POST GRADUATE DEGREE PROGRAMME (CBCS) IN

MATHEMATICS

SEMESTER IV

SELF LEARNING MATERIAL

PAPER : COR 4.1
(Pure and Applied Streams)

Discrete Mathematics
Probability and Statistical Methods

Directorate of Open and Distance Learning
University of Kalyani
Kalyani, Nadia
West Bengal, India

Content Writers

Block I Discrete Mathematics	Ms. Audrija Choudhury Assistant Professor Department of Mathematics Directorate of Open and Distance Learning University of Kalyani
Block II Probability and Statistical Methods	Dr. Arun Sarkar Assistant Professor Department of Mathematics Heramba Chandra College

April, 2024

Directorate of Open and Distance Learning, University of Kalyani

Published by the Directorate of Open and Distance Learning

University of Kalyani, 741235, West Bengal

All rights reserved. No part of this work should be reproduced in any form without the permission in writing from the Directorate of Open and Distance Learning, University of Kalyani.

Director's Message

Satisfying the varied needs of distance learners, overcoming the obstacle of Distance and reaching the unreached students are the three fold functions catered by Open and Distance Learning (ODL) systems. The onus lies on writers, editors, production professionals and other personnel involved in the process to overcome the challenges inherent to curriculum design and production of relevant Self Learning Materials (SLMs). At the University of Kalyani a dedicated team under the able guidance of the Hon'ble Vice-Chancellorhas invested its best efforts, professionally and in keeping with the demands of Post Graduate CBCS Programmes in Distance Mode to devise a self-sufficient curriculum for each course offered by the Directorate of Open and Distance Learning (DODL), University of Kalyani.

Development of printed SLMs for students admitted to the DODL within a limited time to cater to the academic requirements of the Course as per standards set by Distance Education Bureau of the University Grants Commission, New Delhi, India under Open and Distance Mode UGC Regulations, 2020 had been our endeavor. We are happy to have achieved our goal.

Utmost care and precision have been ensured in the development of the SLMs, making them useful to the learners, besides avoiding errors as far as practicable. Further suggestions from the stakeholders in this would be welcome.

During the production-process of the SLMs, the team continuously received positive stimulations and feedback from Professor **(Dr.) Amalendu Bhunia, Hon'ble Vice-Chancellor, University of Kalyani**, who kindly accorded directions, encouragements and suggestions, offered constructive criticism to develop it with in proper requirements. We gracefully, acknowledge his inspiration and guidance.

Sincere gratitude is due to the respective chairpersons as well as each and every member of PGBOS (DODL), University of Kalyani. Heartfelt thanks are also due to the Course Writers-faculty members at the DODL, subject-experts serving at University Post Graduate departments and also to the authors and academicians whose academic contributions have enriched the SLMs. We humbly acknowledge their valuable academic contributions. I would especially like to convey gratitude to all other University dignitaries and personnel involved either at the conceptual or operational level of the DODL of University of Kalyani.

Their persistent and coordinated efforts have resulted in the compilation of comprehensive, learner-friendly, flexible texts that meet the curriculum requirements of the Post Graduate Programme through Distance Mode.

Self Learning Materials (SLMs) have been published by the Directorate of Open and Distance Learning, University of Kalyani, Kalyani-741235, West Bengal and all the copyrights reserved for University of Kalyani. No part of this work should be reproduced in any from without permission in writing from the appropriate authority of the University of Kalyani.

All the Self Learning Materials are self written and collected from e-book, journals and websites.

Director

Directorate of Open and Distance Learning

University of Kalyani

Post Graduate Board of Studies
Department of Mathematics
Directorate of Open and Distance Learning
University of Kalyani

Sl No.	Name & Designation	Role
1	Dr. Samares Pal, Professor & Head Department of Mathematics, University of Kalyani	Chairperson
2	Dr. Pulak Sahoo, Professor Department of Mathematics, University of Kalyani	Member
3	Dr. Sahidul Islam, Associate Professor Department of Mathematics, University of Kalyani	Member
4	Dr. Sushanta Kumar Mohanta, Professor Department of Mathematics, West Bengal State University	External Nominated Member
5	Ms. Audrija Choudhury, Assistant Professor Department of Mathematics Directorate of Open and Distance Learning University of Kalyani	Member
6	Director Directorate of Open and Distance Learning University of Kalyani	Convener

Core Paper

COR 4.1
(Pure and Applied Streams)
Marks : 100 (SEE : 80; IA : 20)

Discrete Mathematics(Marks : 60 (SEE: 50; IA: 10))
Probability and Statistical Methods (Marks : 40 (SEE: 30; IA: 10))

Syllabus

- **Unit 1:** Definition of graphs, circuits, cycles, Subgraphs, induced subgraphs, degree of a vertex, Connectivity.
- **Unit 2:** Trees, Euler's formula for connected graphs, Spanning trees, Complete and complete bipartite graphs.
- **Unit 3:** Planar graphs and their properties, Fundamental cut set and cycles. Matrix representation of graphs.
- **Unit 4:** Kuratowski's theorem (statement only) and its use, Chromatic index, chromatic numbers and stability numbers.
- **Unit 5:** Lattices as partial ordered sets. Their properties. Lattices as algebraic system. Some special Lattices e.g. complete complemented and distributed lattices.
- **Unit 6:** Sublattices. Direct products and Homomorphism.
- **Unit 7:** Boolean Algebra Basic Definitions, Duality, Basic theorems, Boolean algebra as lattices.
- **Unit 8:** Boolean functions, Sum and Product of Boolean algebra, Minimal Boolean Expressions, Prime implicants Propositions and Truth tables.
- **Unit 9:** Logic gates and circuits, Applications of Boolean Algebra to Switching theory (using AND, OR, and NOT gates), Karnaugh Map method.
- **Unit 10:** Combinatorics: Introduction, Basic counting principles, Permutation and combination, pigeonhole principle, Recurrence relations and generating functions.
- **Unit 11:** Grammar and Language: Introduction, Alphabet, Words, Free semi group, Languages, Regular expression and regular languages, Grammars.
- **Unit 12:** Finite Automata (FA).
- **Unit 13:** Finite State Machine. Non-deterministic and deterministic FA. Push Down Automation (PDA), Equivalence of PDAs and Context Free Languages (CFLs).

- **Unit 14:** Computable Functions.
- **Unit 15:** Fields and σ-fields of events. Probability as a measure. Random variables. Probability distribution.
- **Unit 16:** Expectation. Moments. Moment inequalities, Characteristic function. Convergence of sequence of random variables-weak convergence, strong convergence and convergence in distribution, continuity theorem for characteristic functions. Weak and strong law of large numbers. Central Limit Theorem.
- **Unit 17:** Definition and classification of stochastic processes. Markov chains with finite and countable state space, classification of states.
- **Unit 18:** Statistical Inference, Estimation of Parameters, Minimum Variance Unbiased Estimator, Method of Maximum Likelihood for Estimation of a parameter.
- **Unit 19:** Interval estimation, Method for finding confidence intervals, Statistical hypothesis, Level of significance; Power of the test.
- **Unit 20:** Analysis of variance, One factor experiments, Linear mathematical model for ANOVA.

