1=2+2+1=2+1+1+1$

$=1+1+1+1+1$

We should like to obtain $p(n)$ for a given n without having to list all the partitions. We need a tool to keep track of the numbers of 1's, 2's, ..., n's that are used as summands for n.

Keep track of 1's: $1+x+x^2+x^3+\cdots$

Keep track of 2's: $1+x^2+x^4+x^6+\cdots$

$\vdots$

Keep track of k's: $1+x^k+x^{2k}+x^{3k}+\cdots$

For example, $p(10)$ is the coefficient of x^{10} in

$$f(x)=(1+x+x^2+\cdots)(1+x^2+x^4+\cdots)\cdots(1+x^{10}+\cdots)$$

$$=\frac{1}{1-x}\frac{1}{1-x^2}\frac{1}{1-x^3}\cdots\frac{1}{1-x^{10}}=\prod_{i=1}^{10}\frac{1}{1-x^i}.$$

In general, $P(x)=\prod_{i=1}^{\infty}\frac{1}{1-x^i}$ generate the sequence $p(0),p(1),\cdots$

Restricted Partitions

NOTES

A restricted partition is a partition in which the parts are constrained in several different ways. For example, we could count partitions that contain only odd numbers. Among the 22 partitions of the number 8, there are 6 that contain only odd parts as follows:

- 7 + 1
- 5 + 3
- 5 + 1 + 1 + 1
- 3 + 3 + 1 + 1
- 3 + 1 + 1 + 1 + 1 + 1
- 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1

Alternatively, we could count partitions in which no number occurs more than once. If we count the partitions of 8 with distinct parts, we also obtain 6 as follows:

- 8
- 7 + 1
- 6 + 2
- 5 + 3
- 5 + 2 + 1
- 4 + 3 + 1

For all positive numbers the number of partitions with odd parts equals the number of partitions with distinct parts. This result was proved by Leonhard Euler in 1748 and is a special case of Glaisher's theorem.

Following are some results concerning restricted partitions:

- The number of partitions of n in which the greatest part is m is equal to the number of partitions of n into m parts.
- The number of partitions of n in which each part is less than or equal to m is equal to the number of partitions of n into m or fewer parts.
- The number of partitions of n in which all parts are equal is the number of divisors of n.
- The number of partitions of n in which all parts are 1 or 2 (or, equivalently, the number of partitions of n into 1 or 2 parts) is:

$$\left\lfloor \frac{n}{2} + 1 \right\rfloor.$$

- The number of partitions of n in which all parts are 1, 2 or 3 (or, equivalently, the number of partitions of n into 1, 2 or 3 parts) is the nearest integer to $(n + 3)^2 / 12$.

NOTES

One class of restricted partition is specified by limiting the partitions to have at most M parts, each of size at most N. Let $p(N,M;n)$ denote the number of such partitions of n. There is a recurrence relation of the form:

$$p(N,M;n) = p(N,M-1;n) + p(N-1,M;n-M)$$

This is obtained by observing that $p(N,M;n) - p(N,M-1;n)$ counts the partitions of n into exactly M parts of size at most N, and subtracting 1 from each part of such a partitions yields a partition of $n–M$.

Generating Function

The generating function for $p(n)$ is given by:

$$\sum_{n=0}^{\infty} p(n)x^n = \prod_{k=1}^{\infty}\left(\frac{1}{1-x^k}\right).$$

Expanding each term on the right-hand side as a geometric series, we can rewrite it as follows:

$$(1 + x + x^2 + x^3 + ...)(1 + x^2 + x^4 + x^6 + ...)(1 + x^3 + x^6 + x^9 + ...) ....$$

The x^n term in this product counts the number of ways to write:

$n = a_1 + 2a_2 + 3a_3 + ... = (1 + 1 + ... + 1) + (2 + 2 + ... + 2) + (3 + 3 + ... + 3) + ...,$

Here each number i appears a_i times. This is precisely the definition of a partition of n, so the product is the desired generating function. More generally, the generating function for the partitions of n into numbers from a set A can be found by taking only those terms in the product where k is an element of A. This result is due to Euler.

The formulation of Euler's generating function is a special case of a q-Pochhammer symbol and is similar to the product formulation of many modular forms, and specifically the Dedekind eta function. The denominator of the product is Euler's function and can be written, by the pentagonal number theorem, as follows:

$$(1-x)(1-x^2)(1-x^3)\ldots = 1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + x^{22} + x^{26} + \ldots$$

Here the exponents of x on the right-hand side are the generalized pentagonal numbers; i.e., numbers of the form $\frac{1}{2}m(3m-1)$, where m is an integer. The signs in the summation alternate as $(-1)^m$. This theorem can be used to derive a recurrence for the partition function:

$p(k) = p(k - 1) + p(k - 2) - p(k - 5) - p(k - 7) + p(k - 12) + p(k - 15) - p(k - 22) - ...$

Where $p(0)$ is taken to equal 1 and $p(k)$ is taken to be zero for negative k.

Another way of stating this is that the value of $p(n)$ can be created from the formula:

$$p(n) = \begin{array}{c|cccccccccc} \text{GPN's} & & & & & & & & & & \\ 0 & 1 & & & & & & & & & \\ 1 & 1 & -1 & & & & & & & & \\ 2 & 1 & 1 & -1 & & & & & & & \\ & 0 & 1 & 1 & -1 & & & & & & \\ & 0 & 0 & 1 & 1 & -1 & & & & & \\ 5 & -1 & 0 & 0 & 1 & 1 & -1 & & & & \\ & 0 & -1 & 0 & 0 & 1 & 1 & -1 & & & \\ 7 & -1 & 0 & -1 & 0 & 0 & 1 & 1 & -1 & & \\ & 0 & -1 & 0 & -1 & 0 & 0 & 1 & 1 & -1 & \\ & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 1 & 1 & \\ \vdots & \vdots & & & & & & & & & \ddots \end{array}_{n\times n}.$$

This illustrates that $p(n)$ is the determinant of the $n\times n$ truncation of the infinite dimensional Toeplitz matrix shown above. The only non-zero diagonals of this matrix start on a row labeled by a generalized pentagonal number q_m. The super diagonal is taken to start on row '0'. On these diagonals, the matrix element is $(-1)^{m+1}$. This follows from a general formula for the quotients for power series.

The generating function for $q(n)$ partitions into distinct parts is given by:

$$\sum_{n=0}^{\infty} q(n)x^n = \prod_{k=1}^{\infty}(1+x^k) = \prod_{k=1}^{\infty}\left(\frac{1}{1-x^{2k-1}}\right).$$

The second product can be written as $\phi(x^2)/\phi(x)$ where ϕ is Euler's function; the pentagonal number theorem can be applied to this as well giving a recurrence for q:

$q(k) = a_k + q(k-1) + q(k-2) - q(k-5) - q(k-7) + q(k-12) + q(k-15) - q(k-22) - \ldots$

Where a_k is $(-1)^m$ if $k = 3m^2 - m$ for some integer m and is 0 otherwise.

The determinant formula for the quotient of power series can be applied to the expression $\phi(x^2)/\phi(x)$ to produce the expression:

$$q(n) = \begin{vmatrix} 1 & & & & & & & & 1 \\ -1 & 1 & & & & & & & 0 \\ -1 & -1 & 1 & & & & & & -1 \\ 0 & -1 & -1 & 1 & & & & & 0 \\ 0 & 0 & -1 & -1 & 1 & & & & -1 \\ 1 & 0 & 0 & -1 & -1 & 1 & & & 0 \\ 0 & 1 & 0 & 0 & -1 & -1 & 1 & & 0 \\ 1 & 0 & 1 & 0 & 0 & -1 & -1 & & 0 \\ \vdots & & & & & & & \ddots & \vdots \end{vmatrix}_{(n+1)\times(n+1)},$$

NOTES

Here the diagonals in the first n columns are constants equal to the coefficients in the power series for $\phi(x)$ and the last column has values a_k given above.

NOTES

Check Your Progress

1. Define generating function.
2. Explain about the algebra of formal power series.
3. What is an ordinary generating function?
4. Define rational numbers.
5. What do you understand by generating functions for partitions?

4.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. A generating function is a formal power series in one indeterminate, whose coefficients encode information about a sequence of numbers a_n, i.e., indexed by the natural numbers. Generating functions were originally established by Abraham de Moivre in 1730 for solving the general linear recurrence problem. There are various types of generating functions, such as ordinary generating functions, exponential generating functions, Lambert series, Bell series and Dirichlet series. The specifications of generating function, if any, depends upon the characteristic feature of the sequence and the aspects of the problem being addressed**.**
2. A formal power series is a generalization of a polynomial, where the number of terms is allowed to be infinite, with no requirements of convergence. Thus, the series may no longer represent a function of its variable, merely a formal sequence of coefficients, in contrast to a power series, which defines a function by taking numerical values for the variable within a radius of convergence.
3. The ordinary generating function of a sequence an is:

$$G(a_n; x) = \sum_{n=0}^{\infty} a_n x^n.$$

4. The ordinary generating function of a sequence can be expressed as a rational function (the ratio of two polynomials) if and only if the sequence is a linear recursive sequence with constant coefficients; this generalizes the examples above. Going in the reverse direction, every sequence generated by a fraction of polynomials satisfies a linear recurrence with constant coefficients; these coefficients are identical to the coefficients of the fraction denominator polynomial.

5. In number theory and combinatorics, a partition of a positive integer n, also called an integer partition, is a way of writing n as a sum of positive integers. Two sums that differ only in the order of their summands are considered the same partition. If order matters, the sum becomes a composition.

NOTES

4.7 SUMMARY

- A generating function is a formal power series in one indeterminate, whose coefficients encode information about a sequence of numbers an, i.e., indexed by the natural numbers.
- In a formal power series, the powers of the variable are used only as position-holders for the coefficients, so that the coefficient of x^5 is the fifth term in the sequence.
- The method of generating functions uses formal power series to represent numerical sequences and multisets, for instance allowing concise expressions for recursively defined sequences regardless of whether the recursion can be explicitly solved.
- The ordinary generating function of a sequence an is:

$$G(a_n; x) = \sum_{n=0}^{\infty} a_n x^n.$$

- Polynomials are a unique case of ordinary generating functions, corresponding to finite sequences or equivalently sequences that disappear after a certain point. These are significant in that many finite sequences can usefully be interpreted as generating functions, such as the Poincare polynomial, and others.
- The ordinary generating function of a sequence can be expressed as a rational function (the ratio of two polynomials) if and only if the sequence is a linear recursive sequence with constant coefficients; this generalizes the examples above. Going in the reverse direction, every sequence generated by a fraction of polynomials satisfies a linear recurrence with constant coefficients; these coefficients are identical to the coefficients of the fraction denominator polynomial.
- Multiplication of ordinary generating functions yields a discrete convolution (the Cauchy product) of the sequences.
- When the series converges absolutely,

$$G\left(a_n; e^{-iw}\right) = \sum_{n=0}^{\infty} a_n e^{-iwn}$$

is the discrete-time Fourier transform of the sequence $a_0, a_1, \ldots$.

NOTES

- In calculus, often the growth rate of the coefficients of a power series can be used to deduce a radius of convergence for the power series. The reverse can also hold; often the radius of convergence for a generating function can be used to deduce the asymptotic growth of the underlying sequence.
- The generating functions can be defined in several variables for arrays with several indices. These are called multivariate generating functions or sometimes super generating functions. For two variables, these are often called bivariate generating functions.
- Exponential generating functions are generally more convenient than ordinary generating functions for combinatorial enumeration problems that involve labelled objects.
- In number theory and combinatorics, a partition of a positive integer n, also called an integer partition, is a way of writing n as a sum of positive integers. Two sums that differ only in the order of their summands are considered the same partition. If order matters, the sum becomes a composition.
- A restricted partition is a partition in which the parts are constrained in several different ways. For example, we could count partitions that contain only odd numbers.

4.8 KEY WORDS

- **Generating function:** Generating function is a formal power series in one indeterminate, whose coefficients encode information about a sequence of numbers an, i.e., indexed by the natural numbers.
- **Formal power series:** A formal power series is a generalization of a polynomial, where the number of terms is allowed to be infinite, with no requirements of convergence. Thus, the series may no longer represent a function of its variable, merely a formal sequence of coefficients, in contrast to a power series, which defines a function by taking numerical values for the variable within a radius of convergence.
- **Poincare polynomial:** Polynomials are a unique case of ordinary generating functions, corresponding to finite sequences or equivalently sequences that disappear after a certain point. These are significant in that many finite sequences can usefully be interpreted as generating functions, such as the Poincare polynomial, and others.
- **Rational function:** The ordinary generating function of a sequence can be expressed as a rational function (the ratio of two polynomials) if and only if the sequence is a linear recursive sequence with constant coefficients; this generalizes the examples above. Going in the reverse direction, every sequence generated by a fraction of polynomials satisfies a linear recurrence with constant coefficients; these coefficients are identical to the coefficients of the fraction denominator polynomial.

- **Restricted partitions:** A restricted partition is a partition in which the parts are constrained in several different ways. For example, we could count partitions that contain only odd numbers.

NOTES

4.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define generating functions.
2. What do you understand by the algebra of formal power series?
3. What is meant by generating functions for permutations?
4. How do we define generating functions for partitions?

Long-Answer Questions

1. Explain briefly about the generating functions and also give the various types of generating functions.
2. What is the algebra of formal power series?
3. Discuss the generating functions for permutations.
4. Describe the basic concept of generating functions for partitions.

4.10 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

NOTES

BLOCK II
RECURRENCE RELATIONS AND SYMMETRIC FUNCTIONS

UNIT 5 RECURRENCE RELATIONS

Structure

5.0 Introduction
5.1 Objectives
5.2 Inventory of Maps
5.3 Recurrence Relations
5.4 Answers to Check Your Progress Questions
5.5 Summary
5.6 Key Words
5.7 Self Assessment Questions and Exercises
5.8 Further Readings

5.0 INTRODUCTION

In mathematics, a recurrence relation is an equation that recursively defines a sequence or multidimensional array of values, once one or more initial terms are given; each further term of the sequence or array is defined as a function of the preceding terms.

The term difference equation sometimes refers to a specific type of recurrence relation. However, "difference equation" is frequently used to refer to any recurrence relation.

A recurrence relation is an equation that expresses each element of a sequence as a function of the preceding ones. More precisely, in the case where only the immediately preceding element is involved.

The Fibonacci numbers, commonly denoted F_n, form a sequence, called the Fibonacci sequence, such that each number is the sum of the two preceding ones, starting from 0 and 1.

In this unit, you will study the basics concepts of inventory of maps and recurrence formula.

5.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe the basics concepts of inventory of maps
- Explain the recurrence formula

5.2 INVENTORY OF MAPS

Factorial

The factorial is defined by the recurrence relation

$n! = n(n-1)!$ for $n > 0$,

and the initial condition

$0! = 1$.

Logistic Map

An example of a recurrence relation is the logistic map:

$x_{n+1} = rx_n(1 - x_n)$,

with a given constant r; given the initial term x_0 each subsequent term is determined by this relation.

Solving a recurrence relation means obtaining a closed-form solution: a non-recursive function of n.

Fibonacci Numbers

The recurrence of order two satisfied by the Fibonacci numbers is the archetype of a homogeneous linear recurrence relation with constant coefficients (see below). The Fibonacci sequence is defined using the recurrence

$$F_n = F_{n-1} + F_{n-2}$$

with initial conditions (seed values)

$$F_0 = 0$$

$$F_1 = 1.$$

Explicitly, the recurrence yields the equations

$$F_2 = F_1 + F_0$$

$$F_3 = F_2 + F_1$$

$$F_4 = F_3 + F_2$$

etc. We obtain the sequence of Fibonacci numbers, which begins

0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...

The recurrence can be solved by methods described below yielding Binet's formula, which involves powers of the two roots of the characteristic polynomial $t^2 = t + 1$; the generating function of the sequence is the rational function

$$\frac{t}{1-t-t^2}.$$

NOTES

Binomial Coefficients

A simple example of a multidimensional recurrence relation is given by the binomial coefficients, $\binom{n}{k}$ which count the number of ways of selecting k elements out of a set of n elements. They can be computed by the recurrence relation which count the number of ways of selecting k elements out of a set of n elements. They can be computed by the recurrence relation

$$\binom{n}{k}=\binom{n-1}{k-1}+\binom{n-1}{k},$$

with the base cases $\binom{n}{0}=\binom{n}{n}=1$ Using this formula to compute the values of all binomial coefficients generates an infinite array called Pascal's triangle. The same values can also be computed directly by a different formula that is not a recurrence, but that requires multiplication and not just addition to compute:

$$\binom{n}{0}=\frac{n!}{k!(n-k)!}.$$

5.3 RECURRENCE RELATIONS

The expressions for permutations and combinations are one of the most fundamental tools for counting the elements of finite sets. They often prove to be inadequate and many problems of computer science require a different approach. An important alternate approach uses recurrence relations (often called recurrence equations or difference equation) to define the terms of a sequence. A formal definition of recurrence relations is difficult because of the wide variety of forms in which such relations can be written, but the concept is straight forward. We have already seen an example of a recurrence relation in the definition of the Fibonacci sequence, where for $n \geq 2$, the term f_n is defined by the recurrence relation

$$f_n = f_{n-1} + f_{n-2}$$

The salient characteristic of a recurence relation is the specification of the term f_n as a function of the terms $f_0, f_1, \ldots, f_{n-1}$. By itself, however, a recurrence relation is not sufficient to define the terms of a sequence; we must also specify the values of some initial terms of the sequence. Thus in our definition of the Fibonacci sequence, we set $f_0 = 0$ and $f_1 = 1$. These are called the boundary conditions or initial conditions of the sequence.

Recall that a recursive definition of a sequence specifies one or more initial terms and a rule for determining subsequent terms from those that precede them. Recursive definitions can be used to solve counting problems. When they are, the rule for finding terms from those that precede them is called a recurrence relation.

NOTES

Recurrence Relation: A recurrence relation for the sequence $\{f_n\}$ is a formula that express f_n in terms of one or more of the previous terms of the sequence, namely, $f_0, f_1,...,f_{n-1}$, for all integers n with $n \geq n_0$, when n_0 is a non-negative integer. A sequence is called a solution of a recurrence relation if its terms satisfy the recurrence relation.

Example 5.1: Determine whether the sequence $\{f_n\}$ is a solution of the recurrence relation

$$f_n = 2f_{n-1} - f_{n-2}$$

for $n = 2, 3, 4,...$ where $f_n = 3n$ for every non-negative integer n.

Solution: Suppose that $f_n = 3n$ for every non-negative integer n, Then, for $n \geq 2$,

$$\begin{aligned} f_n &= 2f_{n-1} - f_{n-2} \\ &= 2[3(n-1)] - 3(n-2) \text{ since } f_n = 3n \\ &= 6n - 6 - 3n + 6 = 3n \end{aligned}$$

Therefore $\{f_n\}$, where $f_n = 3n$, is a solution of the recurrence relation.

Example 5.2: Show that the sequence $\{f_n\}$ is a solution of the recurrence relation $f_n = -3f_{n-1} + 4f_{n-2}$ if $f_n = 2(-4)^n + 3$.

Solution Suppose that $f_n = 2(-4)^n + 3$. Then

$$\begin{aligned} f_n &= -3f_{n-1} + 4f_{n-2} \\ &= -3[2(-4)^{n-1} + 3] + 4[2(-4)^{n-2} + 3] \\ &= -6(-4)^{n-1} - 9 + 8(-4)^{n-2} + 12 \\ &= -6(-4)^{n-1} - 2(-4)^{n-1} + 3 \\ &= -8(-4)^{n-1} + 3 \\ &= 2(-4)^n + 3 \end{aligned}$$

Therefore, $\{f_n\}$, where $f_n = 2(-4)^n + 3$, is a solution of the recurrence relation.

In the following, we shall study a class of recurrence relations known as linear recurrence relations with constant coefficients.

Linear Recurrence Relation: A recurrence relation of the form

$$a_0 f_n + a_1 f_{n-1} + a_2 f_{n-2} + ... + a_k f_{n-k} = f(n) \qquad (1)$$

Where a_1′s are constants and is called a linear recurrence relation with constant coefficients. The recurrence relation (1) is known as a kth-order recurrence relation, provided that both a_0 and a_k are non zero.

Note: The phrase "kth-order" means that each term in the sequence depends only on the k previous terms.

For example, Consider the Fibonacci sequence defined by the recurrence relation $f_n = f_{n-1} + f_{n-2}$, $n \geq 2$ and the initial condition $f_0 = 0$ and $f_1 = 1$. The recurrence relation is called a second-order relation because f_n depends on the two previous terms of f.

NOTES

For example, consider the recurrence relation $f(k) - 5f(k-1) + 6f(k-2) = 4k + 10$ defined for $k \geq 2$, together with the initial condition $f(0) = \frac{7}{3}$ and $f(1) = 5$. Clearly, it is a second-order linear recurrence relation.

Homogeneous Recurrence: A kth-order linear relation is a homogeneous recurrence relation if $f(n) = 0$ for all n. Otherwise, it is called non-homogeneous.

For example, consider the recurrence relation $c(k) - 5c\,(k-1) + 8c\,(k-2) = 0$ together with the initial condition $c(0) = 5$ and $c(1) = 2$. It is a second-order homogeneous recurrence relation.

Example 5.3: Which of the recurrence relations of the following are homogenous and which are non- homogenous?

(i) $f_n = nf_{n-2}$

(ii) $a_n = a_{n-1} + a_{n-3}$

(iii) $b_n = b_{n-1} + 2$

(iv) $s\,(n-2) + s\,(n-4)$

Solution: The relation $f_n = nf_{n-2}$, $a_n = a_{n-1} + a_{n-3}$, $s(n) = s(n-2) + s\,(n-4)$ are all homogeneous and the relation $b_n = b_{n-1} + 2$ is non-homogeneous.

Check Your Progress

1. What is meant by Fibonacci Numbers?
2. Define recurrence relations.

5.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The recurrence of order two satisfied by the Fibonacci numbers is the archetype of a homogeneous linear recurrence relation with constant coefficients (see below). The Fibonacci sequence is defined using the recurrence

   $F_n = F_{n-1} + F_{n-2}$

2. Recurrence relation for the sequence $\{f_n\}$ is a formula that express f_n in terms of one or more of the previous terms of the sequence, namely, $f_0, f_1, \ldots, f_{n-1}$, for all integers n with $\underline{n} \geq \underline{n}_0$, when n_0 is a non-negative integer. A sequence is called a solution of a recurrence relation if its terms satisfy the recurrence relation.

5.5 SUMMARY

NOTES

- The recurrence of order two satisfied by the Fibonacci numbers is the archetype of a homogeneous linear recurrence relation with constant coefficients. The Fibonacci sequence is defined using the recurrence

  $F_n = F_{n-1} + F_{n-2}$
- The expressions for permutations and combinations are one of the most fundamental tools for counting the elements of finite sets.
- The expressions for permutations and combinations are one of the most fundamental tools for counting the elements of finite sets.
- A formal definition of recurrence relations is difficult because of the wide variety of forms in which such relations can be written, but the concept is straight forward. We have already seen an example of a recurrence relation in the definition of the Fibonacci sequence, where for $n \geq 2$, the term f_n is defined by the recurrence relation.
- Recurrence relation for the sequence $\{f_n\}$ is a formula that express f_n in terms of one or more of the previous terms of the sequence, namely, f_0, f_1, ..., f_{n-1}, for all integers n with $\underline{n} \geq \underline{n}_0$, when n_0 is a non-negative integer. A sequence is called a solution of a recurrence relation if its terms satisfy the recurrence relation.
- The phrase "*k*th-order" means that each term in the sequence depends only on the k previous terms.
- A *k*th-order linear relation is a homogeneous recurrence relation if $f(n) = 0$ for all n. Otherwise, it is called non-homogeneous.

5.6 KEY WORDS

- **Fibonacci numbers:** The recurrence of order two satisfied by the Fibonacci numbers is the archetype of a homogeneous linear recurrence relation with constant coefficients. The Fibonacci sequence is defined using the recurrence

  $F_n = F_{n-1} + F_{n-2}$
- **Recurrence relation:** Recurrence relation for the sequence $\{f_n\}$ is a formula that express f_n in terms of one or more of the previous terms of the sequence, namely, $f_0, f_1, ..., f_{n-1}$, for all integers n with $n \geq n_0$, when n_0 is a non-negative integer. A sequence is called a solution of a recurrence relation if its terms satisfy the recurrence relation.
- **Homogeneous recurrence:** A kth-order linear relation is a homogeneous recurrence relation if $f(n) = 0$ for all n. Otherwise, it is called non-homogeneous.

NOTES

5.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define the term Fibonacci numbers
2. What do you understand by recurrence relations?

Long-Answer Questions

1. Explain briefly about the inventory of maps and also discuss about the Fibonacci numbers.
2. Describe the concept of recurrence relations.

5.8 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

UNIT 6 SYMMETRIC FUNCTIONS

NOTES

Structure

6.0 Introduction
6.1 Objectives
6.2 Symmetric Functions
6.3 The Monomial Symmetric Functions m_α
6.4 The Complete Homogeneous Symmetric Functions h_k
6.5 Answers to Check Your Progress Questions
6.6 Summary
6.7 Key Words
6.8 Self Assessment Questions and Exercises
6.9 Further Readings

6.0 INTRODUCTION

A function of n variables is symmetric if its value is the same no matter the order of its arguments. For example, if $f = f(x_1, x_2)$ is a symmetric function, then $f(x_1, x_2) = f(x_2, x_1)$for all x_1 and $x_2$such that (x_1, x_2) and (x_2, x_1) are in the domain of f. The most commonly encountered symmetric functions are polynomial functions, which are given by the symmetric polynomials. A related notion is alternating polynomials, which change sign under an interchange of variables. Aside from polynomial functions, tensors that act as functions of several vectors can be symmetric, and in fact the space of symmetric k-tensors on a vector space V is isomorphic to the space of homogeneous polynomials of degree k on V. Symmetric functions should not be confused with even and odd functions, which have a different sort of symmetry.

Given any function f in n variables with values in an abelian group, a symmetric function can be constructed by summing values of f over all permutations of the arguments. Similarly, an anti-symmetric function can be constructed by summing over even permutations and subtracting the sum over odd permutations. These operations are of course not invertible, and could well result in a function that is identically zero for nontrivial functions f. The only general case where f can be recovered if both its symmetrization and anti-symmetrization are known is when $n = 2$ and the Abelian group admits a division by 2 (inverse of doubling); then f is equal to half the sum of its symmetrization and its anti-symmetrization.

In this unit, you will study about the symmetric functions, the monomial symmetric functions m_α, and the complete homogeneous symmetric functions h_k.

NOTES

6.1 OBJECTIVES

After going through this unit, you will be able to:

- Define symmetric functions
- Explain the monomial symmetric functions m_α
- Comprehend the complete homogeneous symmetric functions h_k

6.2 SYMMETRIC FUNCTIONS

Given any function *f* in *n* variables with values in an Abelian group, a symmetric function can be constructed by summing values of *f* over all permutations of the arguments. Similarly, an anti-symmetric function can be constructed by summing over even permutations and subtracting the sum over odd permutations. These operations are of course not invertible, and could well result in a function that is identically zero for nontrivial functions *f*. The only general case where *f* can be recovered if both its symmetrization and anti-symmetrization are known is when *n* = 2 and the Abelian group admits a division by 2 (inverse of doubling); then *f* is equal to half the sum of its symmetrization and its anti-symmetrization.

A function of *n* variables is symmetric if its value is the same no matter the order of its arguments. For example, if $f = f(x_1, x_2)$ is a symmetric function, then $f(x_1, x_2) = f(x_2, x_1)$ for all x_1 and $x_2$such that (x_1, x_2) and (x_2, x_1) are in the domain of *f*. The most commonly encountered symmetric functions are polynomial functions, which are given by the symmetric polynomials.

A related notion is alternating polynomials, which change sign under an interchange of variables. Aside from polynomial functions, tensors that act as functions of several vectors can be symmetric, and in fact the space of symmetric *k*-tensors on a vector space *V* is isomorphic to the space of homogeneous polynomials of degree *k* on *V*. Symmetric functions should not be confused with even and odd functions, which have a different sort of symmetry.

Consider the real function

$$f(x_1, x_2, x_3) = (x - x_1)(x - x_2)(x - x_3)$$

By definition, a symmetric function with *n* variables has the property that

$$f(x_1, x_2, \ldots, x_n) = f(x_2, x_1, \ldots, x_n) = f(x_3, x_1, \ldots, x_n, x_{n-1})$$

etc.

In general, the function remains the same for every permutation of its variables. This means that, in this case,

$$(x - x_1)(x - x_2)(x - x_3) = (x - x_2)(x - x_1)(x - x_3) = (x - x_3)(x - x_1)(x - x_2)$$

And so on, for all permutations of x_1, x_2, x_3.

NOTES

Now, consider the function $f(x, y) = x^2 + y^2 - r^2$

If x and y are interchanged the function becomes

$$f(y, x) = y^2 + x^2 - r^2$$

Which yields exactly the same results as the original $f(x,y)$.

Again, consider now the function

$$f(x, y) = ax^2 + by^2 - r^2$$

If x and y are interchanged, the function becomes

$$f(y, x) = ay^2 + bx^2 - r^2.$$

This function is obviously not the same as the original if $a \neq b$, which makes it non-symmetric.

6.3 THE MONOMIAL SYMMETRIC FUNCTIONS m_α

Suppose $\alpha = (\alpha_1, \alpha_2, ...)$ is a sequence of non-negative integers, only finitely many of which are non-zero. Then we can consider the monomial defined by α: $X^\alpha = X_1^{\alpha 1} X_2^{\alpha 2} X_3^{\alpha 3} ....$ Then m_α is the symmetric function determined by X^α, i.e., the sum of all monomials obtained from X^α by symmetry. For a formal definition, define $\beta \sim \alpha$ to mean that the sequence β is a permutation of the sequence α and set

$$m_\alpha = \sum_{\beta \sim \alpha} X^\beta.$$

This symmetric function corresponds to the monomial symmetric polynomial $m_\alpha(X_1,...,X_n)$ for any n large enough to have the monomial X_α. The distinct monomial symmetric functions are parameterized by the integer partitions (each m_α has a unique representative monomial X^λ with the parts λ_i in weakly decreasing order). Since any symmetric function containing any of the monomials of some m_α must contain all of them with the same coefficient, each symmetric function can be written as an R-linear combination of monomial symmetric functions, and the distinct monomial symmetric functions therefore form a basis of Λ_R as R-module.

Any monomial in $X_1,...,X_n$ can be written as $X_1^{\alpha 1} \ldots X_n^{\alpha n}$ where the exponents α_i are natural numbers (possibly zero); writing $\alpha = (\alpha_1,\ldots,\alpha_n)$ this can be abbreviated to X^α. The monomial symmetric polynomial $m_\alpha(X_1,...,X_n)$ is defined as the sum of all monomials x^β where β ranges over all distinct permutations of $(\alpha_1,\ldots,\alpha_n)$. For instance one has

$$m_{(3,1,1)}(X_1, X_2, X_3) = X_1^3 X_2 X_3 + X_1 X_2^3 X_3 + X_1 X_2 X_3^3,$$

$$m_{(3,2,1)}(X_1, X_2, X_3) = X_1^3 X_2^2 X_3 + X_1^3 X_2 X_3^2 + X_1^2 X_2^3 X_3 + X_1^2 X_2 X_3^3 + X_1 X_2^3 X_3^2 + X_1 X_2^2 X_3^3.$$

NOTES

Clearly $m_\alpha = m_\beta$ when β is a permutation of α, so one usually considers only those m_α for which $\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n$, in other words for which α is a partition of an integer. These monomial symmetric polynomials form a vector space basis: every symmetric polynomial P can be written as a linear combination of the monomial symmetric polynomials. To do this it suffices to separate the different types of monomial occurring in P. In particular if P has integer coefficients, then so will the linear combination.

6.4 THE COMPLETE HOMOGENEOUS SYMMETRIC FUNCTIONS h_k

The complete homogeneous symmetric functions h_k, for any natural number k; h_k is the sum of all monomial symmetric functions m_α where α is a partition of k. As a power series, this is the sum of all monomials of degree k, which is what motivates its name. This symmetric function corresponds to the complete homogeneous symmetric polynomial $h_k(X_1,...,X_n)$ for any $n \geq k$.

The complete homogeneous symmetric polynomials are a specific kind of symmetric polynomials. Every symmetric polynomial can be expressed as a polynomial expression in complete homogeneous symmetric polynomials.

The complete homogeneous symmetric polynomial of degree k in n variables $X_1,...,X_n$, written h_k for $k = 0, 1, 2, \ldots$, is the sum of all monomials of total degree k in the variables. Formally,

$$h_k(X_1, X_2, \ldots, X_n) = \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \leq n} X_{i_1} X_{i_2} \cdots X_{i_k}.$$

The formula can also be written as:

$$h_k(X_1, X_2, \ldots, X_n) = \sum_{\substack{l_1 + l_2 + \cdots + l_n = k \\ l_i \geq 0}} X_1^{l_1} X_2^{l_2} \cdots X_n^{l_n}.$$

Indeed, l_p is just the multiplicity of p in the sequence i_k.

The first few of these polynomials are

$$h_0(X_1, X_2, \ldots, X_n) = 1,$$

$$h_1(X_1, X_2, \ldots, X_n) = \sum_{1 \leq j \leq n} X_j,$$

$$h_2(X_1, X_2, \ldots, X_n) = \sum_{1 \leq j \leq k \leq n} X_j X_k,$$

$$h_3(X_1, X_2, \ldots, X_n) = \sum_{1 \leq j \leq k \leq l \leq n} X_j X_k X_l.$$

Thus, for each nonnegative integer k, there exists exactly one complete homogeneous symmetric polynomial of degree k in n variables.

Check Your Progress

1. Explain the symmetric functions.
2. Define the monomial symmetric functions.
3. Elaborate on the complete homogeneous symmetric functions.

NOTES

6.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. A function of n variables is symmetric if its value is the same no matter the order of its arguments. For example, if $f = f(x_1, x_2)$ is a symmetric function, then $f(x_1, x_2) = f(x_2, x_1)$for all x_1 and $x_2$such that (x_1, x_2) and (x_2, x_1) are in the domain of f. The most commonly encountered symmetric functions are polynomial functions, which are given by the symmetric polynomials.
2. Suppose $\alpha = (\alpha_1, \alpha_2,...)$ is a sequence of non-negative integers, only finitely many of which are non-zero. Then we can consider the monomial defined by α: $X^{\alpha}=X_1^{\alpha 1}X_2^{\alpha 2}X_3^{\alpha 3}$.... Then m_{α} is the symmetric function determined by X^{α}, i.e., the sum of all monomials obtained from X^{α} by symmetry. For a formal definition, define $\beta \sim \alpha$ to mean that the sequence β is a permutation of the sequence α and set

$$m_{\alpha} = \sum_{\beta \sim \alpha} X^{\beta}.$$

3. The complete homogeneous symmetric functions h_k, for any natural number k; h_k is the sum of all monomial symmetric functions m_{α} where α is a partition of k. As a power series, this is the sum of all monomials of degree k, which is what motivates its name. This symmetric function corresponds to the complete homogeneous symmetric polynomial $h_k(X_1,...,X_n)$ for any $n \geq k$.

6.6 SUMMARY

- Given any function f in n variables with values in an Abelian group, a symmetric function can be constructed by summing values of f over all permutations of the arguments. Similarly, an anti-symmetric function can be constructed by summing over even permutations and subtracting the sum over odd permutations.
- A function of n variables is symmetric if its value is the same no matter the order of its arguments. For example, if $f = f(x_1, x_2)$ is a symmetric function, then $f(x_1, x_2) = f(x_2, x_1)$for all x_1 and $x_2$such that (x_1, x_2) and (x_2, x_1) are in the domain of f. The most commonly encountered symmetric functions are polynomial functions, which are given by the symmetric polynomials.

NOTES

- Suppose $\alpha = (\alpha_1, \alpha_2,...)$ is a sequence of non-negative integers, only finitely many of which are non-zero. Then we can consider the monomial defined by α: $X^{\alpha}=X_1^{\alpha 1}X_2^{\alpha 2}X_3^{\alpha 3}$.... Then m_{α} is the symmetric function determined by X^{α}, i.e., the sum of all monomials obtained from X^{α} by symmetry. For a formal definition, define $\beta \sim \alpha$ to mean that the sequence β is a permutation of the sequence α and set

$$m_{\alpha} = \sum_{\beta \sim \alpha} X^{\beta}.$$

- Since any symmetric function containing any of the monomials of some m_{α} must contain all of them with the same coefficient, each symmetric function can be written as an R-linear combination of monomial symmetric functions, and the distinct monomial symmetric functions therefore form a basis of Λ_R as R-module.
- The complete homogeneous symmetric functions h_k, for any natural number k; h_k is the sum of all monomial symmetric functions m_{α} where α is a partition of k. As a power series, this is the sum of all monomials of degree k, which is what motivates its name. This symmetric function corresponds to the complete homogeneous symmetric polynomial $h_k(X_1,...,X_n)$ for any $n \geq k$.

6.7 KEY WORDS

- **Symmetric functions:** A function of n variables is symmetric if its value is the same no matter the order of its arguments. The most commonly encountered symmetric functions are polynomial functions, which are given by the symmetric polynomials.
- **Monomial symmetric functions:** Suppose $\alpha = (\alpha_1, \alpha_2,...)$ is a sequence of non-negative integers, only finitely many of which are non-zero. Then we can consider the monomial, defined by α: $X^{\alpha}=X_1^{\alpha 1}X_2^{\alpha 2}X_3^{\alpha 3}$.... .
- **Complete homogeneous symmetric functions:** The complete homogeneous symmetric functions h_k, for any natural number k; h_k is the sum of all monomial symmetric functions m_{α} where α is a partition of k.

6.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. State the symmetric functions.
2. Illustrate the monomial symmetric functions.
3. Define the complete homogeneous symmetric functions.

Long-Answer Questions

1. Discuss briefly the symmetric functions with the help of example.
2. Analyse the monomial symmetric functions.
3. Describe the complete homogeneous symmetric functions.

NOTES

6.9 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

NOTES

UNIT 7 ELEMENTARY SYMMETRIC FUNCTIONS

Structure

7.0 Introduction
7.1 Objectives
7.2 The Elementary Symmetric Functions
7.3 The Power Sum Symmetric Functions
7.4 Answers to Check Your Progress Questions
7.5 Summary
7.6 Key Words
7.7 Self Assessment Questions and Exercises
7.8 Further Readings

7.0 INTRODUCTION

The study of symmetric functions is based on that of symmetric polynomials. In a polynomial ring in some finite set of indeterminate, a polynomial is called symmetric if it stays the same whenever the indeterminate are permuted in any way. More formally, there is an action by ring automorphisms of the symmetric group Sn on the polynomial ring in n indeterminate, where a permutation acts on a polynomial by simultaneously substituting each of the indeterminate for another according to the permutation used. The invariants for this action form the subring of symmetric polynomials.

The elementary symmetric functions e_k, for any natural number k; one has $e_k = m_\alpha$ where $X^\alpha = \Pi_{i=1}^{k} X_i$. As a power series, this is the sum of all distinct products of k distinct indeterminate. This symmetric function corresponds to the elementary symmetric polynomial $e_k(X_1,...,X_n)$ for any $n \geq k$.

The name "Symmetric Function" for elements of Λ_R is a misnomer: in neither construction the elements are functions, and in fact, unlike symmetric polynomials, no function of independent variables can be associated to such elements (for instance e_1 would be the sum of all infinitely many variables, which is not defined unless restrictions are imposed on the variables).

In this unit, you will study about the elementary symmetric functions, and the power sum symmetric functions.

7.1 OBJECTIVES

After going through this unit, you will be able to:

- Comprehend the elementary symmetric functions
- Analyse the power sum symmetric functions

NOTES

7.2 THE ELEMENTARY SYMMETRIC FUNCTIONS

The elementary symmetric functions e_k, for any natural number k; one has $e_k = m_\alpha$ where $X^\alpha = \Pi_{i=1}^k X_i$. As a power series, this is the sum of all distinct products of k distinct indeterminate. This symmetric function corresponds to the elementary symmetric polynomial $e_k(X_1,...,X_n)$ for any $n \geq k$.

Consider a set of generators for Λ as a ring, and determine the ring structure of Λ. For each $j \in \mathbb{N}$, the jth elementary symmetric function e_j is m_{1^j}, where 1^j denotes the partition with j parts all equal to 1. More clearly,

$$e_j := \sum_{1 \leq i_1 < \cdots < i_j} x_{i_1} \cdots x_{i_j}.$$

In particular, $e_0 = 1$ and $e_1 = x_1+x_2+x_3+\cdots$ and $e_2 = x_1x_2+x_1x_3+x_1x_4+\cdots+x_2x_3+\cdots$.

Another way to view this definition is in terms of the generating function

$$E(t) := \sum_{j=0}^{\infty} e_j t^j = \prod_{i=1}^{\infty} (1 \mid x_i t).$$

The case $e_0 = 1$ is special; the other e_j are referred to as one-part elementary symmetric functions, for reasons which will become clear in a moment.

Proposition: The one-part elementary symmetric functions $\{e_1, e_2, \ldots\}$ generate Λ as a ring and are algebraically independent over $\mathbb{Z}$. Therefore,

$$\Lambda = \mathbb{Z}[e_1, e_2, \ldots]$$

is a polynomial ring over $\mathbb{Z}$ in infinitely many indeterminate $\{e_1, e_2, \ldots\}$.

Proposition: The set $\{e_\lambda\}$ (where λ ranges over all partitions) is a $\mathbb{Z}$-bias for Λ. That is, every symmetric function can be written uniquely as a finite $\mathbb{Z}$-linear combination of elementary symmetric functions.

To prove Propositions, we can restrict our attention to Λ^n, the symmetric functions which are homogeneous of degree n, for each $n \in$ N separately. Since $\{m_\lambda : \lambda \vdash n\}$ is a (finite) basis for Λ^n over $\mathbb{Z}$, we can express each elementary symmetric function of degree n uniquely in terms of this basis.

7.3 THE POWER SUM SYMMETRIC FUNCTIONS

The power sum symmetric functions p_k, for any positive integer k; one has $p_k = m_{(k)}$, the monomial symmetric function for the monomial X_1^k. This symmetric function corresponds to the power sum symmetric polynomial $p_k(X_1,...,X_n) = X_1^k+...+X_n^k$ for any $n \geq 1$.

There is no power sum symmetric function p0: although it is possible (and in some contexts natural) to define $p_0(X_1, \ldots, X_n) = \Sigma_{i=1}^n X_i^0 = n$ as a symmetric polynomial in n variables, these values are not compatible with the morphisms ρ_n.

NOTES

The "Discriminant" $(\prod_{i<j}(X_i - X_j))^2$ is another example of an expression giving a symmetric polynomial for all n, but not defining any symmetric function. The expressions defining Schur polynomials as a quotient of alternating polynomials are somewhat similar to that for the discriminant, but the polynomials $s_\lambda(X_1,...,X_n)$ turn out to be compatible for varying n, and therefore do define a symmetric function.

The power sum symmetric polynomials are a type of basic building block for symmetric polynomials, in the sense that every symmetric polynomial with rational coefficients can be expressed as a sum and difference of products of power sum symmetric polynomials with rational coefficients. However, not every symmetric polynomial with integral coefficients is generated by integral combinations of products of power-sum polynomials: they are a generating set over the rationales, but not over the integers.

The power sum symmetric polynomial of degree k in n variables $x_1, ..., x_n$, written p_k for $k = 0, 1, 2, ...$, is the sum of all kth powers of the variables. Formally,

$$p_k(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_i^k .$$

The first few of these polynomials are

$$p_0(x_1, x_2, \ldots, x_n) = n,$$

$$p_1(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \cdots + x_n ,$$

$$p_2(x_1, x_2, \ldots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2 ,$$

$$p_3(x_1, x_2, \ldots, x_n) = x_1^3 + x_2^3 + \cdots + x_n^3 .$$

Thus, for each nonnegative integer k there exists exactly one power sum symmetric polynomial of degree k in n variables. The polynomial ring formed by taking all integral linear combinations of products of the power sum symmetric polynomials is a commutative ring.

Check Your Progress

1. Explain the elementary symmetric functions.
2. Define the power sum symmetric functions.

7.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The elementary symmetric functions e_k, for any natural number k; one has $e_k = m_\alpha$ where $X^\alpha = \Pi_{i=1}^k X_i$. As a power series, this is the sum of all distinct products of k distinct indeterminate. This symmetric function corresponds to the elementary symmetric polynomial $e_k(X_1,...,X_n)$ for any $n \geq k$.

2. The power sum symmetric functions p_k, for any positive integer k; one has $p_k = m_{(k)}$, the monomial symmetric function for the monomial X_1^k. This symmetric function corresponds to the power sum symmetric polynomial $p_k(X_1,...,X_n) = X_1^k+...+X_n^k$ for any $n \geq 1$.

7.5 SUMMARY

- The elementary symmetric functions e_k, for any natural number k; one has $e_k = m_\alpha$ where $X^\alpha = \Pi_{i=1}^k X_i$. As a power series, this is the sum of all distinct products of k distinct indeterminate. This symmetric function corresponds to the elementary symmetric polynomial $e_k(X_1,...,X_n)$ for any $n \geq k$.
- Consider a set of generators for as a ring, and determine the ring structure of Λ. For each $j \in \mathbb{N}$, the jth elementary symmetric function e_j is m_{1^j}, where 1^j denotes the partition with j parts all equal to 1. More clearly,

$$e_j := \sum_{1 \leq i_1 < \cdots < i_j} x_{i_1} \cdots x_{i_j}.$$

- The one-part elementary symmetric functions $\{e_1, e_2, \ldots\}$ generate Λ as a ring and are algebraically independent over $\mathbb{Z}$.
- The set $\{e_\lambda\}$ (where λ ranges over all partitions) is a $\mathbb{Z}$-bias for Λ. That is, every symmetric function can be written uniquely as a finite $\mathbb{Z}$-linear combination of elementary symmetric functions.
- The power sum symmetric functions p_k, for any positive integer k; one has $p_k = m_{(k)}$, the monomial symmetric function for the monomial X_1^k. This symmetric function corresponds to the power sum symmetric polynomial $p_k(X_1,...,X_n) = X_1^k+...+X_n^k$ for any $n \geq 1$.
- There is no power sum symmetric function p0: although it is possible (and in some contexts natural) to define $p_0(X_1, \ldots, X_n) = \Sigma_{i=1}^n X_i^0 = n$ as a symmetric polynomial in n variables, these values are not compatible with the morphisms ρ_n.
- The power sum symmetric polynomials are a type of basic building block for symmetric polynomials, in the sense that every symmetric polynomial with rational coefficients can be expressed as a sum and difference of products of power sum symmetric polynomials with rational coefficients.