Contents

Unit 1

Course Structure

- Definition of graphs, circuits, cycles
- Subgraphs, induced subgraphs, degree of a vertex
- Connectivity

1.1 Introduction

In the time of Euler, in the town of Königsberg in Prussia, there was a river containing two islands. The islands were connected to the banks of the river by seven bridges (1.1.1). The bridges were very beautiful, and on their days off, townspeople would spend time walking over the bridges (see figure below). As time passed, a question arose: was it possible to plan a walk so that you cross each bridge once and only once? This is known as the Knigsberg seven bridge problem. In the year 1736, Euler represented the problem as a graph and answered the question in negative. This marked the birth of graph theory.

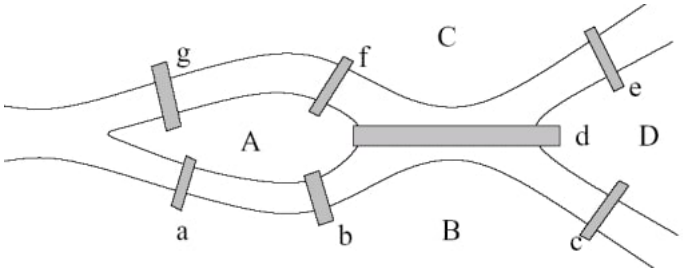


Figure 1.1.1: Königsberg Seven Bridge Problem

Since then it has blossomed in to a powerful tool used in nearly every branch of science and is currently an active area of mathematics research. Over the past 200 years, graph theory has been used in a variety of applications. Graphs are used to model electric circuits, chemical compounds, highway maps, and many more. They are also used in the analysis of electrical circuits, finding the shortest route, project planning, linguistics, genetics and social science.

Objectives

After reading this unit, you will be able to

- define graphs, vertex, edges
- define circuits, cycles and learn their properties
- define subgraphs
- define connected and disconnected graphs and learn their properties
- define planar graphs and trees
- deduce the Euler's formula for connected graphs

1.2 Graphs

Definition 1.2.1. A graph G is a triple (V, E, g), where

1. V is a finite non-empty set, called the set of vertices;
2. E is a finite set (may be empty), called the set of edges;
3. g is a function, called the incidence function, that assigns to each edge, $e \in E$ a one element subset $\{v\}$, or a two-element subset $\{u, v\}$, where u, v are vertices.

For convenience, we will write $g(e) = \{u, v\}$, where v and u may be same in which case we write $g(e) = \{v\}$.

Let $G = (V, E, g)$ be a graph. Suppose e be an edge of this graph. Then there are vertices u and v such that $g(e) = \{u, v\}$; the vertices u and v are called the end vertices or the endpoints of the vertex e. When a vertex v is an endpoint of an edge e, we say that e is incident with vertex v and v is incident with the edge e. Two vertices are said to be adjacent if there exists an edge $e \in E$ such that $g(e) = \{u, v\}$. Two edges e and f are said to be adjacent if they have a common endpoint, that is, if $g(e) = \{u, v\}$ and $g(f) = \{v, w\}$. If e is an edge such that $g(e) = \{u, v\}$ such that $u = v$, that is, $g(e) = \{u\}$, then e is an edge from u to itself, or u is adjacent to itself and such an edge e is called a loop on the vertex u.

From now on, we will simply write the graph $G = (V, E, g)$ as G.

Example 1.2.2. Let $V = \{a, b, c, d\}$ and $E = \{e, f, h, i, j\}$ and g is defined as

$$g(e) = \{a, b\},\ \ g(f) = \{b, c\},\ \ g(h) = \{c, d\},\ \ g(i) = \{d, a\},\ \ g(j) = \{d, b\}.$$

Thus, $G = (V, E, g)$ is a graph. We can also write the above definition of g as follows:

e	f	h	i	j	k
$\{a, b\}$	$\{b, c\}$	$\{c, d\}$	$\{d, a\}$	$\{d, b\}$	$\{d, b\}$

Such a representation of the incidence function g is the incidence table whose columns are indexed by the edges. The vertices adjacent to an edge are placed in the second row below the edge.

In this example, we see that the edge e is incident on the two vertices a and b. Thus, the vertex a and b are adjacent. Similarly, we see that the edges e and f are adjacent since the vertex b is common for both.

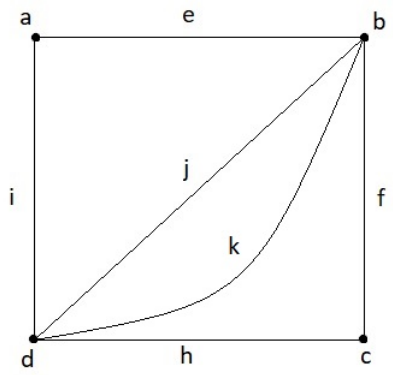


Figure 1.2.1: G as in example 1.2.2

The set of vertices and the set of edges of a graph are finite. Thus, one of the features that make the study of graphs easy and interesting is that they can be represented pictorially. That is, the corresponding diagram for a graph helps us to visualize the facts easily. If we represent the graph in the above example pictorially, then we get something as depicted in the figure 1.2.1.

The incidence function need not be one-to-one. There may be more than one edge having the same end-points. Such edges are called parallel edges. We formally define parallel edges as follows.

Definition 1.2.3. Let $G = (V, E, g)$ be a graph. Two edges e and f are said to be parallel if $g(e) = g(f) = \{u, v\}$ for $u, v \in V$.

In the previous example, the edges j and k are parallel edges since $g(j) = g(k) = \{d, b\}$. This can easily be seen from the figure.

Definition 1.2.4. Let G be a graph and v be a vertex in G. We call v as isolated vertex if it is not incident with any edge, or, v is not an endpoint of any edge.

Definition 1.2.5. Let G be a graph and v be a vertex in G. Then the degree of v is defined as the number of edges incident with v. It is written as $\deg(v)$ or $d(v)$. By convention, it is considered that each loop contributes 2 to the degree of a vertex.

Note that for an isolated vertex v, we will always have $d(v) = 0$. In fact, this is a necessary and sufficient condition for a vertex to be isolated.

Example 1.2.6. $G = (V, E, g)$ is a graph (see figure 1.2.2 where $V = \{A, B, C, D\}$ and $E = \{e, f, h, i, j\}$, where g is defined as

e	f	h	i	j
$\{A, B\}$	$\{B, C\}$	$\{C, B\}$	$\{B, A\}$	$\{A, A\}$

Then, we can see that D is an isolated vertex. Also, $d(A) = 4$, $d(B) = 4$, $d(C) = 2$ and $d(D) = 0$. e and i are parallel edges and f and h are also so. Notice that $g(A) = g(i)$ irrespective of the order in which A and B are written in the incidence table of g. But it is not the case always (as we will study in case of the directed graphs). These graphs that we are studying now are also called undirected graphs (or simply, graphs).

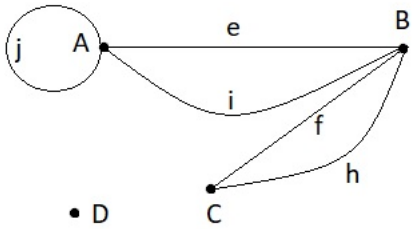


Figure 1.2.2: G in example 1.2.6

Exercise 1.2.7. Represent the following graphs pictorially and find the degree of each of its vertices. Also state the parallel vertices and loops, if any.

1. $V = \{v_1, v_2, \ldots, v_7\}$ and $E = \{e_1, e_2, \ldots, e_7\}$ where g is defined as

e_1	e_2	e_3	e_4	e_5	e_6	e_7
$\{v_1, v_2\}$	$\{v_1, v_2\}$	$\{v_4, v_3\}$	$\{v_6, v_3\}$	$\{v_2, v_4\}$	$\{v_6, v_3\}$	$\{v_6, v_3\}$

2. $V = \{v_1, v_2, v_3, v_4\}$ and $E = \{e_1, e_2, e_3\}$ where g is defined as

e_1	e_2	e_3
$\{v_1, v_2\}$	$\{v_3, v_3\}$	$\{v_4, v_3\}$

The graphs in which all the vertices are of the same degree are called the regular graphs. The two examples of the graphs we have seen so far are not regular graphs (verify it for example 1.2.2).

Definition 1.2.8. Let G be a graph and k be a non-negative integer. Then G is called a k-regular graph if the degree of each vertex of G is k.

An interesting k regular graph is the Petersen 3-regular graph as shown in the figure.

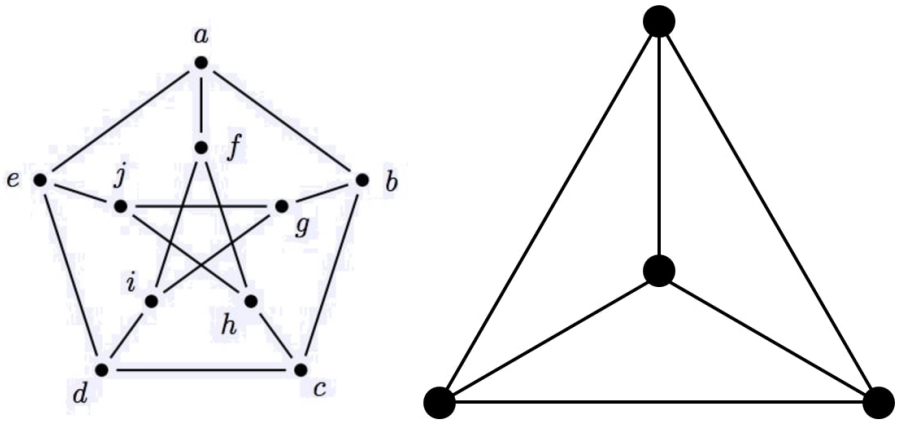


Figure 1.2.3: 3-regular graphs (Petersen 3 regular graph on the left)

Definition 1.2.9. Let G be a graph and v be a vertex of G. v is called an even degree vertex if $d(v)$ is an even number. Similarly, v is odd degree vertex if $d(v)$ is odd.

The Petersen 3-regular graph has every vertex an odd vertex.

Definition 1.2.10. Let $n_1, n_2, \ldots, n_k$ be the degrees of vertices of a graph G such that $n_1 \leq n_2 \leq \cdots \leq n_k$. Then the finite sequence $n_1, n_2, \ldots, n_k$ is called the degree sequence of the graph.

Clearly, every graph has a unique degree sequence. But, we can construct completely different graphs having the same degree sequence.

Exercise 1.2.11.

1. State the even and odd vertices of the graphs in exercise 1.2.7. Also find the degree sequence of them.
2. Construct a 1-regular graph having 3 vertices.
3. Construct two different graphs having the same degree sequence.

Consider the degree sequence of the graph in 1.2.2 which is, 2, 2, 4, 4 and adding them gives $(2+2+4+4 =)12$, which is an even number. In fact, the sum of the degrees of all the vertices is always an even number which is given in the following theorem due to Euler.

Theorem 1.2.12. The sum of the degrees of all the vertices of a graph is twice the number of edges.

Proof. Let G be a graph with n edges and m vertices, say $v_1, v_2, \ldots, v_m$. We want to determine

$$d(v_1) + d(v_2) + \cdots + d(v_m).$$

Now the degree, $d(v_i)$, of v_i is the number of edges incident with v_i. Each edge e is either a loop or incident with two distinct vertices. If e is a loop on a vertex v, then e contributes 2 to the degree of v. On the other hand, if e is incident with two distinct vertices v and w, then e contributes 1 to the degree of each vertex. Thus we find that when we compute the sum of the degrees, each edge contributes 2 to the sum. Because there are n edges, the total contribution to the above sum is $2n$. Hence

$$d(v_1) + d(v_2) + \cdots + d(v_m) = 2n.$$

□

Corollary 1.2.13. The sum of the degrees of all the vertices of a graph is an even integer.

Proof. Since $2n$ is an even integer, the corollary follows from the previous theorem. □

Corollary 1.2.14. In a graph, the number of odd degree vertices is even.

Proof. Suppose a graph G has k odd vertices, $v_1, v_2, \ldots, v_k$, and t even degree vertices, $u_1, u_2, \ldots, u_t$. Thus, by the above corollary,

$$d(v_1) + d(v_2) + \cdots + d(v_k) + d(u_1) + d(u_2) + \cdots + d(u_t) = 2n,$$

where n is the number of edges. Because each $d(u_j)$ is even, it follows that $d(u_1) + d(u_2) + \cdots + d(u_t)$ is an even integer. Also, $2n$ is even. Hence, $d(v_1) + d(v_2) + \cdots + d(v_k)$ must also be even. Now, the sum of odd number of odd integers is an odd integer. Because each number $d(v_i)$ is an odd number and $d(v_1) + d(v_2) + \cdots + d(v_k)$ is even, it follows that the number k cannot be odd. So k is even and this completes the proof. □

1.3 Directed Graphs

Definition 1.3.1. A directed graph, or digraph G is a triple (V, E, g) such that

1. V is a finite non-empty set of vertices;

2. E is a finite set (may be empty) of directed edges or arcs;

3. $g : E \to V \times V$ is a function, that assigns to each edge, $e \in E$ an ordered pair (u, v), where u, v are vertices (u and v may be same).

We can represent a digraph pictorially. The only difference between the representation of graph and digraph is in the directed edges which are drawn with arrows representing the starting and terminating vertices.

If $g(e) = (u, v)$, then u is called the starting vertex and v is called the terminating vertex of the arc e. The in-degree of a vertex v is the number of arcs with v as the terminating vertex and the out-degree of v is the number of arcs with v as the starting vertex. In computing in-degree and out-degree of a vertex, we assume that each loop contributes 1 to the in-degree and 1 to the out-degree of v.

Theorem 1.3.2. In any digraph $G = (V, E, g)$, the following three numbers are equal:

1. The sum of the in-degrees of all the vertices;

2. The sum of the out-degrees of all the vertices;

3. The number of arcs.

Proof. The proof is similar to that of theorem 1.2.12. We just consider the fact that each arc e with starting vertex u and terminating vertex u contributes 1 to the out-degree and 1 to the in-vertex of v. The details are left as exercise. □

Example 1.3.3. Let G be a digraph such that $V = \{a, b, c, d\}$, $E = \{e, f, h\}$ and $g(e) = (a, a)$, $g(f) = (b, c)$ and $g(h) = (b, d)$. The diagram is as follows:

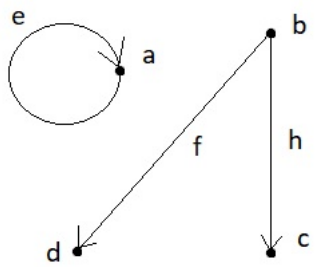


Figure 1.3.1: G in example 1.3.3

The in-degrees of a, b, c and d are 1, 0, 1 and 1 respectively and the out-degrees are 1, 2, 0 and 0 respectively. Then, sum of the in-degrees of all the vertices=sum of the out-degrees of all the vertices=number of arcs=3.