7.6 KEY WORDS

- **Elementary symmetric functions:** The elementary symmetric functions e_k, for any natural number k; one has $e_k = m_\alpha$ where $X^\alpha = \Pi_{i=1}^k X_i$. As a power series, this is the sum of all distinct products of k distinct indeterminate.

NOTES

This symmetric function corresponds to the elementary symmetric polynomial $e_k(X_1,...,X_n)$ for any $n \geq k$.

- **Power sum symmetric functions:** The power sum symmetric functions p_k, for any positive integer k; one has $p_k = m_{(k)}$, the monomial symmetric function for the monomial X_1^k. This symmetric function corresponds to the power sum symmetric polynomial $p_k(X_1,...,X_n) = X_1^k+...+X_n^k$ for any $n \geq 1$.

7.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. State the elementary symmetric functions.
2. Elaborate on the power sum symmetric functions.

Long-Answer Questions

1. Briefly define the elementary symmetric functions.
2. Analyse the power sum symmetric functions.

7.8 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

UNIT 8 MULTINOMIALS

NOTES

Structure

8.0 Introduction
8.1 Objectives
8.2 Multinomials
8.3 Basic Concepts of Multinomials
8.4 Answers to Check Your Progress Questions
8.5 Summary
8.6 Key Words
8.7 Self Assessment Questions and Exercises
8.8 Further Readings

8.0 INTRODUCTION

In mathematics, the multinomial theorem describes how to expand a power of a sum in terms of powers of the terms in that sum. It is the generalization of the binomial theorem from binomials to multinomials.

In statistical mechanics and combinatorics if one has a number distribution of labels then the multinomial coefficients naturally arise from the binomial coefficients. Given a number distribution $\{ni\}$ on a set of N total items, ni represents the number of items to be given the label 'i'. (In statistical mechanics 'i' is the label of the energy state.)

A polynomial is an expression consisting of variables (also called indeterminates) and coefficients, that involves only the operations of addition, subtraction, multiplication, and non-negative integer exponentiation of variables. An example of a polynomial of a single indeterminate x is $x^2 - 4x + 7$. An example in three variables is $x^3 + 2xyz^2 - yz + 1$.

All polynomials with coefficients in a unique factorization domain (for example, the integers or a field) also have a factored form in which the polynomial is written as a product of irreducible polynomials and a constant. This factored form is unique up to the order of the factors and their multiplication by an invertible constant. In the case of the field of complex numbers, the irreducible factors are linear. Over the real numbers, they have the degree either one or two. Over the integers and the rational numbers, the irreducible factors may have any degree.

In this unit, you will study the multinomials and basic concepts of multinomials.

8.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Explain the definition of multinomials
- Describe the basic concept of multinomials

8.2 MULTINOMIALS

Multinomial means a mathematical expression consisting of two or more than two terms, for example; $a_1x_1 + a_2x_2 + a_3x_3 + \ldots + a_kx_k$.

A binomial theorem is given as $(x + y)^n = S\ {}^nC_r\, x^r\, y^{n-r}$ where $r = 0, 1, 2 \ldots n$ and $n = 1, 2, 3 \ldots$ and $0! = 1$. Here ${}^nC_r$ is called binomial coefficient and has following meaning:

1. The number of selections taking r objects out of n given objects.
2. Number of r-element subsets of an n-element set.
3. Number of n-letter sequences having exactly r as power of (x and $(n-r)$ as power of y).
4. The coefficient of $x^r y^{n-r}$ when expanding $(x+y)^n$ and collecting terms.

Extending this, we get multinomial theorem as:

$(x_1 + x_2 + \ldots + x_r)^{n-} = \sum {}^nC_{n1,n2,n3,\ldots,nr}\, x_1^{n1} x_2^{n2} \ldots x_r^{nr}$ where $n = n_1 + n_2 + \ldots + n_r$.

In the above expansion, ${}^nC_{n1,n2,n3,\ldots,nr} = n! / (n_1!n_2!n_3!\ldots n_r!)$ is known as multinomial coefficient which has following interpretations:

1. It stands for number of ways of splitting n distinct objects into r distinct groups of sizes $n_1, \ldots, n_r$, respectively. (In the case $n_1 = \cdots = n_r = 1$ this is the number of ways to permute all n objects.)
2. It is the number of n-letter words formed with r distinct letters, say, $L_1, \ldots, L_r$, used $n_1, \ldots, n_r$ times respectively.
3. It is the coefficient of $x_1^{n1} \ldots x_n^{nr}$ when expanding $(x_1 + \cdots + x_r)^n$ and collecting terms.

For example, if we have to find $(x + y + z)^3$ we may find by expanding its square first and then multiplying the whole term by $(x + y + z)$ once again and get,

$$(x + y + z)^3 = x^3 + y^3 + z^3 + 3x^2y + 3x^2z + 3y^2x + 3y^2z + 3z^2x + 3z^2y + 6xyz$$

This process is slow, and use of the multinomial theorem will provide fast solution and gives us solution for any coefficient desired. We can get the multinomial coefficients of the term, using formulae. If we want to know the coefficient of a term having combination of powers of x, y and z, we can find as below:

Coefficient of $x^2y^0z^1 = {}^3C_{2,0,1} = 3!/(2!.0!.1!) = 3$

Coefficient of $x^1y^1z^1 = {}^3C_{1,1,1} = 3!/(1!1!1!) = 6$

Similarly coefficient of other combination can also be found.

Interpretations of Multinomial Theorem

(1) **Finding ways of putting objects into boxes of a particular size:** Like binomial coefficients, multinomial too can be interpreted in terms of combinations. This can be viewed as a problem of putting *n* distinct objects in *m* number of storage bins such that first bin contains a_1 objects, second contains a_2, a_3 in the third, and so on.

(2) **Unique permutations of words:** Multinomial coefficients have interpretation of representing number of permutations in a multi set of n elements having a_i as the multiplicity of every distinct element and $a_1 + a_2 + a_3 + \ldots\ldots + a_n = n$.

For example, we take the word 'MISSISSIPPI' and try to find numbers of permutations from the letters of the word. We have M – 1, I – 4, S – 4, and P – 2.

Number of desired permutations are: ${}^{11}C_{1,4,4,2} = 11!/(1!.4!.4!.2!) = 34650$.

Binomial theorem is used to make Pascal's triangle. With multinomial theorem, one can make Pascal's pyramid which a generalization from Pascal's triangle. Pascal pyramid is also called Pascal's simplex. This can create a lookup table for multinomial coefficients.

In the multinomial coefficient with $n = 3$, one can easily draw this by hand. If $n = 4$, then this results in a series of growing pyramids.

8.3 BASIC CONCEPTS OF MULTINOMIALS

The multinomial theorem describes how to expand a power of a sum in terms of powers of the terms in that sum. It is the generalization of the binomial theorem from binomials to multinomials.

Theorem

For any positive integer *m* and any non-negative integer *n*, the multinomial formula describes how a sum with *m* terms expands when raised to an arbitrary power *n*:

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{k_1+k_2+\cdots+k_m=n} \binom{n}{k_1, k_2, \ldots, k_m} \prod_{t=1}^{m} x_t^{k_t},$$

where

$$\binom{n}{k_1, k_2, \ldots, k_m} = \frac{n!}{k_1!\, k_2! \cdots k_m!}$$

NOTES

NOTES

is a multinomial coefficient. The sum is taken over all combinations of nonnegative integer indices k_1 through k_m such that the sum of all k_i is n. That is, for each term in the expansion, the exponents of the x_i must add up to n. Also, as with the binomial theorem, quantities of the form x_0 that appear are taken to equal 1 (even when x equals zero).

In the case $m = 2$, this statement reduces to that of the binomial theorem.

Example

The third power of the trinomial $a + b + c$ is given by

$$(a+b+c)^3 = a^3 + b^3 + c^3 + 3a^2b + 3a^2c + 3b^2a + 3b^2c + 3c^2a + 3c^2b + 6abc.$$

This can be computed by hand using the distributive property of multiplication over addition, but it can also be done (perhaps more easily) with the multinomial theorem. It is possible to "read off" the multinomial coefficients from the terms by using the multinomial coefficient formula. For example:

$$a^2b^0c^1 \text{ has the coefficient } \binom{3}{2,0,1} = \frac{3!}{2!\cdot 0!\cdot 1!} = \frac{6}{2\cdot 1\cdot 1} = 3.$$

$$a^1b^1c^1 \text{ has the coefficient } \binom{3}{1,1,1} = \frac{3!}{1!\cdot 1!\cdot 1!} = \frac{6}{1\cdot 1\cdot 1} = 6.$$

Alternate Expression

The statement of the theorem can be written concisely using multiindices:

$$(x_1 + \cdots + x_m)^n = \sum_{|\alpha|=n} \binom{n}{\alpha} x^\alpha$$

were

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$$

and

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_m^{\alpha_m}$$

Proof

This proof of the multinomial theorem uses the binomial theorem and induction on m.

First, for $m = 1$, both sides equal x_1^n since there is only one term $k_1 = n$ in the sum. For the induction step, suppose the multinomial theorem holds for m. Then

$$(x_1 + x_2 + \cdots + x_m + x_{m+1})^n = (x_1 + x_2 + \cdots + (x_m + x_{m+1}))^n$$

$$= \sum_{k_1+k_2+\cdots+k_{m-1}+K=n} \binom{n}{k_1, k_2, \ldots, k_{m-1}, K} x_1^{k_1} x_2^{k_2} \cdots x_{m-1}^{k_{m-1}} (x_m + x_{m+1})^K$$

by the induction hypothesis. Applying the binomial theorem to the last factor,

$$= \sum_{k_1+k_2+\cdots+k_{m-1}+K=n} \binom{n}{k_1, k_2, \ldots, k_{m-1}, K} x_1^{k_1} x_2^{k_2} \cdots x_{m-1}^{k_{m-1}} \sum_{k_m+k_{m+1}=K} \binom{K}{k_m, k_{m+1}} x_m^{k_m} x_m^{k_n}$$

$$= \sum_{k_1+k_2+\cdots+k_{m-1}+k_m+k_{m+1}=n} \binom{n}{k_1, k_2, \ldots, k_{m-1}, k_m, k_{m+1}} x_1^{k_1} x_2^{k_2} \cdots x_{m-1}^{k_{m-1}} x_m^{k_m} x_{m+1}^{k_{m+1}}$$

which completes the induction. The last step follows because

$$\binom{n}{k_1, k_2, \ldots, k_{m-1}, K} \binom{K}{k_m, k_{m+1}} = \binom{n}{k_1, k_2, \ldots, k_{m-1}, k_m, k_{m+1}},$$

as can easily be seen by writing the three coefficients using factorials as follows:

$$\frac{n!}{k_1!k_2!\cdots k_{m-1}!K!} \frac{K!}{k_m!k_{m+1}!} = \frac{n!}{k_1!k_2!\cdots k_{m+1}!}.$$

Check Your Progress

1. Define the term multinomials.
2. What do you mean by binomial coefficient?
3. What do you understand by interpretations of multinomial theorem?
4. Explain about the basic concept of multinomial.

8.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Multinomial means a mathematical expression consisting of two or more than two terms, for example; $a_1x_1 + a_2x_2 + a_3x_3 + \ldots. + a_kx_k$. A binomial theorem is given as $(x + y)^n = S\,{}^nC_r\, x^r\, y^{n-r}$ where $r = 0,1,2, \ldots n$ and $n = 1,2,3, \ldots$ and 0 = 1. Here ${}^nC_r$ is called binomial coefficient.
2. Binomial coefficient and has following meaning:
   - The number of selections taking *r* objects out of n given objects.
   - Number of *r*-element subsets of an *n*-element set.
   - Number of *n*-letter sequences having exactly *r* as power of (*x* and (*n* – *r*) as power of *y*).
   - The coefficient of $x^r y^{n-r}$ when expanding $(x + y)^n$ and collecting terms.
3. Finding ways of putting objects into boxes of a particular size: Like binomial coefficients, multinomial too can be interpreted in terms of combinations. This can be viewed as a problem of putting n distinct objects in *m* number of storage bins such that first bin contains a_1 objects, second contains a_2, a_3 in the third, and so on. Unique permutations of words: Multinomial coefficients have interpretation of representing number of permutations in a multi set of *n* elements having a_i as the multiplicity of every distinct element and $a_1 + a_2 + a_3 + \ldots\ldots + a_n = n$.

NOTES

4. The multinomial theorem describes how to expand a power of a sum in terms of powers of the terms in that sum. It is the generalization of the binomial theorem from binomials to multinomials.

NOTES

8.5 SUMMARY

- Multinomial means a mathematical expression consisting of two or more than two terms, for example; $a_1x_1 + a_2x_2 + a_3x_3 + ... + a_kx_k$. A binomial theorem is given as $(x + y)^n = S\ {}^nC_r x^r y^{n-r}$ where $r = 0,1,2, \ldots n$ and $n = 1,2,3, \ldots$ and $0 = 1$. Here ${}^nC_r$ is called binomial coefficient.
- The number of selections taking r objects out of n given objects. Number of r-element subsets of an n-element set. Number of n-letter sequences having exactly r as power of (x and $(n – r)$ as power of y). The coefficient of $x^r y^{n-r}$ when expanding $(x + y)^n$ and collecting terms.
- It stands for number of ways of splitting n distinct objects into r distinct groups of sizes $n_1, ..., n_r$, respectively. (In the case $n_1 = \cdots = n_r = 1$ this is the number of ways to permute all n objects.)
- It is the number of n-letter words formed with r distinct letters, say, $L_1, \ldots, L_r$, used $n_1, ..., n_r$ times respectively.
- It is the coefficient of $x_1^{n1} ... x_n^{nr}$ when expanding $(x_1 + \cdots + x_r)^n$ and collecting terms.
- Like binomial coefficients, multinomial too can be interpreted in terms of combinations. This can be viewed as a problem of putting n distinct objects in m number of storage bins such that first bin contains a_1 objects, second contains a_2, a_3 in the third, and so on.
- Multinomial coefficients have interpretation of representing number of permutations in a multi set of n elements having a_i as the multiplicity of every distinct element and $a_1 + a_2 + a_3 + \ldots\ldots + a_n = n$.
- The multinomial theorem describes how to expand a power of a sum in terms of powers of the terms in that sum. It is the generalization of the binomial theorem from binomials to multinomials.

8.6 KEY WORDS

- **Binomial coefficient:** The number of selections taking r objects out of n given objects. Number of r-element subsets of an n-element set. Number of n-letter sequences having exactly r as power of (x and $(n – r)$ as power of y). The coefficient of $x^r y^{n-r}$ when expanding $(x + y)^n$ and collecting terms.
- **Multinomial coefficient:** It stands for number of ways of splitting n distinct objects into r distinct groups of sizes $n_1, ..., n_r$, respectively. (In the case

$n_1 = \cdots = n_r = 1$ this is the number of ways to permute all n objects.) It is the number of n-letter words formed with r distinct letters, say, $L_1, ..., L_r$, used $n_1, ..., n_r$ times respectively. It is the coefficient of $x_1^{n1} \ldots x_n^{nr}$ when expanding $(x_1 + \cdots + x_r)^n$ and collecting terms.

- **Pascal's simplex:** Binomial theorem is used to make Pascal's triangle. With multinomial theorem, one can make Pascal's pyramid which a generalization from Pascal's triangle. Pascal pyramid is also called Pascal's simplex.

8.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define multinomials.
2. How do we define the basic concept of multinomials?

Long-Answer Questions

1. What do you mean by binomial coefficient? Explain the interpretations of multinomial theorem.
2. Describe the basic concept of multinomials.

8.8 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

NOTES

NOTES

BLOCK III
PRINCIPLES, PERMUTATIONS AND POLYA THEORY

UNIT 9 INCLUSION AND EXCLUSION PRINCIPLES

Structure

9.0 Introduction
9.1 Objectives
9.2 Inclusion and Exclusion Principle and Theorems
9.3 Answers to Check Your Progress Questions
9.4 Summary
9.5 Key Words
9.6 Self Assessment Questions and Exercises
9.7 Further Readings

9.0 INTRODUCTION

In combinatorics, a branch of mathematics, the inclusion–exclusion principle is a counting technique which generalizes the familiar method of obtaining the number of elements in the union of two finite sets.

The Principle of Inclusion and Exclusion (PIE) is a counting technique that computes the number of elements that satisfy at least one of several properties while guaranteeing that elements satisfying more than one property are not counted twice.

PIE is based on the principle of preventing double counting by adding the number of elements that fit into at least one of two groups and subtracting the overlap. For example, the number of people who own at least one cat or dog can be calculated by adding the number of people who own a cat to the number of people who own a dog, then subtracting the number of people who own both.

In this unit, you will study the basics concepts of inclusion and exclusion principle.

9.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe the basics concepts of inclusion and exclusion principle
- Define finite set and infinite set
- Explain the term cardinality

9.2 INCLUSION AND EXCLUSION PRINCIPLE AND THEOREMS

NOTES

The basic principle of counting the things defines that each object should be counted only once. If there are N number of objects given and out of these N objects some objects have property *a* (denoted as N_a) and some possess property *b* (denoted as N_b) while some have both the properties *a* and *b* (denoted as N*ab*). If V_{ab} denotes number of objects having either of these properties *a* or *b*, then $V_{ab} = N_a + N_b - N_{ab}$.

This is just like a simple addition on elements of sets like $n(A \cup B) = n(A) + n(B) - n(A \cap B)$.

Applying formula $V_{ab} = N_a + N_b - N_{ab}$ recursively we include one more property c, and now we can write as: $V_{abc} = N_a + N_b + N_c - N_{ab} - N_{bc} - N_{ac} + N_{abc}$ and going further for properties *abc*…*n* we can write as: $V_{abc....n} = N_a + N_b + N_c + ... + N_n - N_{ab} - N_{bc} - N_{ac} - .... - N_{mn} + N_{abc} + ... + N_{lmn} - N_{abcd} - ....$

This formula can also be written in a set theoretic form for sets A, B, C, …, etc., and N_a can be written as |A| or #A or *n*(A) and similarity for N_b, N_c or other and N_{ab} can be written as $|A \cap B|$ or $\#(A \cap B)$ or $n(A \cap B)$. This is the principle of inclusion and exclusion (written in short as PIE) with the condition that each object is counted only once. We take an arbitrary object, say S out of a set of given N number of objects and assume that the object S has property *k* out of the many properties in the set of N. To prove that the formulae is correct according to principle of inclusion and exclusion we have to first prove that as per the above given formulae the object S is counted only once.

For example, consider the following solved examples.

Example 9.1: A coin is flipped 5 times. In how many ways can it be done so that there is an exact sequence of 3 heads in a row? The case of 4 heads in a row is not counted. If N*i* be the number of sequences of tosses having an exact sequence of 3 heads and starts on the *i*th throw then what will be V_{12345}?

Solution: V_{12345} is equal to V_{123} because an exact sequence of 3 heads cannot directly start on the 4th or 5th throws. According to PIE formulae: $V_{123} = N_1 + N_2 + N_3 - N_{12} - N_{23} - N_{13} + N_{123}$

We may not have exact sequences of heads starting on the 1st and 2nd throws. Let us write H for head and T for tails. If N_1 is the number of sequences exactly starting with 3 H then the 4th throw must be the tail (T) leaving two possibilities for the 5th throw. Thus, $N_1 = 2$.

In a similar way, if there is an exact sequence of 3 H starting with the 2nd toss then this means that both the 1st and 5th tosses must be T or tail. This leads to $N_2 = 1$. Hence, if a sequence of H starts on the 3rd throw, then the 2nd throw

NOTES

has to be T while the 1st throw may be either H or T. This gives $N_3 = 2$. Hence, $V_{12345} = V_{123} = 2 + 1 + 2 = 5$.

Example 9.2: A coin is flipped 5 times. In how many ways can it be done so that there must be a sequence of at least 3 heads in a row? If N*i* be the number of sequences of tosses with at least 3 heads beginning on the *i*th throw then what will be V_{12345}?

Solution: In this case, N*i* is the number of exact sequences of tosses where there is a sequence of at least 3 heads starting on the *i*th throw. Here, at least 3 means that 3 or more heads. Like previous example, $V_{12345} = V_{123}$ and $V_{123} = N_1 + N_2 + N_3 - N_{12} - N_{23} - N_{13} + N_{123}$

Also in this case the first three terms only can be non-zero and $N_1 = 4$ because each 4th and 5th tosses can be H or T. Also, $N_2 = 2$, because 5th throw can either be H or T. In a similar way $N_3 = 2$ as the 1st throw may be H or T but the 2nd throw must be tails. $V_{12345} = V_{123} = 8$.

Example 9.3: In a class room there are 15 students. Out of these 6 study mechanics, 9 study general science, and 9 study computer science. Also, 2 study mechanics and general science, 3 study mechanics and computer science, and 5 study general science and computer science. One student in the class studies all three subjects. How many of these students study none of the three subjects?

Solution: Let M, G, and C denote the sets of students who study mechanics, general science and computer science respectively and let U be the entire set of 15 students. Then $|M| = 6$, $|G| = 9$, and $|C| = 9$. Also, $|MG| = |M \cap G| = 2$, $|MC| = |M \cap C| = 3$, and $|GC| = |G \cap C| = 6$ and $|MGS| = |M \cap G \cap C| = 1$. Then $(MGC)^c = |U|-(|M|+|G|+|C|-|MG|-|MC|-|GC|+|MGC|)$ = complement of MGC $= 15-(6 + 9 + 9 - 2 - 3 - 6 +1) = 3 = 15 - (24 - 11 + 1) = 1$.

Nature of sets often leads us to many interesting observations. We consider here a nature that distinguishes sets.

Definition: A set is called *finite* if either it is empty or its elements can be exhausted by counting in some definite time period. A set is called *infinite* if it is not finite.

For example, the following sets are finite:

1. ϕ
2. The set of vowels in the English alphabet.
3. $S = [1, 2, 3, ..., 99, 100\}$.
4. The set of students in St. Stephens' College in 2008.
5. The set of Indian nationals residing in USA on 1st Jan. 2008.

The following sets are infinite:

1. The set $\mathbb{N}$ of natural numbers , *i.e.*, $\mathbb{N} = \{1, 2, 3, ...\}$
2. The set $\mathbb{Z}$ of integers, *i.e.*, $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, 3, ...\}$
3. The set $\mathbb{P}$ of prime numbers = $\{2, 3, 5, 7, 11, 13,...\}$

NOTES

4. The set $\mathbb{Q}$ of rational numbers = $\{p/q; q \neq 0, p, q \in \mathbb{Z}\}$

5. The set $\mathbb{I}$ of irrational numbers = $\left\{\sqrt{2}, \sqrt[3]{5}, 3 + 5\sqrt[3]{11}, e, \pi, ...\right\}$

6. The set $\mathbb{R}$ of real numbers, *i.e.*, $\mathbb{Q} \cup \mathbb{I}$.

7. The set $\mathbb{C}$ of complex numbers = $\{a + ib, a, b \in \mathbb{R}, i = \sqrt{-1}\}$

Clearly $\mathbb{P} \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$

Definition: The cardinality of a finite set is defined to be the number of elements in that set. The cardinality of the empty set is defined to be zero. The cardinality of a set A is denoted by card (A) or (A)

A set whose cardinality is one is called a singleton set.

A set whose cardinality is two is called a doubleton set.

A set whose cardinality is three is called a tripleton set.

Thus the set {4} is a singleton set, the set {3,5} is a doubleton set and the cardinality of the set $V = \{a, e, i, o u,\}$ is 5, *i.e.*, card$(V) = 5$.