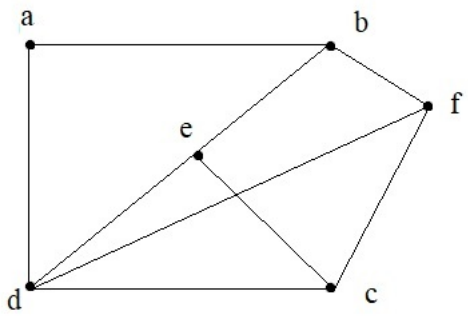


Figure 1.4.1: A simple graph

1.4 Simple Graphs

Definition 1.4.1. Let G be a graph. Then it is called a simple graph if it does not contain any parallel edges or loop.

The graph in the figure 1.4.1 has no loop or parallel edges.

Theorem 1.4.2. Let G be a simple graph with at least two vertices. Then G has at least two vertices of same degree.

Proof. Let G be a simple graph with $n \geq 2$ vertices. G has no loops or parallel edges. Thus, the degree of a vertex v is the same as the number of vertices adjacent to it. The graph G has n vertices. Thus, a vertex v has at most $n-1$ adjacent vertices, because v is not adjacent to itself. Hence, for any vertex v, the degree of v is one of integers: $0, 1, 2, \ldots, n-1$.

We now show that if there exists a vertex v such that $d(v) = 0$, then for each vertex u of G, $d(u) < n-1$. On the contrary, suppose that in G, v is a vertex with degree 0 and u is a vertex with degree $n-1$. Then v is an isolated vertex and u has $n-1$ adjacent vertices. Because G is a simple graph, u is not adjacent to itself. From this and the fact that G is simple and $d(u) = n-1$, it follows that every vertex of G other than u is adjacent to u. This implies that v is adjacent to u, which is a contradiction since v is an isolated vertex. This proves our claim.

In a similar manner, we can prove that if there exists a vertex v in G such that the degree of v is $n-1$, then for each vertex u in G, $d(u) > 0$.

We now conclude that the degree of all the vertices of G are either in the set $\{0, 1, 2, \ldots, n-2\}$ or in the set $\{1, 2, \ldots, n-1\}$.

Let $v_1, v_2, \ldots, v_n$ be the n vertices of G. Then, either for all of $i = 1, 2, \ldots, n$, $d(v_i) \in \{0, 1, 2, \ldots, n-2\}$ or $d(v_i) \in \{1, 2, \ldots, n-1\}$. Thus, by the pigeonhole principle, there exists i and j, $1 \leq i \leq n$ $1 \leq j \leq n$, $i \neq j$, such that $d(v_i) = d(v_j)$. Hence there are atleast two vertices of same degree. □

Remark 1.4.3. The converse of the above theorem is not true in general. For example, a and c have equal degree in example 1.2.2, but the graph G is not simple.

Definition 1.4.4. A simple graph with n vertices in which there is an edge between every pair of distinct vertices is called a complete graph on n vertices. This is denoted by K_n.

A complete graph of three vertices is a triangle.

Theorem 1.4.5. The number of edges in a complete graph with n vertices is $\frac{n(n-1)}{2}$.

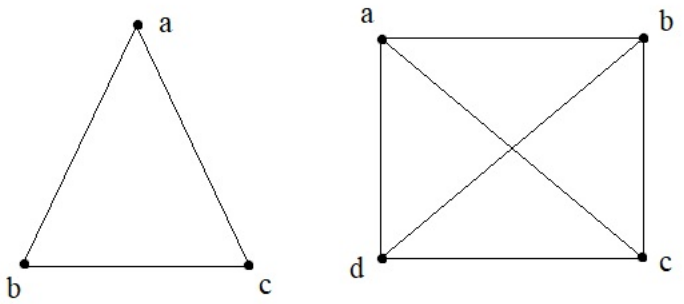


Figure 1.4.2: K_3 and K_4

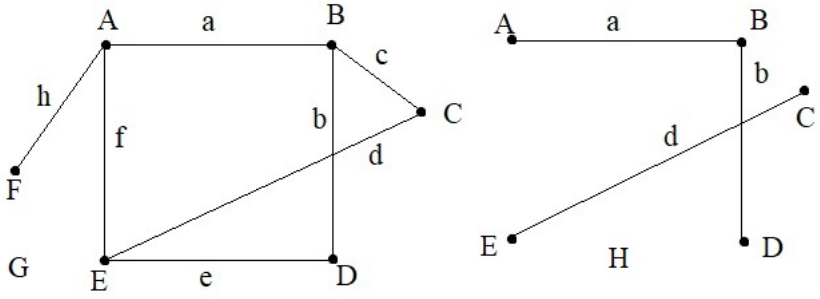
Proof. Let G be a complete graph with n vertices. Then G is a simple graph such that there exists an edge between any two distinct vertices. Hence, for any vertex v of G, each of the remaining $n-1$ vertices is adjacent to v. Hence the degree of each vertex is $n-1$. Also, since G has n vertices, so the sum of the degree of all the vertices is $n(n-1)$. We know that the sum of the degree of all the vertices is 2 times the number of edges. Let the number of edges be m. So, we have, $n(n-1) = 2m$ and thus, we get,

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

□

1.5 Subgraph

Consider the graph $G = (V, E, g)$ and $H = (V_1, E_1, g_1)$ such that $V = \{A, B, C, D, E, F\}$, $E = \{a, b, c, d, e, f, h\}$ and $V_1 = \{A, B, C, D, E\}$, $E_1 = \{a, b, d\}$ and g and g_1 are as shown in the figure. It is worthy to note



that $V_1 \subset V$ and $E_1 \subset E$. Also, g_1 is the function g restricted over E_1. Such a graph H is called a subgraph of G. We will now formally define subgraphs.

Definition 1.5.1. Let $G = (V, E, g)$ be a graphs. A graph $H = (V_1, E_1, g_1)$ is called a subgraph of G if V_1 is a non-empty subset of V and E_1 is a subset of E and g_1 is a restriction of g on E_1 such that for all $e \in E_1$, we have $g_1(e) = g(e) = \{u, v\}$ for $u, v \in V_1$.

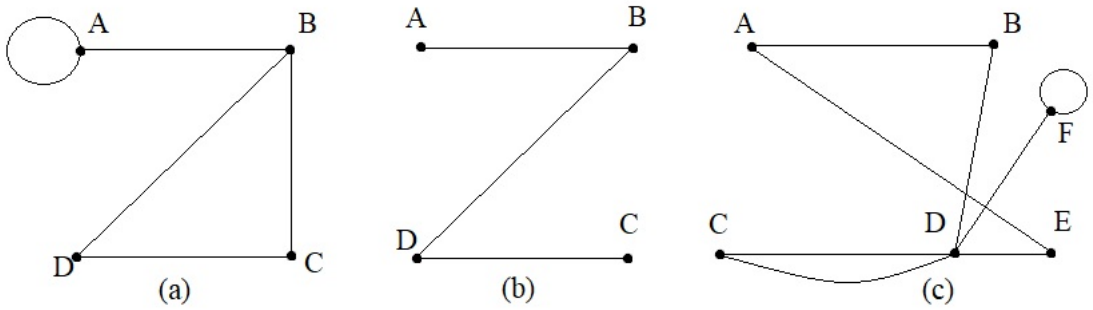
Remark 1.5.2. Let $G = (V, E)$ be a graph and $H = (V_1, E_1)$ be a subgraph of G. From the previous definition, it follows that if $e \in E_1$, and u, v are the end vertices of e in G, then $u, v \in V_1$.

Let G be a graph with vertex set V and edge set E. Suppose that V contains more than one vertex. Then for any vertex $v \in V$, $G \setminus \{v\}$ denotes the subgraph whose vertex set is $V_1 = V \setminus \{v\}$ and the edge set is $E_1 = \{e \in E |\ v \text{ is not an end vertex of } e\}$. Then $G \setminus \{v\}$ is called a subgraph obtained from G by deleting the vertex v.

Let $e \in E$, and $G \setminus \{e\}$ denote the subgraph whose edge set is $E \setminus \{e\}$ and the vertex set is $V_1 = V$. Then $G \setminus \{e\}$ is the subgraph obtained by deleting the edge e.

Remark 1.5.3. $G \setminus \{v\}$ is obtained by deleting the vertex v and at the same time deleting all the edges incident with v. However, the graph $G \setminus \{e\}$ is obtained from G by deleting only the edge e without deleting any of the vertices of G.

Exercise 1.5.4. 1. Determine which of the following graphs are simple:



2. Draw a graph having the following properties and explain why no such graph exists:
   (a) Simple graph, five vertices, each of degree 2
   (b) Simple graph having degree sequence $3, 3, 3, 3, 4$
   (c) Six edges and having the degree sequence $1, 2, 3, 4, 6$
3. Find three subgraphs of G in the figure 1.5.1 with at least four vertices and six edges:
4. How many vertices are there in a graph with 20 edges if each vertex is of degree 5?
5. Does there exist a simple graph with degree sequence $1, 2, 3, 4, 5$? Justify.
6. Does there exist a graph with five edges and degree sequence $1, 2, 3, 4$?

1.6 Walks, Path, Cycles, Circuits

Definition 1.6.1. Let u and v be two vertices in a graph G. A walk from u to v in G, is an alternating sequence of $n + 1$ vertices and n edges of G

$$(u = v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n, e_n, v_{n+1} = v)$$

beginning with vertex u, called the initial vertex, and ending with vertex v, called the terminal vertex, in which v_i and v_{i+1} are endpoints of edge e_i, for $i = 1, 2, \ldots, n$.

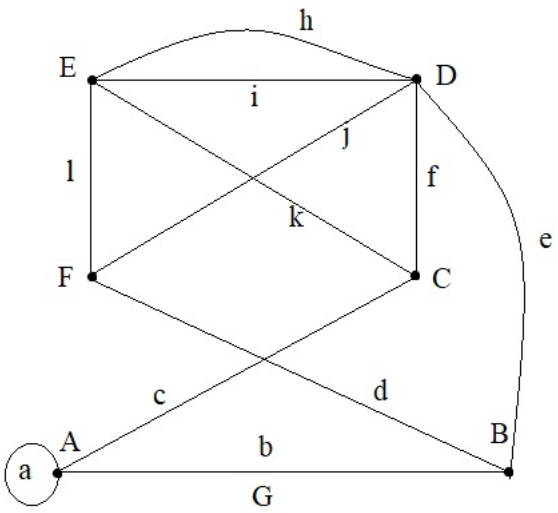


Figure 1.5.1

Definition 1.6.2. Let u and v be two vertices in a digraph G. A directed walk from u to v in G, is an alternating sequence of $n+1$ vertices and n arcs of G

$$(u = v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n, e_n, v_{n+1} = v)$$

beginning with vertex u and ending with vertex v, in which each e_i is an arc from v_i to v_{i+1} for $i = 1, 2, \ldots, n$.

Definition 1.6.3. The length of a walk(or a directed walk) is the total number of occurrences of edges(or, arcs) in the walk(or, directed walk). A walk or directed walk of length zero is only a vertex.

A walk (or, directed walk) from a vertex u to v in a graph (or, digraph) G is also called a $u - v$ walk (or, directed walk). If u and v are the same, then $u - v$ walk (or, directed walk) is called a closed walk (or, directed walk). Otherwise, it is called an open walk (or, directed walk).

Definition 1.6.4. A walk with no repeated edges is called a trail, and a walk with no repeated vertices except possibly the initial and terminal vertices is called a path.

Thus, from the previous definitions, it is clear that in a path, no edge can be repeated. Hence, every path is a trail, but not conversely.

Definition 1.6.5. A walk, path, or trail is called trivial if it has only one vertex and no edges. A walk, path, or trail that is not trivial is called nontrivial.

Definition 1.6.6. A nontrivial closed trail from a vertex u to itself is called a circuit.

Hence, a circuit is a closed walk of nonzero length from a vertex u to itself with no repeated edges.

Example 1.6.7. Consider the graph in figure 1.6.1. In this graph

$$(A, a, B, b, C, f, E, e, B, d, D)$$

is a walk of length 5. It is an open walk from A to D. This is a walk with no repeated edges. Hence this walk is a trail since B appears twice. But

$$(B, b, C, f, E, i, D, j, G)$$

is a path of length 4 from vertex B to G.

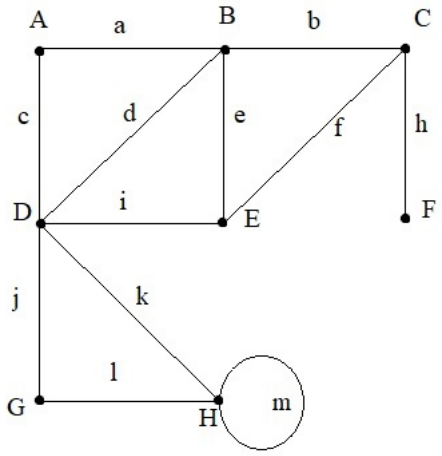


Figure 1.6.1

Definition 1.6.8. A circuit that does not contain any repetition of vertices except the starting and terminal vertices is called a cycle. A cycle of length k is called a k-cycle. A cycle is called even (odd) if it contains an even (odd) number of edges.

It follows from definition that a 3-cycle is a triangle.
Directed walks, trails, paths, circuits, cycles are defined analogously.

Definition 1.6.9. Let $P = (v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n)$ be a walk in a graph G. A subwalk of P is a subsequence of **consecutive** entries $Q = (v_i, e_i, v_{i+1}, e_{i+1}, \ldots, v_{k-1}, e_{k-1}, v_k)$, $1 \leq i \leq k \leq n$, that begins at a vertex and ends at a vertex.

From the definition, it follows that every subwalk is a walk.
Let $P = (v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n)$ be a walk in a graph G and $Q = (v_i, e_i, v_{i+1}, e_{i+1}, \ldots, v_{k-1}, e_{k-1}, v_k = v_i)$ be a closed subwalk of P. If we delete this subwalk Q from P except for the vertex v_i, then we obtain a new walk. This walk is denoted by $P - Q$ and is called the reduction of P by Q.

Theorem 1.6.10. Let G be a graph and u, v be two vertices of G. If there is a walk from u to v, there is a path from u to v.