A result of great importance on cardinality is the following theorem, called cardinality theorem or the theorem of inclusion and exclusion.

Cardinality Theorem: If A and B are two finite sets, then card $(A \cup B)$ = card (A) + card (B) – card $(A \cap B)$

Proof: Let x be the number of elements which belong to A only and let y be the number of elements which belong to B only. If Z be the number of elements which belong to both A and B, then

card $(A) = x + z$ and card $(B) = y + z$.

As card $(A \cap B) = z$, and card $(A \cup B) = x + y + z$, we see

card (A) + card (B) – card $(A \cap B)$

$= (x + z) + (y + z) – z$

$= x + y + z$ = card $(A \cup B)$

Hence the proof.

Corollary 1: If $A \cap B = \phi$, then card $(A \cup B)$ = card (A) + card (B).

Corollary 2: If A, B, C are three finite sets, then

$$\text{card}(A \cup B \cup C) = \text{card}(A) + \text{card}(B) + \text{card}(C) - \text{card}(A \cap B) - \text{card}(B \cap A) - \text{card}(C \cap A) + \text{card}(A \cap B \cap C)$$

Remark: Sometime the notation $n(A)$ is used instead of card (A) to denote cardinality of a finite set.

Example 9.4: If 30 students of a class play football, 20 students play cricket and only 5 of them play both, then find the number of students who play either football or cricket.

NOTES

Solution: Let F denote the set of students playing football and let C denote the set of students playing cricket.

Clearly, card $(F) = 30$, card $(C) = 20$, card $(F \cap C) = 5$

Then the required number of students playing football or cricket

$= \text{card}(F \cup C)$

$= \text{card}(F) + \text{card}(C) - \text{card}(F \cap C)$, by cardinality theorem

$= 30 + 20 - 5 = 45$

Check Your Progress

1. What is meant by inclusion and exclusion principle?
2. When a set is called finite or infinite?
3. Define cardinality of a finite set.

9.3 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. $|A \cap B|$ or $\#(A \cap B)$ or $n(A \cap B)$ is the Principle of Inclusion and Exclusion (written in short as PIE) with the condition that each object is counted only once. We take an arbitrary object, say S out of a set of given N number of objects and assume that the object S has property k out of the many properties in the set of N. To prove that the formulae is correct according to principle of inclusion and exclusion we have to first prove that as per the above given formulae the object S is counted only once.
2. A set is called finite if either it is empty or its elements can be exhausted by counting in some definite time period. A set is called infinite if it is not finite.
3. The cardinality of a finite set is defined to be the number of elements in that set. The cardinality of the empty set is defined to be zero. The cardinality of a set A is denoted by card (A) or (A).

9.4 SUMMARY

- The basic principle of counting the things defines that each object should be counted only once.
- If there are N number of objects given and out of these N objects some objects have property a (denoted as N_a) and some possess property b (denoted as N_b) while some have both the properties a and b (denoted as Nab).
- A set is called finite if either it is empty or its elements can be exhausted by counting in some definite time period. A set is called infinite if it is not finite.

- The cardinality of a finite set is defined to be the number of elements in that set. The cardinality of the empty set is defined to be zero. the cardinality of a set A is denoted by card (A) or (A).
- Sometime the notation $n(A)$ is used instead of card (A) to denote cardinality of a finite set.

9.5 KEY WORDS

- **Finite set:** A set is called finite if either it is empty or its elements can be exhausted by counting in some definite time period.
- **Infinite set:** A set is called infinite if it is not finite.
- **Cardinality:** The cardinality of a finite set is defined to be the number of elements in that set. The cardinality of the empty set is defined to be zero. The cardinality of a set A is denoted by card (A) or (A).

9.6 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. How can we define inclusion and exclusion principles?
2. Differentiate between finite and infinite set.

Long-Answer Questions

1. Discuss briefly the inclusion and exclusion principles giving appropriate examples.
2. Explain the finite and infinite set with the help of suitable examples.
3. Elaborate on the cardinality of a finite set giving an appropriate example.
4. If 25 students of a class play football, 20 students play cricket and only 5 students play both the games, then find the number of students who play either football or cricket.

9.7 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

NOTES

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

UNIT 10 PERMUTATIONS

NOTES

Structure

10.0 Introduction
10.1 Objectives
10.2 Permutations with Forbidden Positions
10.3 The Ménage Problem
10.4 Answers to Check Your Progress Questions
10.5 Summary
10.6 Key Words
10.7 Self Assessment Questions and Exercises
10.8 Further Readings

10.0 INTRODUCTION

Permutations differ from combinations, which are selections of some members of a set regardless of order. For example, written as tuples, there are six permutations of the set {1,2,3}, namely: (1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), and (3,2,1). These are all the possible orderings of this three-element set. Anagrams of words whose letters are different are also permutations: the letters are already ordered in the original word, and the anagram is a reordering of the letters. The study of permutations of finite sets is an important topic in the fields of combinatorics and group theory.

The number of permutations of *n* distinct objects is *n* factorial, usually written as *n*!, which means the product of all positive integers less than or equal to *n*. Technically, a permutation of a set *S* is defined as a bijection from *S* to itself. That is, it is a function from *S* to *S* for which every element occurs exactly once as an image value. This is related to the rearrangement of the elements of *S* in which each element *s* is replaced by the corresponding *f*(*s*). For example, the permutation (3, 1, 2) mentioned above is described by the function α defined as:

$$\alpha(1) = 3, \quad \alpha(2) = 1, \quad \alpha(3) = 2.$$

The collection of all permutations of a set form a group called the symmetric group of the set. The group operation is the composition (performing two given rearrangements in succession), which results in another rearrangement. As properties of permutations do not depend on the nature of the set elements, it is often the permutations of the set $\{1, 2, \ldots, n\}$ that are considered for studying permutations. In mathematics, a permutation of a set is, loosely speaking, an arrangement of its members into a sequence or linear order, or if the set is already ordered, a rearrangement of its elements. The word "Permutation" also refers to the act or process of changing the linear order of an ordered set.

NOTES

In elementary combinatorics, the k-permutations, or partial permutations, are the ordered arrangements of k distinct elements selected from a set. When k is equal to the size of the set, these are the permutations of the set.

In this unit, you will study about the permutations with forbidden positions, and the ménage problem.

10.1 OBJECTIVES

After going through this unit, you will be able to:

- Interpret the permutations with forbidden positions
- Elaborate on the ménage problem

10.2 PERMUTATIONS WITH FORBIDDEN POSITIONS

A permutation of the set $S = \{1, ..., n\}$ where each element of S is restricted to not being in certain positions (here the permutation is considered as an ordering of the elements of S) is called a permutation with forbidden positions. For example, with $S = \{1,2,3,4\}$, the permutations with the restriction that the element 1 cannot be in positions 1 or 3, and the element 2 cannot be in position 4 are: 2134, 2143, 3124, 4123, 2341, 2431, 3241, 3421, 4231 and 4321.

In the given example, there are $12 = 2(3!)$ permutations with property P_1, $6 = 3!$ permutations with property P_2 and no permutations have properties P_3 or P_4 as there are no restrictions for these two elements. The number of permutations satisfying the restrictions is thus:

$$4! - (12 + 6 + 0 + 0) + (4) = 24 - 18 + 4 = 10.$$

By letting A_i be the set of positions that the element i is not allowed to be in, and the property P_i to be the property that a permutation puts element i into a position in A_i, the principle of inclusion–exclusion can be used to count the number of permutations which satisfy all the restrictions.

The final 4 in this computation is the number of permutations having both properties P_1 and P_2. There are no other non-zero contributions to the formula.

Theorem 10.1: For $n \geq 1$,

$$Q_n = \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} (n-k)!.$$

Proof: Let S be the set of permutations of $\{1, 2, \ldots, n\}$. Let P_i be the property that in a permutation the pattern $i(i+1)$ does occur, where $1 \leq i \leq n-1$. Let A_i be the set of all permutations satisfying the property P_i. Then Q_n is the number of permutations satisfying none of the properties $P_1, \ldots, P_{n-1}$, i.e.,

$$Q_n = |\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_{n-1}|.$$

Now, $|A_i| = (n-1)!, \quad 1 \le i \le n-1.$

Similarly,

$$|A_i \cap A_j| = (n-2)!, \quad 1 \le i < j \le n-1.$$

More generally,

$$|A_{i_1} \cap \cdots \cap A_{i_k}| = (n-k)!, \quad 1 \le i_1 < \cdots < i_k \le n-1.$$

Thus by the Inclusion-Exclusion Principle,

$$\begin{aligned} Q_n &= |S| + \sum_{k=1}^{n-1} (-1)^k \sum_{1 \le i_1 < \cdots < i_k \le n-1} |A_{i_1} \cap \cdots \cap A_{i_k}| \\ &= \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} (n-k)!. \end{aligned}$$

We assign numbers 1, 2, . . . , 8 to the eight persons so that the number *i* is assigned to the *i*th person (counted from the front). The problem is then to find the number of permutations of {1, 2, . . . , 8} in which the patterns 12, 23, . . ., 78 do not occur. For instance, 31542876 is an allowed permutation, while 83475126 is not. Then the answer will be

$$P = \frac{Q_8}{8!} = \sum_{k=0}^{7} (-1)^k \binom{7}{k} \frac{(8-k)!}{8!} \approx 0.413864.$$

10.3 THE MÉNAGE PROBLEM

In combinatorial mathematics, the ménage problem asks for the number of different ways in which it is possible to seat a set of male-female couples at a round dining table so that men and women alternate and nobody sits next to his or her partner. This problem was formulated in 1891 by Édouard Lucas and independently, a few years earlier, by Peter Guthrie Tait in connection with knot theory. For a number of couples equal to 3, 4, 5, ... the number of seating arrangements is 12, 96, 3120, 115200, 5836320, 382072320, 31488549120, ...

Mathematicians have developed formulas and recurrence equations for computing these numbers and related sequences of numbers. Along with their applications to etiquette and knot theory, these numbers also have a graph theoretic interpretation: they count the numbers of matching and Hamiltonian cycles in certain families of graphs.

There are 2×*n*! ways of seating the women: there are two sets of seats that can be arranged for the women, and there are *n*! ways of seating them at a particular set of seats. For each seating arrangement for the women, there are

NOTES

$$A_n = \sum_{k=0}^{n} (-1)^k \frac{2n}{2n-k} \binom{2n-k}{k} (n-k)!$$

NOTES

Ways of seating the men; this formula simply omits the $2\times n!$ factor from Touchard's formula. The resulting smaller numbers (again, starting from $n = 3$), 1, 2, 13, 80, 579, 4738, 43387, 439792, ...

These are called the ménage numbers.

The ménage numbers satisfy the recurrence relation

$$A_n = nA_{n-1} + \frac{n}{n-2} A_{n-2} + \frac{4(-1)^{n-1}}{n-2}$$

And the simpler four-term recurrence

$$A_n = nA_{n-1} + 2A_{n-2} - (n-4)A_{n-3} - A_{n-4},$$

From which the ménage numbers themselves can easily be calculated.

Theorem 10.2: The number of ways to seat n couples, alternating men and women, so that no one is sitting next to their spouse is $2n!u_n$.

Proof: Let us start by seating the men. They can occupy either the odd or the even numbered seats and they can be arranged in $n!$ ways. The women can now be seated in un ways, so as to avoid sitting next to their spouses. So we have a solution to the problem if we can determine a value for u_n.

In order to find u_n, we appeal to the principle of inclusion/exclusion. The number of ways to arrange the women without any stipulations is $n!$. From this we need to remove all the permutations that meet any of the conditions. Let A_i be the set of all permutations satisfying condition i given above. Then using inclusion/exclusion

$$\begin{aligned} u_n &= n! - |A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_{2n}| \\ &= n! - [|A_1| + |A_2| + \ldots + |A_{2n}| - |A_1 \cap A_2| - |A_1 \cap A_3| - \\ &\quad |A_2 \cap A_3| - \ldots + |A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| \cdots] \end{aligned}$$

This expression can be reduced considerably. For instance, the cardinality of the set $A_1 \cap A_2$ is zero since woman number 1 cannot be in both the first and second seat. In general, if any two consecutive conditions are chosen thecardinality of the corresponding set will be zero.

The number of ways to pick two non-consecutive conditions from the $2n$ conditions is

$$\binom{2n-2}{2} \frac{2n}{(2n-2)}.$$

Therefore, the solution to the ménage problem can be written as

$$2n! \sum_{i=0}^{n} (-1)^i \frac{2n}{2n-i} \binom{2n-i}{i} (n-i)!.$$

However, as stated above, this expression over counts. To eliminate cyclic equivalence, we must divide by $2n$ which results in

$$(n-1)!\sum_{i=0}^{n}(-1)^i\frac{2n}{2n-i}\binom{2n-i}{i}(n-i)!.$$

As the number of couples increases, the solution to the ménage problem grows rapidly. Table 10.1 show the solution to the ménage problem for small n.

Table 10.1 *Solutions to the ménage problem*

n	2	3	4	5	6	7
$(n-1)!u_n$	0	2	12	312	9600	416880

Check Your Progress

1. Define the permutations with forbidden positions.
2. Explain the ménage problem.

10.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. A permutation of the set $S = \{1, ..., n\}$ where each element of S is restricted to not being in certain positions (here the permutation is considered as an ordering of the elements of S) is called a permutation with forbidden positions.
2. The ménage problem asks for the number of different ways in which it is possible to seat a set of male-female couples at a round dining table so that men and women alternate and nobody sits next to his or her partner. This problem was formulated in 1891 by Édouard Lucas and independently, a few years earlier, by Peter Guthrie Tait in connection with knot theory.

10.5 SUMMARY

- A permutation of the set $S = \{1, ..., n\}$ where each element of S is restricted to not being in certain positions (here the permutation is considered as an ordering of the elements of S) is called a permutation with forbidden positions.
- By letting A_i be the set of positions that the element i is not allowed to be in, and the property P_i to be the property that a permutation puts element i into a position in A_i, the principle of inclusion–exclusion can be used to count the number of permutations which satisfy all the restrictions.
- In combinatorial mathematics, the ménage problem asks for the number of different ways in which it is possible to seat a set of male-female couples at a round dining table so that men and women alternate and nobody sits next to his or her partner.

NOTES

NOTES

- Mathematicians have developed formulas and recurrence equations for computing these numbers and related sequences of numbers. Along with their applications to etiquette and knot theory, these numbers also have a graph theoretic interpretation: they count the numbers of matching and Hamiltonian cycles in certain families of graphs.
- The number of ways to seat n couples, alternating men and women, so that no one is sitting next to their spouse is $2n!u_n$.

10.6 KEY WORDS

- **Permutations with forbidden positions:** A permutation of the set $S = \{1, ..., n\}$ where each element of S is restricted to not being in certain positions (here the permutation is considered as an ordering of the elements of S) is called a permutation with forbidden positions.
- **Ménage problems:** The ménage problem asks for the number of different ways in which it is possible to seat a set of male-female couples at a round dining table so that men and women alternate and nobody sits next to his or her partner.

10.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Explain the permutations with forbidden positions.
2. State the ménage problem.

Long-Answer Questions

1. Briefly discuss the permutations with forbidden positions.
2. Analyse the ménage problem with the help of example.

10.8 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

NOTES

NOTES

UNIT 11 PROBLEM OF FIBONACCI AND POLYA'S THEORY

Structure

11.0 Introduction
11.1 Objectives
11.2 Problems of Fibonacci
11.3 Polya Theory
11.4 Answers to Check Your Progress Questions
11.5 Summary
11.6 Key Words
11.7 Self Assessment Questions and Exercises
11.8 Further Readings

11.0 INTRODUCTION

In mathematics, Fibonacci is a well-known mathematical figure. Leonardo Pisano, the mathematician who bears his name, will be surprised by this. And he may have been equally shocked that he was immortalized in the famous sequence – 0, 1, 1, 2, 3, 5, 8, 13, ... rather than for what is considered his much greater mathematical achievement assisting in the popularization of our modern number systems.

The Pólya enumeration theorem, also known as the Redfield–Pólya theorem and Pólya counting, is a theorem in combinatorics that both follows from and ultimately generalizes Burnside's lemma on the number of orbits of a group action on a set. The theorem was first published by J. Howard Redfield in 1927. In 1937 it was independently rediscovered by George Pólya, who then greatly popularized the result by applying it to many counting problems, in particular to the enumeration of chemical compounds.

Finding the number of different ways to arrange a necklace with n different coloured beads is a popular mathematical puzzle. Nonetheless $(n-1)!$ as well as (n-1)!/2 Since the query does not specify what it means for necklaces to be distinct, 2 and 3 are both correct responses. The former counts the number of distinct necklaces that can be rotated, while the latter counts the number of distinct necklaces that can be rotated and reflected. When we consider "distinctness" up to arbitrary transformations, objects with more components, and non-standard symmetries, questions like these become more complicated. The quest for a general response leads us to concepts in group theory and symmetry, which leads us to Polya's enumeration.

In this unit, you will study the basic concept of problems of Fibonacci and Polya theory.

11.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Understand the basic concept of problems of Fibonacci
- Analyse the concept of Polya theory

11.2 PROBLEMS OF FIBONACCI

The numbers in the sequence 0, 1, 2, 3, 5, 8, 13, 21,... in which each new term is the sum of the previous two terms are called the Fibonacci numbers. If we denote the $(n+1)$th Fibonacci number by f_n, we have

$$f_n = f_{n-1} + f_{n-2} \text{ for } n \geq 2$$

and $f_0 = 0$ and $f_1 = 1$. This is called a recursive definition, in which each element of the sequence is defined in terms of previous numbers in the sequence.

We can define a function with the set of non-negative integers as its domain by

(a) Specifying the value of the function at zero, and

(b) Giving a rule for finding its value at an integer from its values at smaller integers.

Such a function is called a recursively defined function or such a definition is called recursive definition. For example, consider the sequence of powers of 3 given by

$$a_n = 3^n \text{ for } n = 0,1,2...$$

This sequence can also be defined by giving the first term of the sequence, namely $a_0 = 1$ and a rule for finding a term of the sequence from the previous one, namely

$$a_{n+1} = 3a_n \text{ for } n = 0,1,2,....$$

Example 11.1: Find a recursive definition of binomial coefficients.

Solution: Denote the binomial coefficient ${}_nC_K$ by $C(n,k)$. Then the recursive definition for $C(n,k)$ where $n \geq 0$, k ≥ 0 and $n \geq k$ is given by

$$C(n,0) = 1;$$

$$C(n,n) = 1; \text{ and}$$

$$C(n,k) = C(n-1, k) + C(n-1, k-1) \text{ for } n > k > 0.$$

Example 11.2: Suppose that f is defined recursively by

$$f(0) = 2$$

$$f(n+1) = 3f(n)+2$$

NOTES

Find $f(1), f(2), f(3)$ and $f(4)$.

Solution: From the recursive definition, it follows that

$$f(1) = 3f(0) + 2 = 3.2 + 2 = 8$$
$$f(2) = 3f(1) + 2 = 3.8 + 2 = 26$$
$$f(3) = 3f(2) + 2 = 3.26 + 2 = 80$$
$$f(4) = 3f(3) + 2 = 3.80 + 2 = 242$$

Example 11.3: Give a recursive definition of $f(n) = n!$.

Solution: Since $0! = 1$ and $(n+1)! = (n+1)n!$ the desired rule is

$$f(0) = 1; \text{ and}$$
$$f(n+1) = (n+1)f(n)$$

Example 11.4: Give a recursive definition of a^n where a is a non zero real number and n is non-negative integer.

Solution: It is easy to see that

$$a^0 = 1; \text{ and } \quad a^{n+1} = a.a^n \text{ for } n \geq 0$$

uniquely define a^n for all non-negative integers n.

Example 11.5: Give a recursive definition for polynomial expression.

Solution: An expression of the form

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$$

where a's are constants ($a_0 \neq 0$) and $n \geq 0$, is called polynomial in x of degree n.

Let S be a set of coefficients. Then recursive definition for polynomial is

(a) A zeroth degree (constant) polynomial is an element of S.

(b) For $n \geq 1$, an nth degree a polynomial expression is an expression of the form $q(x)x+a$ where $q(x)$ is an $(n-1)$th degree polynomial and $a \in S$.

Example 11.6: Using recursive definition for polynomial expression, prove that $f(n) = 3n^3 - 8n^2 + 2n + 4$ is a third-degree polynomial expression.

Solution: Given $f(n) = 3n^3 - 8n^2 + 2n + 4$

$$= (3n^2 - 8n + 2)\, n + 4 = ((3n - 8)\, n - 8)\, n + 4$$
$$= (((3)\, n - 8)\, n + 2)n + 4.$$

Now, 3 is a zeroth degree polynomial

$\Rightarrow (3)n - 8$ is a first degree polynomial

$\Rightarrow ((3)n - 8)\, n + 2$ is a second degree polynomial

$\Rightarrow f(n) = (((3)n - 8)\, n + 2)\, n + 4$ is a third degree polynomial.

NOTES

Note: The final expression for

$$f(n) = (((3)n - 8)\ n + 2)\ n + 4$$

is called its telescoping form. If we use it to calculate $f(6)$, we need only three multiplications and three additions or subtractions. This is called Horner's method for evaluating a polynomial expression.

Example 11.7: Write $p(x) = x^3 - 6x^2 + 11x - 6$ in telescoping form.

Solution: The telescoping form of

$$p(x) = x^3 - 6x^2 + 11x - 6$$
$$= (((1)x - 6)\ x + 11)\ x - 6$$

Example 11.8: Write $g(x) = -x^5 + 3x^4 + 3x^3 + 2x^2 + x$ in telescoping form.

Solution: The telescoping form of

$$g(x) = x^5 + 3x^4 + 3x^3 + 2x^2 + x$$
$$= (((((-1)\ x + 3)\ x + 3)\ x + 2)\ x + 1)\ x + 0.$$

Recursion and Iteration

There is another way to evaluate a function from its recursive definition. Instead of successively reducing the computation to evaluate of the function at small integers, we can start with the basis and successively apply the recursive definition to find the values of the function at successive larger integers. Such a procedure is called iterative procedure.

To find $n!$ using an iterative procedure, we start with 1, the value of the factorial function at 0, and multiply successively by each positive integer less than or equal to n.

Note: Often an iterative approach for the evaluation of a recursively defined sequence requires much less computation than a procedure using recursion. But it is sometimes preferable to use a recursive procedure even if it is less efficient than the iterative procedure.

Closed form Expression

Let $f(x_1, x_2, ..., x_n)$ be an algebraic expression involving variables. $x_1, x_2, ..., x_n$ which are allowed to take values from some predetermined set. f is a closed form expression if there exists a number F such that evaluation of f with any allowed values $x_1, x_2 ..., x_n$ will take no more than F operations.

For example, (i) A closed form expression for

$$1 + 2 + ... + n = \frac{n(n+1)}{2}$$

(ii) A closed form expression for

$$1^2 + 2^2 + ... + n^2 = \frac{n(n+1)(2n+1)}{2}.$$

NOTES

Sequence of Integers

A sequence of integers is a function from the natural numbers into integers. That is, if f is a sequence of integers, then $f: N \rightarrow Z$ is a function. We use the notation f_n or $f(n)$ to denote the image of any natural n. We call f_n an nth term of the sequence.

Note: A sequence is often called a discrete function.

Example 11.9: Prove by induction that $f(k) = 3k + 2$, $k \geq 0$ is a closed form expression of the sequence f defined recursively by $f(0) = 2$ and $f(k) = f(k-1) + 3$ for $k \geq 1$.