Proof. Let $P = (u = v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n = v)$ be a walk. If $u = v$, then this is a closed walk. In this case, (u) from u to u consisting of a single vertex and no edge. Suppose $P = (u = v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n = v)$ is an open walk. If this is not a path, then $v_i = v_j$ for some $1 \leq i < j \leq n$. This shows that there is a closed subwalk Q from v_i to v_j. We reduce P to $P - Q$. Now, $P - Q$ is a new walk from u to v. If this walk is not a path, we repeat this deletion process of subwalks. Because the number of closed subwalks in P is finite, we eventually obtain a path from u to v. □

We can also follow the proof of the above theorem and deduce an analogous result for circuit.

Theorem 1.6.11. Every circuit contains a subwalk that is a cycle.

Proof. Let T be a circuit. Let S be the collection of all closed nontrivial subwalks of T. Because $T \in S$, S is nonempty. Now S is a finite set. Thus we can find a member of S of minimum length. Let T_1 be a nontrivial closed subwalk $(u = v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n = u)$ of T of minimum length. Since T_1 is of minimum length, T_1 cannot contain a nontrivial closed subwalk other than T_1. This implies that T_1 has no repeated vertices except the vertex u. Hence T_1 is a cycle. □

Definition 1.6.12. Let G be a graph. A vertex u is said to be connected to a vertex v of G if there is a $u-v$ walk in graph G. And G is said to be connected if for any two vertices u and v of G, there is a $u-v$ walk in G, otherwise G is called a disconnected graph.

We can show that a graph G is connected if and only if for any two vertices $u, v \in G$, there is a $u-v$ path in G. We assume that a graph with only a single vertex and no edges is connected.

We now define a relation R on the vertex set V of a graph G as

$$R = \{(u, v) \in V \times V : \text{ there is a } u - v \text{ walk in } G\}.$$

Since the trivial walk (u) is a $u-u$ walk in G, R is reflexive. Suppose there is a $u-v$ walk $(u = v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n = v)$. Then, $(v = v_n, e_{n-1}, v_{n-1}, \ldots, v_2, e_1, v_1 = u)$ is a $v-u$ walk in G. Thus, R is symmetric. Again, suppose there is a $u-v$ walk

$$(u = v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n = v)$$

from a vertex u to v. Also, suppose there is a $v-w$ walk $(v = u_1, f_1, u_2, f_2, \ldots, v_m = w)$ from vertex v to another vertex w. Then clearly,

$$(u = v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n = v = u_1, f_1, u_2, f_2, \ldots, v_m = w)$$

is a walk from vertex u to vertex w. Thus, the relation R is transitive. Hence R partitions the vertex set V into disjoint equivalence classes. Let V_1 be an equivalence class of R and E_1 be the set of edges joining the vertices in V_1 in the graph G. Then $G_1 = (V_1, E_1)$ is a subgraph of G. In this subgraph, we see that any two vertices are connected. This subgraph is called a component of G.

Definition 1.6.13. A subgraph H of a graph G is called a component of G if

1. any two vertices of H are connected in H, and
2. H is not properly contained in any connected subgraph of G.

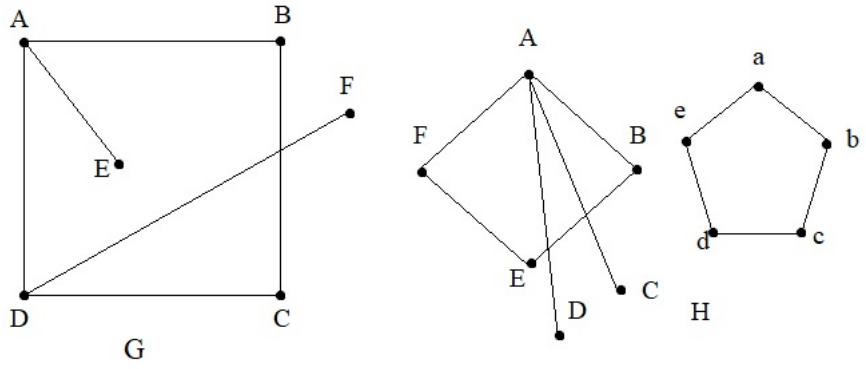


Figure 1.6.2: Connected and Disconnected Graphs

Graph G in the above figure has only one component, which is G itself. Graph H, on the other hand, has two components with vertices $\{A, B, C, D, E, F\}$ and $\{a, b, c, d, e\}$.

From the definition, it follows that any component of a graph is always connected. Now, every equivalence class of the equivalence relation R gives a component of G. Hence, every graph can be partitioned into finite number of components. It follows that a graph G is connected if and only if G has only one component.

Theorem 1.6.14. A connected graph with n vertices has at least $n-1$ edges.

Proof. We prove the result by induction on n. If $n = 1$, then the result is trivially true. Assume that any connected graph with n vertices has at least $n-1$ edges. Consider a connected graph G with $n+1$ vertices. Because G is a connected graph, the degree of each vertex of G is ≥ 1. Suppose the degree of each vertex of G is ≥ 2. Then the sum of the degree of the vertices of G is $\geq 2(n+1) > 2n$. Thus, the number of edges of G is $> n$. Suppose now that G has a vertex v of degree 1. We construct a graph G_1 by deleting the vertex v and the edge incident with v. The graph G_1 is a connected graph with n vertices. By the induction hypothesis, the number of edges of G_1 is at least $n-1$. Therefore, the number of edges of G is at least n. Thus, the result is true for a graph with $n+1$ vertices. Hence, by induction for any connected graph with n vertices, the number of edges is at least $n-1$. □

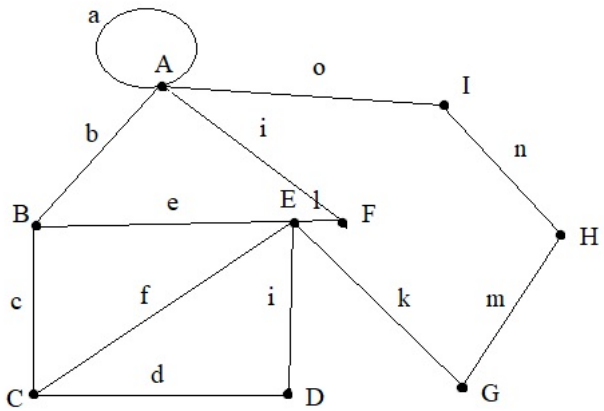
We prove another interesting theorem for connected graphs.

Theorem 1.6.15. Let G be a simple graph with at most $2n$ vertices. If the degree of each vertex is at least n, then the graph is connected.

Proof. Suppose that G is not connected. Then G can be partitioned into components $C_1, C_2, \ldots, C_m$, $m \geq 2$. Since the degree of each vertex of G is at least n and the graph is simple, we find that each vertex has at least n adjacent vertices. Then each component contains at least $n+1$ vertices. This implies that the number of vertices of G is at least $m(n+1) \geq 2(n+1) > 2n$. This contradiction implies that the given graph is connected. □

Few Probable Questions

1. Consider the graph below:



   (a) Find an open walk of length 4. Is is a trail? Is your walk a path?

   (b) Find a closed walk of length 5. Is your walk a circuit?

2. Does there exist a graph with 20 edges if each vertex is of degree 3?

3. Draw s simple graph such that every vertex is adjacent to two vertices and every edge is adjacent to two edges.

4. Define a path of a graph G. If G has exactly two vertices of odd degree, then show that there exists a path between these two vertices.

5. Define simple graph. If there is a trail from a vertex u to another vertex v of a graph G, then show that there is a path from u to v.

6. Define connected graph. Show that a simple graph with n vertices and m components can have at most $\frac{(n-m)(n-m+1)}{2}$ edges.

7. Let G be a connected graph with at least two vertices. If the number of edges in G is less than the number of vertices, then prove that G has a vertex of degree 1.

Unit 2

Course Structure

- Trees, Euler's formula for connected graphs, Spanning trees
- Complete and complete bipartite graphs

2.1 Introduction

In the previous unit, we learnt about the basic definitions of graph theory and certain properties related to them. This unit is a continuation of the previous unit.

Objectives

After reading this unit, you will be able to

- define complete graphs, bipartite graphs, and complete bipartite graphs
- define trees and spanning trees
- learn various properties of connected graphs due to Euler

2.2 Bipartite graphs

Definition 2.2.1. A simple graph G is called a bipartite graph if the vertex set V of G can be partitioned into nonempty subsets V_1 and V_2 such that each edge of G is incident with one vertex in V_1 and one vertex in V_2. $V_1 \cup V_2$ is called a bipartition of G.

In the figure 2.2.1, the graph in (a) is a bipartite graph with partition $\{A\}$ and $\{B, C, D\}$. Whereas, the second graph is not bipartite as we can easily verify. (Verify!)

Definition 2.2.2. A bipartite graph G with bipartition $V_1 \cup V_2$ is called a complete bipartite graph on m and n vertices if the subsets V_1 and V_2 contain m and n vertices, respectively, such that there is an edge between each pair of vertices $v_1 \in V_1$ and $v_2 \in V_2$. A complete bipartite graph with m and n vertices is denoted by $K_{m,n}$.

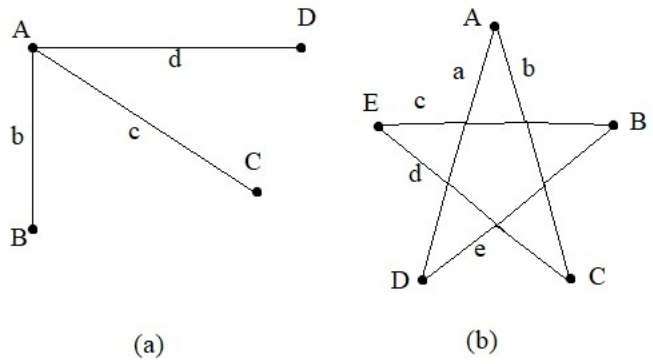
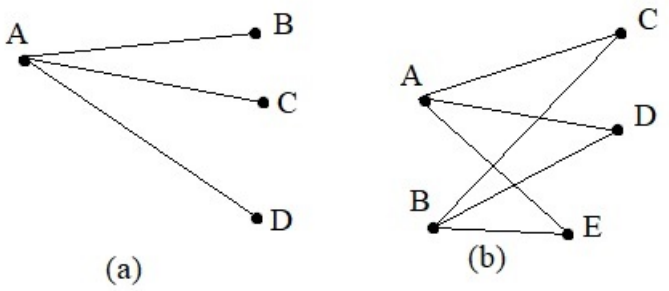


Figure 2.2.1



The two graphs in the above figure represents two complete bipartite graphs. (a) is $K_{1,3}$ while (b) is $K_{2,3}$. Note that the number of edges in the graph $K_{m,n}$ is mn.

Definition 2.2.3. Let G be a graph. Then the distance between two vertices u, v of G, written as $d(u, v)$, is the length of a shortest path, if any exists, from u to v.

If G is a connected graph, then we can prove that

1. $d(u, v) \geq 0$, and equality holds if and only if $u = v$;
2. $d(u, v) = d(v, u)$;
3. $d(u, v) + d(v, w) \geq d(u, w)$, for all vertices $u, v, w \in G$.

We will now deduce a necessary and sufficient condition for a graph to be bipartite.

Theorem 2.2.4. A graph is bipartite if and only if it does not contain any cycle of odd length.

Proof. Let $G = (V, E)$ be a bipartite graph with bipartition $V = V_1 \cup V_2$. Now, each edge of G is incident with one vertex in V_1 and one vertex in V_2. Let $(v_1, e_1, v_2, e_2, \ldots, v_k, e_k, v_1)$ be a cycle in G. Because v_i and v_{i+1} are end vertices of e_i, for $i = 1, 2, \ldots, k$ (assuming $v_{k+1} = v_1$), it follows that for $i = 1, 2, \ldots, k$, if $v_i \in V_1$, then $v_{i+1} \in V_2$. Suppose $v_1 \in V_1$. This implies that $v_k \in V_2$. Also it follows that $v_i \in V_1$ if and only if i is odd. Now $v_k \in V_2$, which implies that k is even and hence the length of this cycle is even. This implies that the length of each cycle is even.

Conversely, let G be a graph such that G has no odd cycle. Suppose G is partitioned into components $C_1, C_2, \ldots, C_m$, $m \geq 1$. If we can show that each C_i is a bipartite graph, then G will be also so. We therefore assume that G is connected. Let u be an arbitrary but fixed vector of G. Define the subsets V_1 and V_2 by

$$V_1 = \{v \in V |\ d(u, v) \text{ is even}\} \quad V_2 = \{w \in V |\ d(u, w) \text{ is odd}\}.$$

From our assumption that G is a connected graph, it follows that every vertex of G is either in V_1 or in V_2. Then $\{V_1, V_2\}$ is a partition of V. Because $d(u, u) = 0$, it follows that $u \in V_1$. Let v be an adjacent vertex of u. Then $d(u, v) = 1$. Hence, $v \in V_2$.

Suppose there are two distinct vertices v and w in V_1 and suppose there exists an edge e with v, w as end vertices. Then there is a walk from u to v in G and hence there is a shortest path, say P_1, from u to v. Similarly, we have a shortest path P_2, from u to w. Because v and w belong to V_1, these two shortest paths are of even length. Paths P_1 and P_2 may have several vertices and edges in common.

Now starting from u, let x be the last vertex common to both P_1 and P_2. Let P_1^* be the section of the path of P_1 from u to x and let P_2^* be the section of the path of P_2 from u to x. Because P_1 and P_2 are the shortest paths, P_1^* and P_2^* have equal lengths, which are either both even or both odd. Let $P_1^{'}$ be the part of P_1 from x to v and $P_2^{'}$ be the part of P_2 from x to w. It follows that the lengths of $P_1^{'}$ and $P_2^{'}$ are both either even or odd. Now the walk $P_1^{'}$ followed by e followed by $P_2^{'}$ forms a closed walk C from x to x. Moreover, C does not contain any repetitions of the vertices. Hence C is a cycle. Because the lengths of paths $P_1^{'}$ and $P_2^{'}$ are both even or odd, C must be of odd length, which is a contradiction. Thus, v and w cannot be both in V_1. Similarly, we can show that v and w cannot both belong to V_2. Hence each edge of G connects one vertex of V_1 with one vertex of V_2. Consequently, G is bipartite. □

Exercise 2.2.5. 1. Draw a complete bipartite graph on 3 and 4 vertices.