Solution:

1. Basis of induction: If $k = 0$, then $f(0) = 3(0) + 2 = 2$, as defined.
2. Induction step: Now assume that for some $k > 0$,

$$f(k) = 3k+2 \qquad (1)$$

we have,

$$f(k+1) = f(k)+3 \text{ by the recursive definition}$$
$$= 3k+2+3 = 3k + 3 + 2$$
$$= 3(k + 1) + 2$$

Therefore, by the principle of mathematical induction the result follows. Hence $f(k) = 3k + 2$, $k \geq 0$ is a closed form expression for the sequence f.

Example 11.10: Define the sequence of numbers f by $f_0 = 100$ and $f_n = 1.08\, f_{n-1}$ for, $n \geq 1$. Prove by induction that $f_n = 100\,(1.08)^n$, $n \geq 0$ is a closed form of expression for the sequence f.

Solution:

1. Basis of induction: If $k = 0$, then $f_0 = 100(1.08)^0 = 100$, as defined.
2. Induction step: Now assume that for some $k > 0$

$$f_k = 100(1.08)^k$$

Then

$$f_{(k+1)} = (1.08) f_k \text{ by the recursive definition}$$
$$= (1.08)100^k$$
$$= 100(1.08)^{k+1} \qquad \text{by Equation (1)}$$

Therefore, by the principle of mathematical induction the result follows. Hence $f_n = 100(1.08)^n$, $n \geq 0$ is a closed form expression for the sequence f.

Notes:

1. The sequence $f: N \rightarrow Z$ defined by $f(n) = n^3 - n$, is a sequence of integers.
2. The sequence $f: N \rightarrow Z$ defined by $f(0) = 1$ and $f(n) = f(n-1) + 2$ for $n \geq 1$ is a sequence of integers.

3. The codomain of a sequence can be any set. We use the notation $\{f_n\}$ to describe the sequence. We describe sequences by listing the terms of the sequence in the order of increasing subscripts.

For example,

(i) Consider the sequence $f: N \to S$ defined by $f(n) = f_n = \frac{1}{n+1}$. The list of the terms of this sequence $f_0, f_1, f_2, \ldots$ begins with

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$$

(ii) Consider the sequence $f: N \to Z$ defined by $f_n = 4^n$. The list of the terms of the sequence $f_0, f_1, f_2, \ldots$ begins with

1, 4, 16, 64,...

Recurrence Relations

The expressions for permutations and combinations are one of the most fundamental tools for counting the elements of finite sets. They often prove to be inadequate and many problems of computer science require a different approach. An important alternate approach uses recurrence relations (often called recurrence equations or difference equation) to define the terms of a sequence. A formal definition of recurrence relations is difficult because of the wide variety of forms in which such relations can be written, but the concept is straight forward. We have already seen an example of a recurrence relation in the definition of the Fibonacci sequence, where for $n \geq 2$, the term f_n is defined by the recurrence relation

$$f_n = f_{n-1} + f_{n-2}$$

The salient characteristic of a recurence relation is the specification of the term f_n as a function of the terms $f_0, f_1, \ldots, f_{n-1}$. By itself, however, a recurrence relation is not sufficient to define the terms of a sequence; we must also specify the values of some initial terms of the sequence. Thus in our definition of the Fibonacci sequence, we set $f_0 = 0$ and $f_1 = 1$. These are called the boundary conditions or initial conditions of the sequence.

Recall that a recursive definition of a sequence specifies one or more initial terms and a rule for determining subsequent terms from those that precede them. Recursive definitions can be used to solve counting problems. When they are, the rule for finding terms from those that precede them is called a recurrence relation.

Recurrence relation: A recurrence relation for the sequence $\{f_n\}$ is a formula that express f_n in terms of one or more of the previous terms of the sequence, namely, $f_0, f_1, \ldots, f_{n-1}$, for all integers n with $n \geq n_0$, when n_0 is a non-negative integer. A sequence is called a solution of a recurrence relation if its terms satisfy the recurrence relation.

NOTES

Example 11.11: Determine whether the sequence $\{f_n\}$ is a solution of the recurrence relation

$$f_n = 2f_{n-1} - f_{n-2}$$

for $n = 2, 3, 4,...$ where $f_n = 3n$ for every non-negative integer n.

Solution: Suppose that $f_n = 3n$ for every non-negative integer n, Then, for $n \geq 2$,

$$\begin{aligned} f_n &= 2f_{n-1} - f_{n-2} \\ &= 2[3(n-1)] - 3(n-2) \text{ since } f_n = 3n \\ &= 6n - 6 - 3n + 6 = 3n \end{aligned}$$

Therefore $\{f_n\}$, where $f_n = 3n$, is a solution of the recurrence relation.

Example 11.12: Show that the sequence $\{f_n\}$ is a solution of the recurrence relation $f_n = -3f_{n-1} + 4f_{n-2}$ if $f_n = 2(-4)^n + 3$.

Solution: Suppose that $f_n = 2(-4)^n + 3$. Then

$$\begin{aligned} f_n &= -3f_{n-1} + 4f_{n-2} \\ &= -3\,[2(-4)^{n-1} + 3] + 4\,[2(-4)^{n-2} + 3] \\ &= -6(-4)^{n-1} - 9 + 8(-4)^{n-2} + 12 \\ &= -6(-4)^{n-1} - 2(-4)^{n-1} + 3 \\ &= -8(-4)^{n-1} + 3 \\ &= 2(-4)^n + 3 \end{aligned}$$

Therefore, $\{f_n\}$, where $f_n = 2(-4)^n + 3$, is a solution of the recurrence relation.

In the following, we shall study a class of recurrence relations known as linear recurrence relations with constant coefficients.

Linear Recurrence Relation: A recurrence relation of the form

$$a_0 f_n + a_1 f_{n-1} + a_2 f_{n-2} + ... + a_k f_{n-k} = f(n) \qquad (1)$$

Where a_1′s are constants and is called a linear recurrence relation with constant coefficients. The recurrence relation (1) is known as a kth-order recurrence relation, provided that both a_0 and a_k are non zero.

Note: The phrase "kth-order" means that each term in the sequence depends only on the k previous terms.

For example, Consider the Fibonacci sequence defined by the recurrence relation $f_n = f_{n-1} + f_{n-2}$, $n \geq 2$ and the initial condition $f_0 = 0$ and $f_1 = 1$. The recurrence relation is called a second-order relation because f_n depends on the two previous terms of f.

For example, Consider the recurrence relation $f(k) - 5f(k-1) + 6f(k-2) = 4k + 10$ defined for $k \geq 2$, together with the initial condition $f(0) = \frac{7}{3}$ and $f(1) = 5$. Clearly, it is a second-order linear recurrence relation.

Homogeneous Recurrence: A kth-order linear relation is a homogeneous recurrence relation if $f(n) = 0$ for all n. Otherwise, it is called non-homogeneous.

For example, Consider the recurrence relation $c(k) - 5c(k-1) + 8c(k-2) = 0$ together with the initial condition $c(0) = 5$ and $c(1) = 2$. It is a second-order homogeneous recurrence relation.

Example 11.13: Which of the recurrence relations of the following are homogenous and which are non- homogenous?

(i) $f_n = nf_{n-2}$

(ii) $a_n = a_{n-1} + a_{n-3}$

(iii) $b_n = b_{n-1} + 2$

(iv) $s(n-2) + s(n-4)$

Solution: The relation $f_n = nf_{n-2}$, $a_n = a_{n-1} + a_{n-3}$, $s(n) = s(n-2) + s(n-4)$ are all homogeneous and the relation $b_n = b_{n-1} + 2$ is non-homogeneous.

Recurrence Relations Obtained from Solutions

Before giving an algorithm for solving a recurrence relations, we will examine a few recurrence relations that arise from certain closed form expressions. The procedure is illustrated by the following examples.

Example 11.14: Form the recurrence relation given $f_n = 3.5^n, n \geq 0$.

Solution: If $n \geq 1$, then

$$f_n = 3.5^n = 3.5.5^{n-1}$$
$$= 5.3.5^{n-1} = 5.f_{n-1}.$$

So the recurrence relation is $f_n - 5f_{n-1} = 0$ with $f_0 = 3$.

Example 11.15: Find the recurrence relation, satisfying $y_n = A(3)^n + B(-2)^n$.

Solution: Given $y_n = A(3)^n + B(-2)^n$

Therefore, $y_{n+1} = A(3)^{n+1} + B(-2)^{n+1} = 3A(3)^n - 2B(-2)^n$

and $y_{n+2} = A(3)^{n+2} + B(-2)^{n+2} = 9A(3)^n + 4B(-2)^n$.

Eliminating A and B from these equation, we get

$$\begin{vmatrix} y_n & 1 & 1 \\ y_{n+1} & 3 & -2 \\ y_{n+2} & 9 & 4 \end{vmatrix} = 0$$

(or) $y_{n+2} - y_{n+1} - 6y_n = 0$ which is the required recurrence relation.

Example 11.16: Find the recurrence relation satifying $y_n = A(3)^n + B(-4)^n$.

Solution: Given $y_{n+1} = A(3)^n + B(-4)^n$

Therefore, $y_{n+1} = 3A(3)^n - 4B(-4)^n$

NOTES

and $y_{n+2} = 9A(3)^n + 16B(-4)^n$. Eliminating A and B from these equation, we get

$$\begin{vmatrix} y_n & 1 & 1 \\ y_{n+1} & 3 & -4 \\ y_{n+2} & 9 & 16 \end{vmatrix} = 0$$

(or) $y_{n+2} + y_{n+1} - 12y_n = 0$ which is the required recurrence relation.

Example 11.17: Find the recurrence relation satisfing $y_n = (A+Bn)4^n$

Solution: Given $y_n = (A + Bn)4^n$

$= A4^n + nB4^n$

Therefore, $y_{n+1} = 4A4^n + 4(n+1)B4^n$

and $y_{n+2} = 16A4^n + 16(n+2)B4^n$.

Eliminating A and B from these equation, we get

$$\begin{vmatrix} y_n & 1 & n \\ y_{n+1} & 4 & 4(n+1) \\ y_{n+2} & 16 & 16(n+2) \end{vmatrix} = 0$$

(or) $y_{n+2} - 8y_{n+1} - 16y_n = 0$, which is the required recurrence relation.

Solving Linear Homogeneous Recurrence Relations

Consider a linear homogeneous recurrence relations of degree k with constant coefficients

$$f_n = a_1 f_{n-1} + a_2 f_{n-2} + \ldots + a_k f_{n-k}$$

Where $a_1, a_2, \ldots, a_k$ are real numbers and $ak \neq 0$. The basic approach for solving linear homogeneous recurrence relations is to look for solutions of the form $f_n = r^n$, where r is a constant. Note that $f_n = r^n$ is solution of the recurrence relation $f_n = a_1 f_{n-1} + a_2 f_{n-2} + \ldots + a_k f_{n-k}$ if and only if

$$r^n = C_1 r^{n-1} + C_2 r^{n-2} + \ldots + C_k r^{n-k}$$

when both sides of this equation are divided by r^{n-k} and the right-hand side is substracted from the left, we obtain

$$r^k - C_1 r^{k-1} - C_2 r^{k-2} - \ldots - C_{k-1} T - C_k = 0$$

consequently, the sequence $\{f_n\}$ with $f_n = r^n$ is a solution if and only if r is a solution of this last equation,

Characteristic Equation: The characteristic equation of the homogeneous k^{th} order linear relation

$$f_n + a_1 f_{n-1} + a_2 f_{n-2} + \ldots + a_k f_{n-k} = 0$$

is the k^{th} degree polynomial equation

$$r^k + C_1 r^{k-1} + C_2 r^{k-2} + \cdots + C_{k-1} r + C_k = 0$$

The solutions of this equation are called the characteristic roots of the recurrence relation.

Example 11.18: What is the characteristic equation of

$$Q(k) + 2Q(k-1) - 3Q(k-2) - 6Q(k-4) = 0$$

Solution: The characteristic equation of

$$Q(k) + 2Q(k-1) - 3Q(k-2) - 6Q(k-4) = 0$$

is $r^4 + 2r^3 - 3r^2 - 6 = 0$.

Note that the absence of a $Q(k-3)$ term means that there is no $r^{4-3} = r$ term in the characteristic equation.

Example 11.19: What is the characteristic equation of $T(k) - 7T(k-2) + 6T(k-3) = 0$

The characteristic equation is $r^3 - 7r + 6 = 0$ and 1,2 and -3 are the characteristic roots.

Algorithm for Solving *k*th Order Homogeneous Linear Recurrence Relation.

Step 1. If $f_n + a_1 f_{n-1} + a_2 f_{n-2} + ... + a_k f_{n-k} = 0$ is a given recurrence relation, then write its characteristic equation.

That is, $r^k + C_1 r^{k-1} + C_2 r^{k-2} + ... + C_{k-1} r + C_k = 0$

Step 2. Find all the characteristic roots of this equation.

Step 3. *Case* (i) If there are k distinct roots, say $c_1, c_2, ..., c_k$ then the general solution of the recurrence relation is

$$f_n = A_1 c_1^k + A_2 c_r^k + ... + A_k c^k{}_k$$

Case (ii) Suppose that c_1 is a root of multiplicity m. Then the corresponding solution is

$$f_n = (A_1 r^{m-1} + A_2 r^{m-2} + ... + A_{m-2} r^2 + A_{m-1}{}^T + A_m) c^T{}_1 + ...$$

Step 4. Use the boundary conditions to determine the constants $A_1, A_2, ..., A_k$.

Example 11.20: Solve the Fibonacci sequence $\{f_n\}$ defined by $f_n = f_{n-1} + f_{n-2}$ for $n \geq 2$ with the initial conditions $f_0 = 1$ and $f_1 = 1$.

Solution: The first step is to form the characteristic equation corresponding to the given difference equation. In this case it is

$$r^2 - r - 1 = 0$$

Solving, we get $c_1 = \frac{1+\sqrt{5}}{2}$, and $c_2 = \frac{1-\sqrt{5}}{2}$ the characteristic roots. It follows that the general solution is

$$f_n = A_1 c_1^m + A_2 c_2^m$$

where A_1 and A_2 are constants. Since $f_0 = 1$ and $f_1 = 1$, we get

NOTES

NOTES

$$0 = A_1 + A_2$$

and $$1 = A_1 c_1 + A_2 c_2$$

$$1 = A_1\left(\frac{1+\sqrt{5}}{2}\right) + A_2\left(\frac{1-\sqrt{5}}{2}\right)$$

On solving, we get

$$A_1 = \frac{1}{\sqrt{5}} \quad \text{and} \quad A_2 = \frac{-1}{\sqrt{5}}$$

Hence the solution is

$$f_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

Example 11.21: If the recurrence relation is $u_{n+1} - 2u_n = 0$, find the closed form expression (solution) for u_n.

Solution: The characteristic equation of the given recurrence relation is $r-2 = 0$, i.e., $r = 2$. Therefore, the general solution is $u_n = A2^n$. Hence $u_n = u_0 \cdot 2^n$ is the closed form expression, where the value of u_0 is the initial condition.

Example 11.22: Find $f(n)$ if $f(n) = 7f(n-1) - 10f(n-2)$, given that $f(0) = 4$ and $f(1) = 17$.

Solution: The characteristic equation of the given recurrence relation is $r^2 - 7r + 10 = 0$

It's characteristic roots are $r = 2, 5$. So the general solution of the recurrence relation is $f(n) = A_1(2)^n A_2(5)^n$.

Since $f(0) = 4$, $4 = A_1 + A_2$

Again, $f(1) = 17$ implies $17 = 2A_1 + 5A_2$

Solving, we get $A_1 = 1$ and $A_2 = 3$. Therefore, $f(n) = (2)^n + 3(5)^n$

Example 11.23: Find $T(k)$ if $T(k) - 7T(k-2) + 6T(k-3) = 0$, where $T(0) = 8$, $T(1) = 6$ and $T(2) = 22$.

Solution: The characteristic equation is $r^3 - 7r + 6 = 0$

Its roots are 1,2 and –3. Therefore, the general solution is

$$T(k) = A_1(1)^k + A_2(2)^k + A_3(-3)^k.$$

Now

$$T(0) = 8 \Rightarrow A_1 + A_2 + A_3 = 8.$$
$$T(1) = 6 \Rightarrow A_1 + 2A_2 - 3A_3 = 6.$$
$$T(2) = 22 \Rightarrow A_1 + 4A_2 + 9A_3 = 22.$$

Solving, we get $A_1 = 5$, $A_2 = 2$ and $A_3 = 1$.

Hence,

$$T(k) = 5 + 2(2)^k + 1(-3)^k$$
$$= 5 + 2^{k+1} + (-3)^k$$

Example 11.24: Solve $f_k - 8f_{k-1} + 16f_{k-2} = 0$ where $f_2 = 16$ and $f_3 = 80$.

Solution: The characteristic equation is $r^2 - 8r + 16 = 0$ (or) $(r-4)^2 = 0$.

So $r = 4$ is a double characteristic root. Therefore, the general solution is $f_k = (A_1 + A_2 k)4^k$.

Now, $f_2 = 16 \Rightarrow (A_1 + 2A_2)16 = 16$

and $f_3 = 80 \Rightarrow (A_1 + 3A_2)64 = 80$

Solving, we get $A_1 = \frac{1}{2}$ and $A_2 = \frac{1}{4}$.

Hence the solution is

$$f_k = \left(\frac{1}{2} + \frac{1}{4}k\right)4k$$
$$= (2 + k)4^{k-1}$$

Example 11.25: Find a solution to the recurrence relation

$$C_n = -3C_{n-1} - 3C_{n-2} - C_{n-3}$$

for $n \geq 3$ with initial conditions $C_0 = 1$, $C_1 = -2$ and $C_2 = 1$.

Solution: The characteristic equation is

$$r^3 + 3r^2 + 3r + 1 = 0 \quad \text{(or)} \quad (r+1)^3 = 0.$$

So $r = -1$ is a characteristic root of multiplicity 3. Therefore, the general solution is

$$C_n = (A_1 + A_2 n + A_3 n^2)(-1)^n$$

Now,

$$C_0 = 1 \Rightarrow A_1 = 1$$
$$C_1 = -2 \Rightarrow -(A_1 + A_2 + A_3) = -2$$
$$C_2 = 1 \Rightarrow A_1 + 2A_2 + 4A_3 = 1$$

Solving, we get $A_1 = 1$, $A_2 = 2$, $A_3 = 1$. Hence the solution is $C_n = (1 + 2n - n^2)(-1)^n$.

Solving Linear Non Homogeneous Recurrence Relations

The solution of a linear non homogeneous recurrence relation with constant coefficient is the sum of the two parts, the homogeneous solution, which satisfies the recurrence relation when the right hand side of the equation is set 1to 0, and the particular solution, which satisfies the difference equation with $f(n)$ on the right hand side.

NOTES

There is no general procedure for determining the particular solution of a difference equation. However, in simple cases, this solution can be obtained by the method of inspection. To determine the particular solution we use the following rules.

***Rule* 1**. When $f(n)$ is of the form of a polynomial of degree m in n

$$k_0 + k_1 n + k_2 n^2 + \ldots + k_{m-1} n^{m-1} + k_m n^m$$

the corresponding particular solution will be of the form

$$Q_0 + Q_1 n + Q_2 n^2 + \ldots + Q_{m-1} n^{m-1} + Q_m n^m$$

***Rule* 2**. When $f(n)$ is of the form

$$(k_0 + k_1 n + k_2 n^2 + \ldots + k_{m+1} n^{m-1} + k_m n^m) a^n$$

the corresponding particular solution is of the form

$$(Q_0 + Q_1 n + Q_2 n^2 + \ldots + Q_{m-1} n^{m-1} + Q_m n^m) a^n$$

if a is not a characteristic root of the recurrence relation.

Example 11.26: Solve $S(k) - S(k-1) - 6S(k-2) = -30$, where $S(0) = 20$, $S(1) = -5$.

Solution: The associated homogeneous relation is $S(k) - S(k-1) - 6S(k-2) = 0$.

The characteristic equation is $r^2 - r - 6 = 0$.

Its characteristic roots are $r = -2, 3$. So the homogeneous solution is $A_1(-2)^k + A_2(3)^k$.

Since the right hand side of $S(k) - S(k-1) - 6S(k-2) = -30$ (1) is a constant, by Rule 1, the particular solution will also be a constant, say Q into (1), we obtain

$$Q - Q - 6Q = -30 \Rightarrow -6Q = -30$$

$$\Rightarrow Q = 5$$

Therefore, the general solution is $S(k) = A_1(-2)^k + A_2(3)^k + 5$

Using the initial conditions, $S(0) = 20$, $S(1) = -5$

$$20 = A_1 + A_2 + 5$$

$$-5 = -2A_1 + 3A_2 + 5$$

which yield, $A_1 = 11$ and $A_2 = 4$. Hence the complete solution is $S(k) = 11(-2)^k + 4(3)^k + 5$.

***Rule* 3**. If a is a characteristic root of multiplicity $r-1$, when $f(n)$ is of the form

$$(k_0 + k_1 n + k_2 n^2 + \ldots + k_{m-1} n^{m-1} + k_m n^m) a^n$$

the corresponding particular solution is of the form

$$n^{r-1}(Q_0 + Q_1 n + Q_2 n^2 + \ldots + Q_{m-1} n^{m-1} + Q_m n^m) a^n.$$

Note: The general solution of the recurrence relation is the sum of the homogeneous solution and particular solution. If no initial conditions are given, then you have finished. If m initial conditions are given, obtain m linear equations in m unknowns and solve the system, if possible, to get a complete solution.

Example 11.27: Solve the recurrence relation $f_n - 5f_{n-1} + 6f_{n-2} = 1$. (1)

Solution: The associated homogeneous relation is $f_n - 5f_{n-1} + 6f_{n-2} = 0$.

The characteristic equation is $r^2 - 5r + 6 = 0$

Its characteristic roots are $r = 2, 3$ so the homogeneous solution is

$$A_1(2)^n + A_2(3)^n \quad (2)$$

Since the right hand side $f(n) = 1$ is a constant, by Rule 1, the particular solution will also be a constant, say Q. Substituting Q into (1), we obtain

$$Q - 5Q + 6Q = 1$$

This implies that $Q = \frac{1}{2}$. Therefore, the complete solution

$$f_n = A_1(2)^n + A_2(3)^n + \frac{1}{2}.$$

Example 11.28: Find the particular solution of the recurrence relation

$$f(n) + 5f(n-1) + 6f(n-2) = 3n^2 - 2n + 1 \quad (1)$$

Solution: By Rule 1, the particular solution is of the form

$$Q_0 + Q_1 n + Q_2 n^2 \quad (2)$$

Substituting Equation (2) into (1), we obtain

$$(Q_0 + Q_1 n + Q_2 n^2) + 5(Q_0 + Q_1 (n-1) + Q_2 (n-1)^2) + 6(Q_0 + Q_1(n-2) + Q_2(n-2)^2) = 3n^2 - 2n + 1$$

which simplifies to

$$(12Q_0 - 17Q_1 + 29Q_2) + (12Q_1 - 34Q_2)n + 12Q_2 n^2 = 3n^2 - 2n + 1 \quad (3)$$

Comparing the two sides of Equation (3), we obtain

$$12Q_2 = 3; \quad 12Q_1 - 34Q_2 = -2; \quad 12Q_0 - 17Q_1 + 29Q_2 = 1$$

which yields $Q_2 = \frac{1}{4}, Q_1 = \frac{13}{24}, Q_0 = \frac{71}{288}$. Therefore, the particular solution is

$$\frac{71}{288} + \frac{13}{24}n + \frac{1}{4}n^2.$$

Example 11.29: Solve $a_r + 5a_{r-1} = 9$ with initial condition $a_0 = 6$.

Solution: The associated homogeneous recurrence relation is $a_r + 5a_{r-1} = 0$.

It's characteristic equation is $r + 5 = 0$.

NOTES

Therefore, $r = -5$. So the homogeneous solution is $A_1(-5)^r$. Since the right hand side of the given relation is a constant, the particular solution will also be a constant Q. Substituting in the recurrence relation, we get

$$Q + 5Q = 9$$

So $Q = \frac{3}{2}$. Therefore the general solution is $a^r = A_1(-5)^r + \frac{3}{2}$. Using the initial condition $a_0 = 6$, we get

$$6 = A_1 + \frac{3}{2} \Rightarrow A_1 = \frac{9}{2}$$

Hence the complete solution is

$$a_r = \frac{9}{2}(-5)^r + \frac{3}{2}$$

Example 11.30: Solve the recurrence relation $f(n) - 7f(n-1) + 10f(n-2) = 6 + 8n$ with $f(0) = 1$ and $f(1) = 2$. (1)

Solution: The homogeneous solution is $A_1(2)^n + A(2)^n + A_2(5)^n$. By Rule 1, the particular solution is of the form $Q_0 + Q_1{}^n$.

Substituting in Equation (1), we get

$$(Q_0 + Q_1{}^n) - 7(Q_0 + Q_1(n-1)) + 10(Q_0 + Q_1(n-2)) = 6 + 8n$$

Comparing the two sides, we obtain

$$4Q_0 - 13Q_1 = 6 \quad \text{and} \quad 4Q_1 = 8$$

which yields, $Q_0 = 8$, $Q_1 = 2$. Therefore, the particular solution is $8 + 2n$ and the general solution is

$$f(n) = A_1(2)^n + A_2(5)^n + 8 + 2n$$

Now the initial conditions,

$$f(0) = 1 \Rightarrow A_1 + A_2 + 8 = 1$$

and

$$f(1) = 2 \Rightarrow 2A_1 + 5A_2 + 10 = 2$$

Solving, we get $A_1 = -9$ and $A_2 = 2$. Hence the complete solution is

$$f(n) = -9(2)^n + 2(5)^n + 8 + 2n$$

Example 11.31: Find the particular solution of the recurrence relation

$$a_n + 5a_{n-1} + 6a_{n-2} = 42(4)^n. \quad (1)$$

Solution: Now $r^2 + r + 6 = 0$, $r = -2, -3$ are characteristic roots. Since 4 is not a characteristic root, by Rule 2, we assume that the general form of the particular solution is $Q.(4)^n$. Substituting in Equation (1), we obtain

$$Q.(4)^n + 5Q.(4)^{n-1} + 6Q.(4)^{n-2} = 42(4)^n$$

$\Rightarrow Q.4^{n-2}[16+20+6] = 42\,(4)^n \Rightarrow Q.4^{n-2}\,(42) = 42\,(4)^n$

$\Rightarrow Q = 4^2 = 16$

Therefore, the particular solution is $16\,(4)^n = 4^{n+2}$.

NOTES

Example 11.32: Find the particular solution of the recurrence relation $f_n + f_{n-1} = 3n2^n$ (1)

Solution: The characteristic equation is $r+1=0$. Therefore $r=-1$ is a characteristic root. Since 2 is not a characteristic root, by Rule 2, the general form of the particular solution is $(Q_0+Q_1n)2^n$. Substituting in Equation (1), we obtain

$$(Q_0+Q_1n)2^n+(Q_0+Q_1(n-1))2^{n-1} = 3n2^n$$

which simplifies to

$$Q_0 2^n + Q_1 n2^n + \frac{1}{2}Q_0 2^n + \frac{1}{2}Q_1 n2^n - \frac{1}{2}Q_1 n2^n = 3n2^n$$

$$\Rightarrow (\frac{3}{2}Q_0 - \frac{1}{2}Q_1)2^n + \frac{3}{2}Q_1 n2^n = 3n2^n$$

Comparing the two sides, we obtain

$$\frac{3}{2}Q_0 - \frac{1}{2}Q_1 = 0$$

$$\frac{3}{2}Q_1 = 3$$

Thus, $Q_0 = \frac{2}{3}$ and $Q_1 = 2$ and the particular solution is $(\frac{2}{3}+2n)2^n$.

Example 11.33: Find the particular solution of the recurrence relation $f(n) - 2f(n-1) = 3.2^n$ (1)

Solution: The characteristic equation is $r-2=0$. Since $r=2$ is the characteristic root of multiplicity 1, by Rule (3), the general form of the particular solution is $Q_n 2^n$. Substituting in Equation (1), we obtain

$$Qn2^n - 2Q.\,(n-1)2^{n-1} = 3.2^n$$

$$\Rightarrow Q \cdot 2^n = 3 \cdot 2^n$$

$$\Rightarrow Q = 3$$

Thus the particular solution is $3n2^n$.

Example 11.34: Find the general solution of $f(n) - 3f(n-1) - 4f(n-2) = 4^n$ (1)

Solution: The associated homogeneous relation is

$$f(n) - 3f(n-1) - 4f(n-2) = 0.$$

NOTES

Its characteristic equation is $r^2 - 3r - 4 = 0$.

Solving, we get $r = -1, 4$ the characteristic roots. Therefore, the homogeneous solution is $A_1(-1)^n + A_2(4)^n$. Since 4 is a characteristic root, by Rule 3, we assume that the general form of the particular solution is $Qn4^n$.

Substituting in (1), we get

$$Qn4^n - 3Q(n-1)4^{n-1} - 4Q(n-2)4^{n-2} = 4^n$$

$$\Rightarrow Qn4^n - 3Qn4^{n-1} + 3Q4^{n-1} - 4Qn4^{n-2} + 8Q4^{n-2} = 4^n$$

$$\Rightarrow (16Qn - 12Qn + 12Q - 4Qn + 8Q)4^{n-2} = (16)4^{n-2}$$

$$\Rightarrow 20Q = 16 \Rightarrow Q = \frac{4}{5}$$

Therefore, the particular solution is $\frac{4}{5} n4^n$. Hence, the general solution of the recurrence relation of

$$f(n) = A_1(-1)^n + A_2(4)^n + \frac{4}{5} n4^n$$

Note: What if the characteristic equation gives rise to complex roots? Here, our methods are still valid, but the method for expressing the solutions of the recurrence relations is different. Since an understanding of these representations require some background in complex numbers, we suggest that an interested reader refer to a more advanced treatment of recurrence relations.

11.3 POLYA THEORY

"Polya theorem" redirects here. For Pólya's theorem for positive polynomials on simplex, see Positive polynomial. For the recurrence of lattice random walks, see random walk higher dimensions.

The Pólya enumeration theorem, also known as the Redfield–Pólya theorem and Pólya counting, is a theorem in combinatorics that both follows from and ultimately generalizes Burnside's lemma on the number of orbits of a group action on a set. The theorem was first published by J. Howard Redfield in 1927. In 1937 it was independently rediscovered by George Pólya, who then greatly popularized the result by applying it to many counting problems, in particular to the enumeration of chemical compounds.

The Pólya enumeration theorem has been incorporated into symbolic combinatorics and the theory of combinatorial species.

Simplified, Unweighted Version

Let X be a finite set and let G be a group of permutations of X (or a finite symmetry group that acts on X). The set X may represent a finite set of beads, and G may be a chosen group of permutations of the beads. For example, if X is a necklace of n

beads in a circle, then rotational symmetry is relevant so G is the cyclic group C_n, while if X is a bracelet of n beads in a circle, rotations and reflections are relevant so G is the dihedral group D_n of order $2n$. Suppose further that Y is a finite set of colors — the colors of the beads — so that Y^X is the set of colored arrangements of beads (more formally: YX is the set of functions X'!Y) Then the group G acts on Y^X. The Pólya enumeration theorem counts the number of orbits under G of colored arrangements of beads by the following formula:

$$|Y^X/G| = \frac{1}{|G|}\sum_{g\in G} m^{c(g)}$$

where $m = |Y|$ is the number of colors and $c(g)$ is the number of cycles of the group element g when considered as a permutation of X.

Full, Weighted Version

In the more general and more important version of the theorem, the colors are also weighted in one or more ways, and there could be an infinite number of colors provided that the set of colors has a generating function with finite coefficients. In the univariate case, suppose that

$$f(t) = f_0 + f_1 t + f_2 t^2 + \cdots$$

is the generating function of the set of colors, so that there are fw colors of weight w for each integer $w \geq 0$. In the multivariate case, the weight of each color is a vector of integers and there is a generating function $f(t_1, t_2, ...)$ that tabulates the number of colors with each given vector of weights.

The enumeration theorem employs another multivariate generating function called the cycle index:

$$Z_G(t_1, t_2, \ldots, t_n) = \frac{1}{|G|}\sum_{g\in G} t_1^{c_1(g)} t_2^{c_2(g)} \cdots t_n^{c_k(g)}$$

where n is the number of elements of X and $ck(g)$ is the number of k-cycles of the group element g as a permutation of X.

A colored arrangement is an orbit of the action of G on the set $'Y^X$ (where Y is the set of colors and YX denotes the set of all functions $\varphi: X \to Y$). The weight of such an arrangement is defined as the sum of the weights of $\varphi(x)$ over all x in X. The theorem states that the generating function F of the number of colored arrangements by weight is given by:

$$F(t) = Z_G(f(t), f(t^2), f(t^3), \ldots, f(t^n))$$

or in the multivariate case:

$$F(t_1, t_2, \ldots) = Z_G(f(t_1, t_2, \ldots), f(t_1^2, t_2^2, \ldots), f(t_1^3, t_2^3, \ldots), \ldots, f(t_1^n, t_2^n, \ldots)).$$

To reduce to the simplified version given earlier, if there are m colors and all have weight 0, then $f(t) = m$ and

$$\left|Y^X/G\right| = F(0) = Z_G(m, m, \ldots, m) = \frac{1}{|G|}\sum_{g\in G} m^{c(g)}.$$

NOTES

In the celebrated application of counting trees (see below) and acyclic molecules, an arrangement of "colored beads" is actually an arrangement of arrangements, such as branches of a rooted tree. Thus the generating function f for the colors is derived from the generating function F for arrangements, and the Pólya enumeration theorem becomes a recursive formula.

Examples

Colored Cubes: How many ways are there to color the sides of a three-dimensional cube with m colors, up to rotation of the cube? The rotation group C of the cube acts on the six sides of the cube, which are equivalent to beads. Its cycle index is

$$Z_C(t_1, t_2, t_3, t_4) = \frac{1}{24}\left(t_1^6 + 6t_1^2t_4 + 3t_1^2t_2^2 + 8t_3^2 + 6t_2^3\right)$$

which is obtained by analyzing the action of each of the 24 elements of C on the 6 sides of the cube, see here for the details. We take all colors to have weight 0 and find that there are

$$F(0) = Z_C(m, m, m, m) = \frac{1}{24}\left(m^6 + 3m^4 + 12m^3 + 8m^2\right)$$ different colorings.

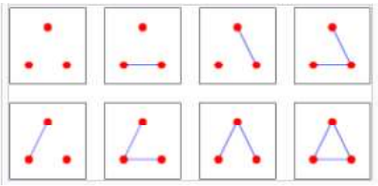
Graphs on Three and Four Vertices

A graph on m vertices can be interpreted as an arrangement of colored beads.

The set X of "beads" is the set of $\binom{m}{2}$ possible edges, while the set of colors Y = {black, white} corresponds to edges that are present (black) or absent (white). The Pólya enumeration theorem can be used to calculate the number of graphs up to isomorphism with a fixed number of vertices, or the generating function of these graphs according to the number of edges they have. For the latter purpose, we can say that a black or present edge has weight 1, while an absent or white edge has weight 0. Thus $f(t) = 1 + t$ is the generating function for the set of colors. The relevant symmetry group is $G = S_m$, the symmetric group on m letters. This group acts on the set X of possible edges: a permutation φ turns the edge $\{a, b\}$ into the edge $\{\varphi(a), \varphi(b)\}$. With these definitions, an isomorphism class of graphs with m vertices is the same as an orbit of the action of G on the set YX of colored arrangements; the number of edges of the graph equals the weight of the arrangement.

The eight graphs on three vertices (before identifying isomorphic graphs) are shown at the right. There are four isomorphism classes of graphs, also shown at the right. The cycle index of the group S_3 acting on the set of three edges is

$$Z_G(t_1, t_2, t_3) = \frac{1}{6}\left(t_1^3 + 3t_1t_2 + 2t_3\right)$$

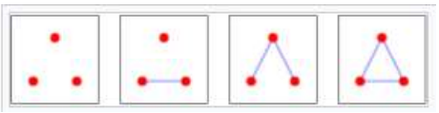


(obtained by inspecting the cycle structure of the action of the group elements; see here). Thus, according to the enumeration theorem, the generating function of graphs on 3 vertices up to isomorphism is

$$F(t) = Z_G\left(t+1, t^2+1, t^3+1\right) = \frac{1}{6}\left((t+1)^3 + 3(t+1)(t^2+1) + 2(t^3+1)\right),$$

which simplifies to

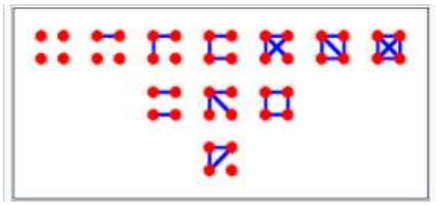
$$F(t) = t^3 + t^2 + t + 1.$$



Thus there is one graph each with 0 to 3 edges.

The cycle index of the group S_4 acting on the set of 6 edges is

$$Z_G(t_1, t_2, t_3, t_4) = \frac{1}{24}\left(t_1^6 + 9t_1^2t_2^2 + 8t_3^2 + 6t_2t_4\right)$$



(see here.) Hence

$$F(t) = Z_G\left(t+1, t^2+1, t^3+1, t^4+1\right) = \frac{(t+1)^6 + 9(t+1)^2(t^2+1)^2 + 8(t^3+1)^2 + 6(t^2+1)(t^4+1)}{24}$$

which simplifies to

$$F(t) = t^6 + t^5 + 2t^4 + 3t^3 + 2t^2 + t + 1.$$

These graphs are shown at the right.

Rooted Ternary Trees

The set T_3 of rooted ternary trees consists of rooted trees where every node (or non-leaf vertex) has exactly three children (leaves or subtrees). Small ternary trees are shown at right. Note that rooted ternary trees with n nodes are equivalent to rooted trees with n vertices of degree at most 3 (by ignoring the leaves). In general, two rooted trees are isomorphic when one can be obtained from the other by permuting the children of its nodes. In other words, the group that acts on the children of a node is the symmetric group S_3. We define the weight of such a ternary tree to be the number of nodes (or non-leaf vertices).

NOTES

Rooted ternary trees on 0, 1, 2, 3 and 4 nodes (= non-leaf vertices). The root is shown in blue, the leaves are not shown. Every node has as many leaves as to make the number of its children equal to 3.

One can view a rooted, ternary tree as a recursive object which is either a leaf or a node with three children which are themselves rooted ternary trees. These children are equivalent to beads; the cycle index of the symmetric group S_3 that acts on them is

$$Z_{S_3}(t_1, t_2, t_3) = \frac{t_1^3 + 3t_1t_2 + 2t_3}{6}.$$

The Polya enumeration theorem translates the recursive structure of rooted ternary trees into a functional equation for the generating function $F(t)$ of rooted ternary trees by number of nodes. This is achieved by "coloring" the three children with rooted ternary trees, weighted by node number, so that the color generating function is given by $f(t) = F(t)$ which by the enumeration theorem gives

$$\frac{F(t)^3 + 3F(t)F(t^2) + 2F(t^3)}{6}$$

as the generating function for rooted ternary trees, weighted by one less than the node number (since the sum of the children weights does not take the root into account), so that

$$F(t) = 1 + t \cdot \frac{F(t)^3 + 3F(t)F(t^2) + 2F(t^3)}{6}.$$

This is equivalent to the following recurrence formula for the number t_n of rooted ternary trees with n nodes:

$$t_0 = 1$$

$$t_{n+1} = \frac{1}{6}\left(\sum_{a+b+c=n} t_a t_b t_c + 3\sum_{a+2b=n} t_a t_b + 2\sum_{3a=n} t_a\right)$$

where a, b and c are nonnegative integers. The first few values of t_n are 1, 1, 1, 2, 4, 8, 17, 39, 89, 211, 507, 1238, 3057, 7639, 1924.

Proof of Theorem

The simplified form of the Pólya enumeration theorem follows from Burnside's lemma, which says that the number of orbits of colorings is the average of the number of elements of Y^X fixed by the permutation g of G over all permutations g. The weighted version of the theorem has essentially the same proof, but with a refined form of Burnside's lemma for weighted enumeration. It is equivalent to apply Burnside's lemma separately to orbits of different weight. For clearer notation, let $x_1, x_2, n \ldots$ be the variables of the generating function f of Y. Given a vector of weights w, let x^w denote the corresponding monomial term of f applying Burnside's lemma to orbits of weight w, the number of orbits of this weight is

NOTES

$$\frac{1}{|G|}\sum_{g\in G}\left|(Y^X)_{\omega,g}\right|$$

were $(Y^X)_{\omega,g}$ is the set of colorings of weight w that are also fixed by g. If we then sum over all possible weights, we obtain

$$F(x_1,x_2,\ldots)=\frac{1}{|G|}\sum_{g\in G,\omega}x^{\omega}\left|(Y^X)_{\omega,g}\right|.$$

Meanwhile a group element g with cycle structure $j_1(g),j_2(g),\ldots,j_n(g)$ will contribute the term $t_1^{j_1(g)}t_2^{j_2(g)}\cdots t_n^{j_n(g)}$ to the cycle index of G. The element g fixes an element $\phi\in Y^X$ if and only if the function φ is constant on every cycle q of g. For every such cycle q, the generating function by weight of $|q|$ identical colors from the set enumerated by f are

$$f\left(x_1^{|q|},x_2^{|q|},x_3^{|q|},\ldots\right).$$

It follows that the generating function by weight of the points fixed by g is the product of the above term over all cycles of g, i.e.

$$\begin{aligned}\sum_{\omega}x^{\omega}\left|(Y^X)_{\omega,g}\right| &= \prod_{q\text{ cycle of }g}f\left(x_1^{|q|},x_2^{|q|},x_3^{|q|},\ldots\right)\\ &= f(x_1,x_2,\ldots)^{j_1(g)}f\left(x_1^2,x_2^2,\ldots\right)^{j_2(g)}\cdots f(x_1^n,x_2^n,\ldots)^{j_n(g)}\end{aligned}$$

Substituting this in the sum over all g yields the substituted cycle index as claimed.

Check Your Progress

1. Define Fibonacci numbers.
2. What do you mean by iterative procedure?
3. How do you define the sequence of integers?
4. What do you understand by Polya theory?

11.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The numbers in the sequence 0, 1, 2, 3, 5, 8, 13, 21, ... in which each new term is the sum of the previous two terms are called the Fibonacci numbers.
2. There is another way to evaluate a function from its recursive definition. Instead of successively reducing the computation to evaluate of the function at small integers, we can start with the basis and successively apply the

NOTES

recursive definition to find the values of the function at successive larger integers. Such a procedure is called iterative procedure.

3. A sequence of integers is a function from the natural numbers into integers. That is, if f is a sequence of integers, then $f: N \rightarrow Z$ is a function. We use the notation fn or $f(n)$ to denote the image of any natural n. We call fn an nth term of the sequence.
4. The Pólya enumeration theorem, also known as the Redfield–Pólya theorem and Pólya counting, is a theorem in combinatorics that both follows from and ultimately generalizes Burnside's lemma on the number of orbits of a group action on a set. The theorem was first published by J. Howard Redfield in 1927. In 1937 it was independently rediscovered by George Pólya, who then greatly popularized the result by applying it to many counting problems, in particular to the enumeration of chemical compounds.

11.5 SUMMARY

- The numbers in the sequence 0, 1, 2, 3, 5, 8, 13, 21, ... in which each new term is the sum of the previous two terms are called the Fibonacci numbers.
- There is another way to evaluate a function from its recursive definition.
- Successively reducing the computation to evaluate of the function at small integers, we can start with the basis and successively apply the recursive definition to find the values of the function at successive larger integers. such a procedure is called iterative procedure.
- Often an iterative approach for the evaluation of a recursively defined sequence requires much less computation than a procedure using recursion. But it is sometimes preferable to use a recursive procedure even if it is less efficient than the iterative procedure.
- A sequence of integers is a function from the natural numbers into integers. That is, if f is a sequence of integers, then $f: N \rightarrow Z$ is a function. We use the notation fn or $f(n)$ to denote the image of any natural n. We call fn an nth term of the sequence.
- A formal definition of recurrence relations is difficult because of the wide variety of forms in which such relations can be written, but the concept is straight forward.
- The Pólya enumeration theorem, also known as the Redfield–Pólya theorem and Pólya counting, is a theorem in combinatorics that both follows from and ultimately generalizes Burnside's lemma on the number of orbits of a group action on a set.
- The theorem was first published by J. Howard Redfield in 1927. In 1937 it was independently rediscovered by George Pólya, who then greatly

popularized the result by applying it to many counting problems, in particular to the enumeration of chemical compounds.

- The Pólya enumeration theorem has been incorporated into symbolic combinatorics and the theory of combinatorial species.

NOTES

11.6 KEY WORDS

- **Fibonacci numbers:** The numbers in the sequence 0, 1, 2, 3, 5, 8, 13, 21, ... in which each new term is the sum of the previous two terms are called the Fibonacci numbers.
- **Iterative procedure:** There is another way to evaluate a function from its recursive definition. Instead of successively reducing the computation to evaluate of the function at small integers, we can start with the basis and successively apply the recursive definition to find the values of the function at successive larger integers. Such a procedure is called iterative procedure.
- **Polya theory:** The Pólya enumeration theorem, also known as the Redfield–Pólya theorem and Pólya counting, is a theorem in combinatorics that both follows from and ultimately generalizes Burnside's lemma on the number of orbits of a group action on a set.

11.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What do you understand by Fibonacci numbers?
2. What is meant by Polya theory?

Long-Answer Questions

1. Describe the concept of Fibonacci with the help of examples.
2. Explain the iterative procedure giving examples.
3. Discuss the Polya theory with the help of appropriate examples.

11.8 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

NOTES

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

NOTES

BLOCK IV
PERMUTATION GROUPS

UNIT 12 NECKLACE PROBLEM AND BURNSIDE'S LEMMA

Structure

12.0 Introduction
12.1 Objectives
12.2 Necklace Problem
12.3 Burnside's Lemma
12.4 Cyclic Index of a Permutation Group
12.5 Answers to Check Your Progress Questions
12.6 Summary
12.7 Key Words
12.8 Self Assessment Questions and Exercises
12.9 Further Readings

12.0 INTRODUCTION

In combinatorics, a *k*-ary necklace of length *n* is an equivalence class (a grouping for which there exists an equivalence relation) of *n*-character strings over an alphabet of size *k*, taking all rotations as equivalent. It represents a structure with *n* circularly connected beads which have *k* available colours.

A *k*-ary bracelet, also referred to as a turnover (or free) necklace, is a necklace such that strings may also be equivalent under reflection. That is, given two strings, if each is the reverse of the other, they belong to the same equivalence class. For this reason, a necklace might also be called a fixed necklace to distinguish it from a turnover necklace. Formally, one may represent a necklace as an orbit of the cyclic group acting on *n*-character strings, and a bracelet as an orbit of the dihedral group. One can count these orbits, and thus necklaces and bracelets, using Pólya's enumeration theorem.

Burnside's lemma, sometimes also called Burnside's counting theorem, the Cauchy–Frobenius lemma, orbit-counting theorem, or The Lemma that is not Burnside's, is a combinatorial result in group theory which is often useful in taking account of symmetry when counting mathematical objects. Its various eponyms are based on William Burnside, George Pólya, Augustin Louis Cauchy, and Ferdinand Georg Frobenius. The result is not due to Burnside himself, who merely quotes it in his book "On the Theory of Groups of Finite Order", attributing it instead to Frobenius (1887).

William Burnside stated and proved this lemma, attributing it to Frobenius 1887, in his 1897 book on finite groups. But, even prior to Frobenius, the formula was known to Cauchy in 1845. In fact, the lemma was apparently so well known that Burnside simply omitted to attribute it to Cauchy. Consequently, this lemma is sometimes referred to as the lemma that is not Burnside's. This is less ambiguous than it may seem: Burnside contributed many lemmas to this field.

In this unit, you will study about the necklace problem, Burnside's lemma, and the cyclic index of a permutation group.

12.1 OBJECTIVES

After going through this unit, you will be able to:

- Interpret the necklace problem
- Explain the Burnside's lemma
- Comprehend the cyclic index of a permutation group

12.2 NECKLACE PROBLEM

The necklace problem is a problem in recreational mathematics concerning the reconstruction of necklaces (cyclic arrangements of binary values) from partial information.

The necklace problem involves the reconstruction of a necklace of n beads, each of which is either black or white, from partial information. The information specifies how many copies the necklace contains of each possible arrangement of k black beads. For instance, for $k=2$ the specified information gives the number of pairs of black beads that are separated by i positions, for $i = 0, \ldots [n/2 - 1]$.

This can be made formal by defining a k-configuration to be a necklace of k-black beads and $n - k$ white beads, and counting the number of ways of rotating a k-configuration so that each of its black beads coincides with one of the black beads of the given necklace.

The necklace problem asks: if is given, and the numbers of copies of each k-configuration are known up to some threshold $k \leq K$ how large does the threshold K need to be before this information completely determines the necklace that it describes? Equivalently, if the information about a k-configuration is provided in stages, where the kth stage provides the numbers of copies of each -configuration how many stages are needed (in the worst case) in order to reconstruct the precise pattern of black and white beads in the original necklace?

For a given set of n beads, all distinct, the number of distinct necklaces made from these beads, counting rotated necklaces as the same, is

$$\frac{n!}{n} = (n-1)!.$$

This is because the beads can be linearly ordered in $n!$ ways, and the n circular shifts of such an ordering all give the same necklace. Similarly, the number of distinct bracelets, counting rotated and reflected bracelets as the same, is

NOTES

$$\frac{n!}{2n}, \text{ for } n \geq 3.$$

If the beads are not all distinct, having repeated colours, then there are fewer necklaces (and bracelets). The above necklace-counting polynomials give the number necklaces made from all possible multi-sets of beads. Polya's pattern inventory polynomial refines the counting polynomial, using variable for each bead colour, so that the coefficient of each monomial counts the number of necklaces on a given multi-set of beads.

Number of Necklaces

There are

$$N_k(n) = \frac{1}{n} \sum_{d|n} \varphi(d) k^{\frac{n}{d}}$$

Different k-ary necklaces of length n, where φ is Euler's totient function. This follows directly from Pólya's enumeration theorem applied to the action of the cyclic group C_n acting on the set of all functions

$$f : \{1, ..., n\} \to \{1, ... , k\}.$$

There are also

$$L_k(n) = \frac{k!}{n} \sum_{d|n} \varphi(d) S(\frac{n}{d}, k)$$

Different necklaces of length n with exactly k different coloured beads, where $S(n, k)$ are the Stirling number of the second kind. (The variable k is overloaded and it is unclear whether k refers to the alphabet size or to the number of distinct elements in the necklace).

$N_k(n)$ and $L_k(n)$ are related via the Binomial coefficients:

$$N_k(n) = \sum_{j=1}^{k} \binom{k}{j} L_k(n)$$

And

$$L_k(n) = \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} N_k(n).$$

NOTES

12.3 BURNSIDE'S LEMMA

Burnside's lemma, sometimes also called Burnside's counting theorem, the Cauchy–Frobenius lemma, orbit-counting theorem, or The Lemma that is not Burnside's, is a result in group theory which is often useful in taking account of symmetry when counting mathematical objects. William Burnside stated and proved this lemma, attributing it to Frobenius 1887, in his 1897 book on finite groups. But, even prior to Frobenius, the formula was known to Cauchy in 1845. In fact, the lemma was apparently so well known that Burnside simply omitted to attribute it to Cauchy. Consequently, this lemma is sometimes referred to as the lemma that is not Burnside's. This is less ambiguous than it may seem: Burnside contributed many lemmas to this field. Its various eponyms are based on William Burnside, George Pólya, Augustin Louis Cauchy, and Ferdinand Georg Frobenius. The result is not due to Burnside himself, who merely quotes it in his book "On the Theory of Groups of Finite Order", attributing it instead to Frobenius.

In the following, let G be a finite group that acts on a set X. For each g in G let X^g denote the set of elements in X that are fixed by g (also said to be left invariant by g), i.e. $X^g = \{x \in X \mid g.x = x\}$. Burnside's lemma asserts the following formula for the number of orbits, denoted $|X/G|$:

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |X^g|.$$

Thus the number of orbits (a natural number or $+\infty$) is equal to the average number of points fixed by an element of G (which is also a natural number or infinity). If G is infinite, the division by $|G|$ may not be well-defined; in this case the following statement in cardinal arithmetic holds:

$$|G||X/G| = \sum_{g \in G} |X^g|.$$

Proof

The first step in the proof of the lemma is to re-express the sum over the group elements $g \in G$ as an equivalent sum over the set of elements $x \in X$:

$$\sum_{g \in G} |X^g| = |\{(g, x) \in G \times X \mid g \cdot x = x\}| = \sum_{x \in X} |G_x|.$$

(Here $X^g = \{x \in X \mid g.x = x\}$ is the subset of all points of X fixed by $g \in G$, whereas $G_x = \{g \in G \mid g.x = x\}$ is the stabilizer subgroup of G that fixes the point $x \in X$).

The orbit-stabilizer theorem says that there is a natural bijection for each $x \in X$ between the orbit of x, $G.x = \{g.x \mid g \in G\} \subseteq X$, and the set of left cosets G/G_x of its stabilizer subgroup G_x. With Lagrange's theorem this implies

$$|G \cdot x| = [G : G_x] = |G|/|G_x|.$$

Our sum over the set X may therefore be rewritten as

$$\sum_{x \in X} |G_x| = \sum_{x \in X} \frac{|G|}{|G \cdot x|} = |G| \sum_{x \in X} \frac{1}{|G \cdot x|}.$$

Finally, notice that X is the disjoint union of all its orbits in X/G, which means the sum over X may be broken up into separate sums over each individual orbit.

$$\sum_{x \in X} \frac{1}{|G \cdot x|} = \sum_{A \in X/G} \sum_{x \in A} \frac{1}{|A|} = \sum_{A \in X/G} 1 = |X/G|.$$

Putting everything together gives the desired result:

$$\sum_{g \in G} |X^g| = |G| \cdot |X/G|.$$

This proof is essentially also the proof of the class equation formula, simply by taking the action of G on itself ($X = G$) to be by conjugation, $g.x = gxg^{-1}$, in which case G_x instantiates to the centralizer of x in G.

12.4 CYCLIC INDEX OF A PERMUTATION GROUP

In combinatorial mathematics a cycle index is a polynomial in several variables which is structured in such a way that information about how a group of permutations acts on a set can be simply read off from the coefficients and exponents. This compact way of storing information in an algebraic form is frequently used in combinatorial enumeration.

Each permutation π of a finite set of objects partitions that set into cycles; the cycle index monomial of π is a monomial in variables $a_1, a_2, \ldots$ that describes the type of this partition (the cycle type of π): the exponent of a_i is the number of cycles of π of size i. The cycle index polynomial of a permutation group is the average of the cycle index monomials of its elements. The phrase cycle indicator is also sometimes used in place of cycle index.

Knowing the cycle index polynomial of a permutation group, one can enumerate equivalence classes due to the group's action. This is the main ingredient in the Pólya enumeration theorem. Performing formal algebraic and differential operations on these polynomials and then interpreting the results combinatorially lies at the core of species theory.

NOTES

Definition: The cycle index of a permutation group G is the average of the cycle index monomials of all the permutations g in G. More formally, let G be a permutation group of order m and degree n. Every permutation g in G has a unique decomposition into disjoint cycles, say $c_1 c_2 c_3 \ldots$.

Let the length of a cycle c be denoted by $|c|$.

Now, let $j_k(g)$ be the number of cycles of g of length k, where

$$0 \leq j_k(g) \leq \lfloor n/k \rfloor \text{ and } \sum_{k=1}^{n} k\, j_k(g) = n.$$

We associate to g the monomial

$$\prod_{c \in g} a_{|c|} = \prod_{k=1}^{n} a_k^{j_k(g)}$$

In the variables $a_1, a_2, \ldots, a_n$.

Then, the cycle index $Z(G)$ of G is given by

$$Z(G) = \frac{1}{|G|} \sum_{g \in G} \prod_{k=1}^{n} a_k^{j_k(g)}.$$

Check Your Progress

1. Elaborate on the necklace problem.
2. Define the Burnside's lemma.
3. Explain the cyclic index of a permutation group.

12.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The necklace problem involves the reconstruction of a necklace of n beads, each of which is either black or white, from partial information. The information specifies how many copies the necklace contains of each possible arrangement of k black beads. For instance, for $k = 2$ the specified information gives the number of pairs of black beads that are separated by i positions, for $i = 0, \ldots [n/2 - 1]$.
2. Burnside's lemma, sometimes also called Burnside's counting theorem, the Cauchy–Frobenius lemma, orbit-counting theorem, or The Lemma that is not Burnside's, is a result in group theory which is often useful in taking account of symmetry when counting mathematical objects.
3. The cycle index of a permutation group G is the average of the cycle index monomials of all the permutations g in G. More formally, let G be a

permutation group of order m and degree n. Every permutation g in G has a unique decomposition into disjoint cycles, say $c_1 c_2 c_3 \ldots$.

12.6 SUMMARY

NOTES

- The necklace problem is a problem in recreational mathematics concerning the reconstruction of necklaces (cyclic arrangements of binary values) from partial information.
- The necklace problem involves the reconstruction of a necklace of n beads, each of which is either black or white, from partial information. The information specifies how many copies the necklace contains of each possible arrangement of k black beads.
- The necklace problem asks: if n is given, and the numbers of copies of each k-configuration are known up to some threshold $k \leq K$ how large does the threshold K need to be before this information completely determines the necklace that it describes?
- Burnside's lemma, sometimes also called Burnside's counting theorem, the Cauchy–Frobenius lemma, orbit-counting theorem, or The Lemma that is not Burnside's, is a result in group theory which is often useful in taking account of symmetry when counting mathematical objects.
- William Burnside stated and proved this lemma, attributing it to Frobenius 1887, in his 1897 book on finite groups. But, even prior to Frobenius, the formula was known to Cauchy in 1845. In fact, the lemma was apparently so well known that Burnside simply omitted to attribute it to Cauchy.
- In combinatorial mathematics a cycle index is a polynomial in several variables which is structured in such a way that information about how a group of permutations acts on a set can be simply read off from the coefficients and exponents.
- The cycle index of a permutation group G is the average of the cycle index monomials of all the permutations g in G. More formally, let G be a permutation group of order m and degree n. Every permutation g in G has a unique decomposition into disjoint cycles, say $c_1 c_2 c_3 \ldots$.

12.7 KEY WORDS

- **Necklace problem:** The necklace problem is a problem in recreational mathematics concerning the reconstruction of necklaces (cyclic arrangements of binary values) from partial information.
- **Burnside's lemma:** Burnside's lemma, sometimes also called Burnside's counting theorem, the Cauchy–Frobenius lemma, orbit-counting theorem,

NOTES

or The Lemma that is not Burnside's, is a result in group theory which is often useful in taking account of symmetry when counting mathematical objects.

- **Cycle index of a permutation group:** The cycle index of a permutation group G is the average of the cycle index monomials of all the permutations g in G.

12.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. State the necklace problem.
2. Illustrate the Burnside's lemma.
3. Define the cyclic index of a permutation group.

Long-Answer Questions

1. Discuss briefly the necklace problem. Give appropriate examples.
2. Describe the Burnside's lemma with its applications.
3. Analyse the cyclic index of a permutation group.

12.9 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

UNIT 13 POLYA'S THEOREMS AND THEIR IMMEDIATE APPLICATIONS

Structure

13.0 Introduction
13.1 Objectives
13.2 Polya's Theorems
13.2.1 Applications
13.3 Answers to Check Your Progress Questions
13.4 Summary
13.5 Key Words
13.6 Self Assessment Questions and Exercises
13.7 Further Readings

13.0 INTRODUCTION

The Pólya enumeration theorem, also known as the Redfield–Pólya theorem and Pólya counting theorem, is a theorem in combinatorics that both follows from and ultimately generalizes Burnside's lemma on the number of orbits of a group action on a set. Polya's theorem can be used to enumerate objects under permutation groups. Using group theory, combinatorics and some examples, Polya's theorem and Burnside's lemma are derived. The examples used are a square, pentagon, hexagon and heptagon under their respective dihedral groups. Generalization using more permutations and applications to graph theory.

The theorem was first published by J. Howard Redfield in 1927. In 1937 it was independently rediscovered by George Pólya, who then greatly popularized the result by applying it to many counting problems, in particular to the enumeration of chemical compounds.

The Pólya enumeration theorem has been incorporated into symbolic combinatorics and the theory of combinatorial species.

In this unit, you will study about the Polya's theorems and their immediate applications.

13.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the Polya's theorem
- Explain the immediate applications of Polya's theorem

NOTES

13.2 POLYA'S THEOREMS

This is a most general theorem which allows the number of discrete combinatorial objects of a given types to be enumerated (counted) as a function of their "Order". The Pólya enumeration theorem, also known as the Redfield–Pólya theorem and Pólya counting, is a theorem in combinatorics that both follows from and ultimately generalizes Burnside's lemma on the number of orbits of a group action on a set. The theorem was first published by J. Howard Redfield in 1927. In 1937 it was independently rediscovered by George Pólya, who then greatly popularized the result by applying it to many counting problems, in particular to the enumeration of chemical compounds.

The Pólya enumeration theorem has been incorporated into symbolic combinatorics and the theory of combinatorial species.

Polya's Theorem (Unweighted)

Let X be a finite set and let G be a group of permutations of X (or a finite symmetry group that acts on X). The set X may represent a finite set of beads, and G may be a chosen group of permutations of the beads. For example, if X is a necklace of n beads in a circle, then rotational symmetry is relevant so G is the cyclic group C_n, while if X is a bracelet of n beads in a circle, rotations and reflections are relevant so G is the dihedral group D_n of order $2n$. Suppose further that Y is a finite set of colours — the colours of the beads — so that Y^X is the set of coloured arrangements of beads (more formally: Y^X is the set of functions $X \to Y$) Then the group G acts on Y^X. The Pólya enumeration theorem counts the number of orbits under G of coloured arrangements of beads by the following formula:

$$\left|Y^X/G\right| = \frac{1}{|G|}\sum_{g\in G} m^{c(g)}$$

Where $m = |\ Y\ |$ is the number of colours and $c(g)$ is the number of cycles of the group element g when considered as a permutation of X.

Polya's Theorem (Weighted)

In the more general and more important version of the theorem, the colours are also weighted in one or more ways, and there could be an infinite number of colours provided that the set of colours has a generating function with finite coefficients. In the univariate case, suppose that

$$f(t) = f_0 + f_1 t + f_2 t^2 + \cdots$$

is the generating function of the set of colors, so that there are f_w colors of weight w for each integer $w \geq 0$. In the multivariate case, the weight of each color is a vector of integers and there is a generating function $f(t_1, t_2, ...)$ that tabulates the number of colors with each given vector of weights.

NOTES

The enumeration theorem employs another multivariate generating function called the cycle index:

$$Z_G(t_1, t_2, \ldots, t_n) = \frac{1}{|G|} \sum_{g \in G} t_1^{c_1(g)} t_2^{c_2(g)} \cdots t_n^{c_k(g)}$$

Where n is the number of elements of X and $c_k(g)$ is the number of k-cycles of the group element g as a permutation of X.

A colored arrangement is an orbit of the action of G on the set Y^X (where Y is the set of colors and Y^X denotes the set of all functions $\varphi: X \to Y$). The weight of such an arrangement is defined as the sum of the weights of $\varphi(x)$ over all x in X. The theorem states that the generating function F of the number of colored arrangements by weight is given by:

$$F(t) = Z_G(f(t), f(t^2), f(t^3), \ldots, f(t^n))$$

Or in the multivariate case:

$$F(t_1, t_2, \ldots) = Z_G(f(t_1, t_2, \ldots), f(t_1^2, t_2^2, \ldots), f(t_1^3, t_2^3, \ldots), \ldots, f(t_1^n, t_2^n, \ldots)).$$

To reduce to the simplified version given earlier, if there are m colors and all have weight 0, then $f(t) = m$ and

$$\left|Y^X/G\right| = F(0) = Z_G(m, m, \ldots, m) = \frac{1}{|G|} \sum_{g \in G} m^{c(g)}.$$

In the celebrated application of counting trees and acyclic molecules, an arrangement of "Colored Beads" is actually an arrangement of arrangements, such as branches of a rooted tree. Thus the generating function f for the colors is derived from the generating function F for arrangements, and the Pólya enumeration theorem becomes a recursive formula.

Proof

The simplified form of the Pólya enumeration theorem follows from Burnside's lemma, which says that the number of orbits of colorings is the average of the number of elements of Y^X fixed by the permutation g of G over all permutations g. The weighted version of the theorem has essentially the same proof, but with a refined form of Burnside's lemma for weighted enumeration. It is equivalent to apply Burnside's lemma separately to orbits of different weight.

For clearer notation, let $x_1, x_2, \ldots$ be the variables of the generating function f of Y. Given a vector of weights w let x^w denote the corresponding monomial term of f. Applying Burnside's lemma to orbits of weight w the number of orbits of this weight is

$$\frac{1}{|G|} \sum_{g \in G} \left|(Y^X)_{\omega, g}\right|$$

NOTES

Where $(Y^X)_{w,g}$ is the set of colorings of weight w that are also fixed by g. If we then sum over all possible weights, we obtain

$$F(x_1, x_2, \ldots) = \frac{1}{|G|} \sum_{g \in G, \omega} x^{\omega} \left|(Y^X)_{\omega, g}\right|.$$

Meanwhile a group element g with cycle structure $j_1(g), j_2(g), \ldots, j_n(g)$ will contribute the term

$$t_1^{j_1(g)} t_2^{j_2(g)} \cdots t_n^{j_n(g)}$$

To the cycle index of G. The element g fixes an element $\phi \in Y^X$ if and only if the function φ is constant on every cycle q of g. For every such cycle q, the generating function by weight of $|q|$ identical colors from the set enumerated by f is

$$f\left(x_1^{|q|}, x_2^{|q|}, x_3^{|q|}, \ldots\right).$$

It follows that the generating function by weight of the points fixed by g is the product of the above term over all cycles of g, i.e.

$$\sum_{\omega} x^{\omega} \left|(Y^X)_{\omega, g}\right| = \prod_{q \text{ cycle of } g} f\left(x_1^{|q|}, x_2^{|q|}, x_3^{|q|}, \ldots\right)$$

$$= f(x_1, x_2, \ldots)^{j_1(g)} f\left(x_1^2, x_2^2, \ldots\right)^{j_2(g)} \cdots f(x_1^n, x_2^n, \ldots)^{j_n(g)}$$

Substituting this in the sum over all g yields the substituted cycle index as claimed.

13.2.1 Applications

Polya's theorem is exclusively convenient because it will act as a picture function - actually producing a polynomial that determines what the different configurations are, and how many of each exist. It has numerous applications such as chemical isomer enumeration, graph theory, and music theory. Here are some important applications.

Colored Cubes

How many ways are there to color the sides of a three-dimensional cube with m colors, up to rotation of the cube? The rotation group C of the cube acts on the six sides of the cube, which are equivalent to beads. Its cycle index is

$$Z_C(t_1, t_2, t_3, t_4) = \frac{1}{24}\left(t_1^6 + 6t_1^2 t_4 + 3t_1^2 t_2^2 + 8t_3^2 + 6t_2^3\right)$$

Which is obtained by analyzing the action of each of the 24 elements of C on the 6 sides of the cube, see here for the details.

We take all colors to have weight 0 and find that there are

$$F(0) = Z_C(m, m, m, m) = \frac{1}{24}\left(m^6 + 3m^4 + 12m^3 + 8m^2\right)$$

different colorings.

Graphs on Three and Four Vertices

The Pólya enumeration theorem can be used to calculate the number of graphs up to isomorphism with a fixed number of vertices, or the generating function of these graphs according to the number of edges they have.

Rooted Ternary Trees

The set T_3 of rooted ternary trees consists of rooted trees where every node (or non-leaf vertex) has exactly three children (leaves or subtrees). Small ternary trees are shown at right. Note that rooted ternary trees with n nodes are equivalent to rooted trees with n vertices of degree at most 3 (by ignoring the leaves). In general, two rooted trees are isomorphic when one can be obtained from the other by permuting the children of its nodes. In other words, the group that acts on the children of a node is the symmetric group S_3. We define the weight of such a ternary tree to be the number of nodes (or non-leaf vertices).

One can view a rooted, ternary tree as a recursive object which is either a leaf or a node with three children which are themselves rooted ternary trees. These children are equivalent to beads; the cycle index of the symmetric group S_3 that acts on them is

$$Z_{S_3}(t_1, t_2, t_3) = \frac{t_1^3 + 3t_1 t_2 + 2t_3}{6}.$$

Check Your Progress

1. Define the Polya's theorems.
2. Explain the unweighted version of Polya's theorems.
3. Analyse the weighted version of Polya's theorems.
4. Elaborate on the applications of Polya's theorems.

13.3 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. This is a most general theorem which allows the number of discrete combinatorial objects of a given types to be enumerated (counted) as a function of their "Order". In 1937 it was independently rediscovered by George Pólya, who then greatly popularized the result by applying it to many counting problems, in particular to the enumeration of chemical compounds.
2. Let X be a finite set and let G be a group of permutations of X (or a finite symmetry group that acts on X). The set X may represent a finite set of beads, and G may be a chosen group of permutations of the beads.

3. In the more general and more important version of the theorem, the colours are also weighted in one or more ways, and there could be an infinite number of colours provided that the set of colours has a generating function with finite coefficients. In the univariate case, suppose that

$$f(t) = f_0 + f_1 t + f_2 t^2 + \cdots$$

is the generating function of the set of colors, so that there are f_w colors of weight w for each integer $w \geq 0$.

4. Polya's theorem is exclusively convenient because it will act as a picture function - actually producing a polynomial that determines what the different configurations are, and how many of each exist. It has numerous applications such as chemical isomer enumeration, graph theory, and music theory. Here are some important applications.

13.4 SUMMARY

- This is a most general theorem which allows the number of discrete combinatorial objects of a given types to be enumerated (counted) as a function of their "Order". In 1937 it was independently rediscovered by George Pólya, who then greatly popularized the result by applying it to many counting problems, in particular to the enumeration of chemical compounds.
- Let X be a finite set and let G be a group of permutations of X (or a finite symmetry group that acts on X). The set X may represent a finite set of beads, and G may be a chosen group of permutations of the beads.
- In the more general and more important version of the theorem, the colours are also weighted in one or more ways, and there could be an infinite number of colours provided that the set of colours has a generating function with finite coefficients. In the univariate case, suppose that

$$f(t) = f_0 + f_1 t + f_2 t^2 + \cdots$$

is the generating function of the set of colors, so that there are f_w colors of weight w for each integer $w \geq 0$.

- Polya's theorem is exclusively convenient because it will act as a picture function - actually producing a polynomial that determines what the different configurations are, and how many of each exist. It has numerous applications such as chemical isomer enumeration, graph theory, and music theory. Here are some important applications.

NOTES

13.5 KEY WORDS

- **Polya's theorem:** This is a most general theorem which allows the number of discrete combinatorial objects of a given types to be enumerated (counted) as a function of their "Order". In 1937 it was independently rediscovered by George Pólya, who then greatly popularized the result by applying it to many counting problems, in particular to the enumeration of chemical compounds.
- **Applications:** Polya's theorem is exclusively convenient because it will act as a picture function - actually producing a polynomial that determines what the different configurations are, and how many of each exist. It has numerous applications such as chemical isomer enumeration, graph theory, and music theory. Here are some important applications.

13.6 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. State the Polya's theorems.
2. Illustrate the unweighted version of Polya's theorems.
3. Explain the weighted version of Polya's theorems.
4. Define the applications of Polya's theorems.

Long-Answer Questions

1. Briefly define the Polya's theorems.
2. Describe the unweighted version of Polya's theorems.
3. Analyse the applications of Polya's theorems.

13.7 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

NOTES

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.

NOTES

UNIT 14 BINARY OPERATIONS AND PERMUTATION GROUPS

Structure

14.0 Introduction
14.1 Objectives
14.2 Binary Operations
  14.2.1 Binary Relation
  14.2.2 Properties of Binary Relations
  14.2.3 Closures of Relation
14.3 Permutation Groups
  14.3.1 Permutation Group
  14.3.2 Even and Odd Permutation
  14.3.3 Alternating Group
14.4 Answers to Check Your Progress Questions
14.5 Summary
14.6 Key Words
14.7 Self Assessment Questions and Exercises
14.8 Further Readings

14.0 INTRODUCTION

In mathematics, a binary operation or dyadic operation is a calculation that combines two elements (called operands) to produce another element. More formally, a binary operation is an operation of arity two.

More specifically, a binary operation on a set is an operation whose two domains and the codomain are the same set. Examples include the familiar arithmetic operations of addition, subtraction, multiplication. Other examples are readily found in different areas of mathematics, such as vector addition, matrix multiplication and conjugation in groups.

An operation of arity two that involves several sets is sometimes also called a binary operation. For example, scalar multiplication of vector spaces takes a scalar and a vector to produce a vector, and scalar product takes two vectors to produce a scalar. Such binary operations may be called simply binary functions.

Binary operations are the keystone of most algebraic structures that are studied in algebra, in particular in semigroups, monoids, groups, rings, fields, and vector spaces.

A permutation group is a group G whose elements are permutations of a given set M and whose group operation is the composition of permutations in G (which are thought of as bijective functions from the set M to itself). The group of all permutations of a set M is the symmetric group of M, often written as Sym(M).

The term permutation group thus means a subgroup of the symmetric group. If $M = \{1,2,...,n\}$ then, Sym(M), the symmetric group on n letters is usually denoted by S_n.

By Cayley's theorem, every group is isomorphic to some permutation group. The way in which the elements of a permutation group permute the elements of the set is called its group action. Group actions have applications in the study of symmetries, combinatorics and many other branches of mathematics, physics and chemistry.

In this unit, you will study the basic concept of binary operations and permutation group.

14.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe the basic concept of binary operations
- Explain the permutation group

14.2 BINARY OPERATIONS

A binary operation or dyadic operation is a calculation that combines two elements (called operands) to producer another element. More formally, a binary operation is an operation of arity two.

14.2.1 Binary Relation

A binary relation R from a set A to a set B is a subset R of the Cartesian product $A \times B$.

For example,

1. Let $A = B = N$, the set of natural numbers.
   (i) Define the relation R as '='
   Now, $R = \{ (1, 1), (2, 2), (3, 3),....\} \subseteq N \times N$
   $\therefore$ R is a binary relation.
   (ii) Define R as '<'
   Then, $R = \{ (1, 2), (2, 3), (3, 4),..... (1, 3), (2, 4), (3, 5), ...\} \subseteq N \times N$
   $\therefore$ R is a binary relation.
2. Let A be the set of all people on earth and $a, b \in A$, and $a\ R\ b$ iff a and b were born on the same day of the same year.

NOTES

Domain and the Range of a Relation

Let R be a binary relation. The set $D(R)$ of all elements x such that, for all y, $(x, y) \in R$ is called the domain of R.

i.e., $D(R) = \{ x : (x, y) \in R, \text{ for all } y\}$

Similarly, $Rg(R)$ of all elements y such that, for all x, $(x, y) \in R$ is called the range of R.

i.e., $Rg(R) = \{y : (x, y) \in R, \text{ for all } x\}$

Operations on Relations

Let R and S be relations from a set A to a set B. Now the union and intersection of R and S is defined as follows:

(i) $R \cup S = \{ (a, b) : (a, b) \in R \text{ or } (a, b) \in S\}$

(ii) $R \cap S = \{(a,b) : (a, b) \in R \text{ and } (a, b) \in S\}$

Example 14.1: Let $X = \{1, 2, 3, 4, 5, 6\}$

Let R and S be relations from X to X as,

$R = \{ (x, y) : (x + y) \text{ is a multiple of } 2\}$

$S = \{ (x, y) : (x + y) \text{ is a multiple of } 3\}$

Find $R \cup S$ and $R \cap S$.

Solution: $R = \{(1, 3), (1, 5)\}$ and $S = \{(2, 4), (1, 5)\}$

$R \cup S = \{(1, 3), (2, 4), (1, 5)\}$

$R \cap S = \{(1, 5)\}$

Inverse of R: Let R be a relation from a set A to set B. The inverse of R is relation from B to A and is given by $R^{-1} = \{(y, x): (x, y) \in R\}$.

Representation of a Solution

(i) A binary relation R from a set A with n elements to a set B with m elements is represented as an $n \times m$ array M_R by marking the positions in M_R. The positions which correspond to the pairs belong to R with 1 and 0 elsewhere.

$$\text{i.e., } M_R = [a_{ij}] \begin{cases} 1 \text{ if } i\text{th element of } A \text{ is related to } j\text{th element of } B \\ 0, \text{ otherwise.} \end{cases}$$

Example 14.2: Let $A = B = X = \{1, 2, 3, 4, 5, 6\}$. Define R as '<' on X.

Solution: $R = \{(1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 3), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5), (3, 6), (4, 5), (4, 6), (5, 6)\}$.

NOTES

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

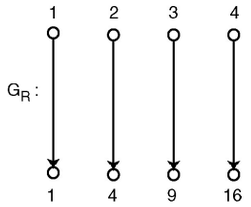
(ii) The relation array can be viewed graphically as elements of sets represented by models and an ordered pair is represented by an edge between the vertices that correspond to the paired elements, with an arrow pointing to the second element of the pair.

Example 14.3: Let $A = \{1, 2, 3, 4\}$, $B = \{1, 4, 9, 16\}$ and the relation $R = \{(1, 1), (2, 4), (3, 9), (4, 16)\}$. Draw the relation graph.

Solution: First we shall write the relation matrix M_R:

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

Now we shall draw the relation graph G_R.

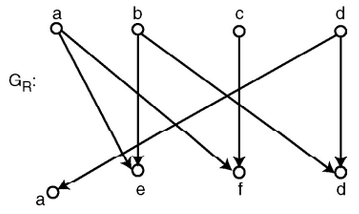


Example 14.4: Let $A = \{a, b, c, d\}$, $B = \{a, e, f, d\}$ and let $R = \{(a, e), (a, f), (b, e), (c, f), (b, d), (d, d), (d, a)\}$. Draw the relation graph.

Solution: First we shall write the relation matrix M_R:

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

The relation graph G_R is given as follows:



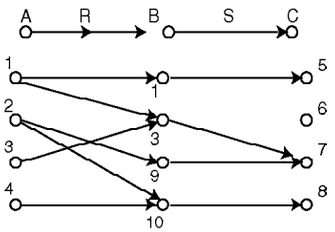
NOTES

Composition of Two Relations

Let R be a binary relation from the set A to the set B and S be a binary relation from the set B to the set C, then, the ordered pair (R, S) is said to be composable. If (R, S) is a composable pair of binary relations, the composite $R \circ S$ and R and S, is a binary relation from the set A to the set C, such that , for $a \in A$ and $c \in C$, $a\ (R \circ S)c$ if for some $b \in B$, both aRb and bSc are binary relations.

Example 14.5: $A = \{1, 2, 3, 4\}$, $B = \{1, 3, 9, 10\}$ $C = \{5, 6, 7, 8\}$, $R = \{(1, 1), (1, 3), (2, 9), (2, 10), (3, 3), (4, 10)\}$ $S = \{(1, 5), (3, 7), (9, 7), (10, 8)\}$. Find $R \circ S$ and its relation graph.

Solution:

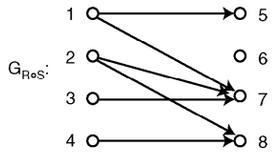


$R \circ S = \{(1, 5), (1, 7), (2, 7), (2, 8), (3, 7), (4, 8)\}$

The corresponding matrix is as follows:

$$M_{R \circ S} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and the corresponding relation graph $G_{R \circ S}$ is as follows:



14.2.2 Properties of Binary Relations

In mathematics, a binary relation is defined as a relation on a set A which is a collection of ordered pairs of elements of A. It can also be defined as a subset of the Cartesian product $A^2 = A \times A$. Basically, a binary relation between two sets A and B is a subset of $A \times B$. The term 2-place relation is synonym for binary relations. A binary relation is the exceptional case $n = 2$ of an n-ary relation $R \subseteq A_1 \times \ldots \times A_n$ which is a set of n-tuples where the jth component of each n-tuple is taken from the jth domain A_j of the relation. In certain specific systems of axiomatic set theory, relations can be extended to classes which are generalized from sets. In set theory, this extension is required to model the concepts of 'Is An Element Of' or 'Is A Subset Of'. The following are the properties of binary relations:

NOTES

Let R be a relation on a set A (i.e., $R \subseteq A \times A$). R is called:

(i) **Reflexive:** If $aRa, \forall\ a \in A$.

(ii) **Symmetric:** If aRb then $bRa\ \forall\ a, b \in A$.

(iii) **Transitive:** If aRb and bRc then $aRc\ \forall\ a,b, c \in A$.

(iv) **Irreflexive:** If $a \not R a$, and $a \in A$.

(v) **Antisymmetric:** If aRb then $b \not R a$, and $a = b$ for $a, b \in A$.

(vi) **Connected:** A relation R in A is connected iff for every two *distinct* elements x and y in A, $\langle x,y \rangle \in R$ or $\langle y,x \rangle \in R$ or both.

Based on these properties, other combinations can be defined using certain classes of relations such as equivalence, tolerance or ordering.

- **Equivalence:** A relation R on a set A is called an **equivalence** relation if R is reflexive, symmetric and transitive. Examples of equivalence relations are a set of parallel lines in a plane.
- **Tolerance:** A relation R in $A \times A$ is called a **tolerance** or a tolerance relation if it is reflexive and symmetric. Tolerance is weaker than equivalence. The notion of tolerance relation is an explication of similarity or closeness.
- **Ordering:** An **ordering** is a binary relation which is basically transitive and further it can either be reflexive and antisymmetric or irreflexive and asymmetric.

14.2.3 Closures of Relation

Let R be any relation on a set A. R may be or may not be symmetric, reflexive and transitive. Let S be any other relation on A such that S contains R, and S is a subset of every relation containing R. Then S is called closure of R.

Reflexive Closure

Let R be a relation on a set X. The reflexive closure S of R is obtained by adding to R all pairs of the form (a,a) which are not in R, $a \in A$. Now S is reflexive, contains R and is in any reflexive relation containing R.

Example 14.6: $R = \{(a, a), (a, b), (b, a), (c, b)\}$ on the set $A = \{a, b, c\}$. Find the reflexive closure of R.

Solution: By inspection, we can see that $(b, b), (c, c)$ are not in R.

$\therefore\ S = \{(a, a), (a, b), (b, a), (c, b), (b, b), (c, c)\}$

Then, S is the reflexive closure of R.

Note: To obtain a reflexive closure of a relation R, just add the diagonal relation elements to R, i.e., Let $D = \{(a,c) / a \in A\}$, the diagonal relation S is the reflexive closure of R, then $S = R \cup D$.

NOTES

Symmetric Closure

Let R be a relation on set A. The symmetric closure S of R is obtained by adding to R all pairs of the form (b, a), if $(a, b) \in R$ and $(b, a) \notin R$.

In other words, symmetric closure of R can be obtained as $S = R \cup R^{-1}$

Example 14.7: Let $R = \{(a, a), (a, b), (b, b), (c, c), (c, b)\}$ on $\{a, b, c\}$. Find the symmetric closure of R.

Solution: Clearly R is not symmetric.

Now $S = R \cup R^{-1} = \{(a, a), (a, b), (b, b), (c, a), (c, b), (b, a), (a, c), (b, c)\}$

Clearly S is symmetric and it contains R and is in any symmetric relation containing R.

14.3 PERMUTATION GROUPS

Let S be a finite set of order n and let $f : S \to S$ such that f is one-one and onto then f is known as a **permutation** of degree n. Two permutations say, f and g of degree n are said to be **equal permutation** if for all $a \in S$, we have $f(a) = g(a)$. If S is a finite set having n elements then there are $n!$ distinct permutations of degree n and the set containing all these permutations is $P_n = \{f : f$ is a permutation of degree $n\}$. If $e : S \to S$ such that for all $a \in S$, we have $e(a) = a$, then e is called an **identity permutation**. Also, $fe = ef = f$.

Composition of Permutations

If f and g are two permutations, then their composition is denoted as fog and is defined as first carrying out the operation defined by f and then the operation defined by g. For example, let $f = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix}$ and $g = \begin{pmatrix} b_1 & b_2 & \cdots & b_n \\ c_1 & c_2 & \cdots & c_n \end{pmatrix}$.

Then $fog = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix} \begin{pmatrix} b_1 & b_2 & \cdots & b_n \\ c_1 & c_2 & \cdots & c_n \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ c_1 & c_2 & \cdots & c_n \end{pmatrix}$. Here, f replaces a_1 by b_1, g replaces b_1 by c_1 so that fog replaces a_1 by c_1 and so on.

14.3.1 Permutation Group

The set of all the permutations of degree n is denoted by P_n and binary operation is composition of permutation, then P_n is a group known as the **permutation group**.

(i) Closure: If $f, g \in P_n \Rightarrow fog \in P_n$ [$\because$ Composition of permutations is also a permutation]

(ii) Associative: For all $f, g, h \in P_n$ we have $(fog)\ o\ h = f\ o\ (goh)$

NOTES

(iii) Existence of Identity: For each $f \in P_n$ we have $e \in P_n$ such that $foe = eof$.

$\therefore e$ is the identity element.

(iv) Existence of Inverse: For each $f \in P_n$ we have $f^{-1} \in P_n$ such that $fof^{-1} = e = f^{-1}of$.

Hence, P_n is a group.

14.3.2 Even and Odd Permutation

A permutation is called **even** permutation if it can be expressed as the product of even number of transpositions and it is called **odd** permutation if it can be expressed as the product of odd number of transpositions.

Properties of Even and Odd Permutations

(i) Any permutation is either even or odd.

(ii) A cycle of length n will be an even (odd) permutation if n is odd (even).

(iii) Product of two even permutations is an even permutation and the product of two odd permutations is an even permutation.

(iv) Product of an even and odd permutation is always odd permutation.

(v) Inverse of an even (odd) permutation is even (odd) permutation.

(vi) Out of the total $n!$ permutations on n symbols, half the permutations i.e. $\frac{n!}{2}$ are even and half are odd.

14.3.3 Alternating Group

The set consisting of all the even permutations is called the **alternating set** of degree n and is denoted by A_n.

Theorem 14.1: The set A_n of all even permutations of degree n is a group of order $\frac{n!}{2}$ with respect to multiplication.

Proof: Closure: Let $f, g \in A_n \Rightarrow f$ and g are even permutation. Also, the product of even permutations is even permutation.

$\therefore fog \in A_n \Rightarrow$ Closure is satisfied

Associative: Multiplication of permutations is always associative.

Existence of Identity: There exists identity permutation which is always even such that $If = f = fI$ for all $f \in A_n$. Hence, I is the identity of A_n.

Existence of Inverse: Inverse of an even permutation is always even permutation.

Hence, A_n is a group with respect to multiplication of groups of order $\frac{n!}{2}$.

NOTES

Example 14.8: Let S = {1, 2, 3, 4, 5}. If $f = (2\ \ 3)$ and $g = (1\ \ 4)$, show that $fog = gof$.

Solution: We have, $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 2 & 4 & 5 \end{pmatrix}$ and $g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 3 & 1 & 5 \end{pmatrix}$.

Then, $fog = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 2 & 4 & 5 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 3 & 1 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 1 & 5 \end{pmatrix}$ … (i)

and $gof = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 3 & 1 & 5 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 2 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 1 & 5 \end{pmatrix}$ … (ii)

From (i) and (ii), we get $fog = gof$.

Check Your Progress

1. What do you mean by binary operations?
2. State the difference between equivalence and tolerance.
3. Define permutations groups.
4. What are the properties of even and odd permutations?

14.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. A binary operation or dyadic operation is a calculation that combines two elements (called operands) to producer another element. More formally, a binary operation is an operation of arity two.
2. Equivalence: A relation R on a set A is called an equivalence relation if R is reflexive, symmetric and transitive. Examples of equivalence relations are a set of parallel lines in a plane.

   Tolerance: A relation R in A × A is called a tolerance or a tolerance relation if it is reflexive and symmetric. Tolerance is weaker than equivalence. The notion of tolerance relation is an explication of similarity or closeness.
3. Let S be a finite set of order n and let $f: S \to S$ such that f is one-one and onto then f is known as a permutation of degree n. Two permutations say, f and g of degree n are said to be equal permutation if for all $a \in S$ we have $f(a) = g(a)$.

If S is a finite set having n elements then there are $n!$ distinct permutations of degree n and the set containing all these permutations is $P_n = \{f : f$ is a permutation of degree $n\}$. If $e : S \to S$ such that for all $a \in S$, we have $e(a) = a$, then e is called an identity permutation. Also, $fe = ef = f$.

4. Properties of Even and Odd Permutations:
   (i) Any permutation is either even or odd.
   (ii) A cycle of length n will be an even (odd) permutation if n is odd (even).
   (iii) Product of two even permutations is an even permutation and the product of two odd permutations is an even permutation.
   (iv) Product of an even and odd permutation is always odd permutation.
   (v) Inverse of an even (odd) permutation is even (odd) permutation.
   (vi) Out of the total $n!$ permutations on n symbols, half the permutations.

14.5 SUMMARY

- A binary operation or dyadic operation is a calculation that combines two elements (called operands) to producer another element. More formally, a binary operation is an operation of arity two.
- A binary relation R from a set A to a set B is a subset R of the Cartesian product $A \times B$.
- A binary relation R from a set A with n elements to a set B with m elements is represented as an $n \times m$ array M_R by marking the positions in M_R. The positions which correspond to the pairs belong to R with 1 and 0 elsewhere.
- Let R be a binary relation from the set A to the set B and S be a binary relation from the set B to the set C, then, the ordered pair (R, S) is said to be composable. If (R, S) is a composable pair of binary relations, the composite $R \circ S$ and R and S, is a binary relation from the set A to the set C, such that , for $a \varepsilon A$ and $c \varepsilon C$, $a(R \circ S)c$ if for some $b \varepsilon B$, both aRb and bSc are binary relations.
- A relation R in A is connected iff for every two distinct elements x and y in A, $\langle x,y \rangle \varepsilon R$ or $\langle y,x \rangle \varepsilon R$ or both.
- Equivalence: A relation R on a set A is called an equivalence relation if R is reflexive, symmetric and transitive. Examples of equivalence relations are a set of parallel lines in a plane.
- Tolerance: A relation R in A × A is called a tolerance or a tolerance relation if it is reflexive and symmetric. Tolerance is weaker than equivalence. The notion of tolerance relation is an explication of similarity or closeness.
- An ordering is a binary relation which is basically transitive and further it can either be reflexive and antisymmetric or irreflexive and asymmetric.

- Let R be any relation on a set A. R may be or may not be symmetric, reflexive and transitive. Let S be any other relation on A such that S contains R, and S is a subset of every relation containing R. Then S is called closure of R.
- Let R be a relation on a set X. The reflexive closure S of R is obtained by adding to R all pairs of the form (a,a) which are not in R, $a \in A$. Now S is reflexive, contains R and is in any reflexive relation containing R.
- A permutation is called even permutation if it can be expressed as the product of even number of transpositions and it is called odd permutation if it can be expressed as the product of odd number of transpositions.
- Product of two even permutations is an even permutation and the product of two odd permutations is an even permutation.
- Product of an even and odd permutation is always odd permutation.
- Inverse of an even (odd) permutation is even (odd) permutation.
- The set consisting of all the even permutations is called the alternating set of degree n and is denoted by A_n

14.6 KEY WORDS

- **Equivalence:** A relation R on a set A is called an equivalence relation if R is reflexive, symmetric and transitive. Examples of equivalence relations are a set of parallel lines in a plane.
- **Tolerance:** A relation R in $A \times A$ is called a tolerance or a tolerance relation if it is reflexive and symmetric. Tolerance is weaker than equivalence. The notion of tolerance relation is an explication of similarity or closeness.
- **Ordering:** An ordering is a binary relation which is basically transitive and further it can either be reflexive and antisymmetric or irreflexive and asymmetric.
- **Permutation group:** The set of all the permutations of degree n is denoted by P_n and binary operation is composition of permutation, then P_n is a group known as the permutation group.

14.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What are binary operations?
2. What do you understand by permutation groups?

NOTES

Long-Answer Questions

1. What do you mean by binary relations? Discuss the properties of binary relations giving examples.
2. What is permutation group? Discuss in detail about the properties of even and odd permutations giving appropriate examples.

14.8 FURTHER READINGS

Krishnamurthy, V. 1985. *Combinatorics: Theory and Applications*. New Delhi: Affiliated East-West Press.

Anderson, Ian. 1974. *A First Course in Combinatorial Mathematics (Oxford Applied Mathematics and Computing Science Series)*. Oxford (UK): Clarendon Press (OUP).

Balakrishnan, V. K. 1996. *Schaum's Outline of Combinatorics: Including Concepts of Graph Theory*. New York: McGraw-Hill Education.

Liu, C. L. 1968. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Iyengar, N Ch S N, V M Chandrasekaran, K A Venkatesh and P S Arunachalam. 2007. *Discrete Mathematics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Gries, David and Fred B Schneider. 1994. *A Logical Approach to Discrete Math*. New York: Springer-Verlag.