2. How many edges are there in each of the following graphs

$$(a)\ K_{2,3} \qquad (b)\ K_{4,3} \qquad (c)\ K_{4,4} \qquad (d)\ K_{n,n}$$

3. Prove that a simple graph with a cycle of length 3 can't be a bipartite graph.

2.3 Special Circuits

2.3.1 Euler Circuits

Let us consider a connected graph with more than one vertex such that every vertex has odd degree. For example consider the graph in the figure 2.3.1. It is a connected graph whose every vertex is of odd degree. This graph has no circuit, so it has no circuit that contains all the edges. Also the graph K_4 contains 4 vertices and 6 edges. The degree of each vertex is 3. And this graph also has no circuit consisting of all the edges. But there are circuits consisting of all the edges for some graphs which are of special interest. Let us write the following

Definition 2.3.1. A circuit in a graph that includes all the edges of the graph is called an **Euler Circuit**. And a graph G is called **Eulerian** if either G is trivial graph or G has an Euler circuit.

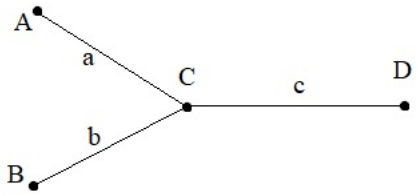


Figure 2.3.1: G with odd vertices

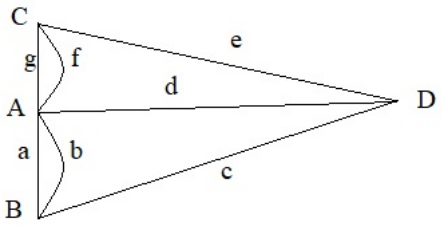
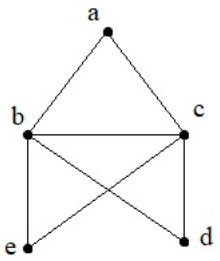


Figure 2.3.2: K*ö*nigsberg Bridge problem

Recall the K*ö*nigsberg bridge problem at the beginning of unit 1. The problem was to determine whether it is possible to take a walk that crosses each bridge exactly once before returning to the starting point. Euler converted this into a problem of graph theory as follows : Each of the islands A, B, C and D are considered as the vertices of a graph and the seven bridges a the seven vertices of the graph. Now the problem reduces to finding a circuit in the graph such that it contains all the edges, or, to find an Euler circuit, or to show that the graph is Eulerian. It is evident from the figure that there does not exist any Euler circuits of the graph.

Example 2.3.2. Consider the graph below.



Each vertex of the above graph are even vertices. In fact, this is a feature of Eulerian graphs as we will soon show.

Theorem 2.3.3. If a connected graph G is Eulerian, then every vertex of G has even degree.

Proof. Suppose that G is Eulerian.

First suppose that G is the trivial graph. Then G has only one vertex v and no edges. Hence the degree of v is 0 which is even.

Next suppose that G contains more than one vertex. Since G is Eulerian, it has an Euler circuit, say

$$C : (v_1, e_1, v_2, e_2, v_3, \ldots, e_{n-1}, v_n = v_1)$$

from a vertex v_1 to $v_n = v_1$. Now, C contains all the vertices (since G is connected) and all the edges of G. However, there are no repeated edges in C, though in C a vertex may appear more than once. Let u be a vertex of G. Since G is connected, u is not an isolated vertex. So u is the end vertex of some edge. Since C contains all the edges, it follows that u is a member of C.

Suppose u is v_1. Let us say that this is the first appearance of u in C. Now, if u is also v_n, we say that v_n is the last appearance of u in C. For each of these two appearances of u, the edge e_1 and the edge e_{n-1} together contribute 2 to the degree of u.

Suppose now u is v_i in C for some i, $1 < i < n$. Then u is an end vertex of the edges e_{i-1} and e_i. These edges together contribute 2 to the degree of u. It now follows that the degree of any vertex in C is even. Hence the degree of any vertex in G is even. □

Suppose G is connected in which every vertex is of even degree. We shall show that G contains an Euler circuit. To do so, we first prove the following lemma.

Lemma 2.3.4. Let G be a connected graph with one or two vertices. If every vertex of G is of even degree, then G has an Euler circuit.

Proof. Suppose G is a graph with only one vertex, say u. Now there may exist zero or more loops at u. However, the number of loops at u must be finite. If there is no loop at u, then (u) is an Euler circuit of G. Also suppose that there are loops $e_1, e_2, \ldots, e_n$, $n \geq 1$, at u. Then $(u, e_1, u, e_2, \ldots, e_n, u)$ is an Euler circuit of G. Hence, G contains an Euler circuit.

Suppose now that G has two vertices u and v such that both are of even degree. Because G is connected, u and v are connected. So there exists an even number of parallel edges between u and v. Let $\{f_1, f_2, \ldots, f_{2k}\}$, $k \geq 1$ be the set of all edges between u and v. Let $e_1, e_2, \ldots, e_n$, $n \geq 0$, be the loops at u and let $g_1, g_2, \ldots, g_m$, $m \geq 0$, be the loops at v. (If $n = 0$, then there are no loops at u. Similarly, if $m = 0$, there are no loops at v). Now,

$$(u, e_1, u, e_2, \ldots, u, e_n, u, f_1, v, g_1, v, g_2, v, \ldots, g_m, v, f_2, u, f_3, v, , f_4, \ldots, f_{2k-1}, v, f_{2k}, u)$$

is a trail that begins at u, traverses all the loops incident with u, traverses one edge from u to v, traverses all the loops at v, then traverses one edge from v to u, and then traverses all the edges between u and v. This trail does not contain any repeated edges. Hence, it is a circuit from u to u. Because this circuit contains all the edges of G, it follows that the graph G has an Euler circuit. □

Theorem 2.3.5. Let G be a connected graph such that every vertex of G is of even degree. Then G has an Euler circuit.

Proof. Suppose G has n edges. We prove by induction on the number of edges of G to show that G has an Euler circuit.

Basic Step: Suppose $n = 0$. Because G has no edges, it follows that G has a single vertex, say u. Then (u) is an Euler circuit.

Inductive hypothesis: Let n be a positive integer. Assume that any connected graph with k edges, $0 \leq k < n$, in which every vertex has even degree has an Euler circuit.

Inductive step: Let $G = (V, E)$ be a connected graph with n edges and the degree of each vertex of G is even. If the number of vertices of G is 1 or 2, then by previous lemma, it follows that G has an Euler circuit. So assume that G has at least three vertices.

Since G is connected, there are vertices v_1, v_2, v_3 and edges e_1, e_2 such that v_1, v_2 are the end vertices of e_1, and v_2, v_3 are the end vertices of e_2. Now consider the subgraph $G_1 = (V_1, E_1)$, where $V_1 = V$ and $E_1 = E - \{e_1, e_2\}$. Next we add a new edge e with v_1, v_3 as end vertices to the subgraph and obtain a new graph $G_2 = (V_2, E_2)$, where $V_2 = V$, $E_2 = E_1 \cup \{e\}$.

Notice that the graph G_2 is obtained from G by deleting edges e_1, e_2, but not removing any vertices, and adding a new edge e with end vertices v_1 and v_3.

In G, suppose $\deg(v_1) = r$, $\deg(v_2) = m$, and $\deg(v_3) = t$. Because we deleted edges e_1, e_2 in G, $\deg(v_1) = r - 1$, $\deg(v_2) = m - 2$, and $\deg(v_3) = t - 1$. Now in graph G_2, we add a new edge e with end vertices v_1 and v_3. Hence, in graph G_2, we have $\deg(v_1) = r$, $\deg(v_2) = m - 2$, $\deg(v_3) = t$. While constructing G_1 from G and G_2 from G_1, the other vertices of G were not disturbed; i.e., their degree in G_2 is the same as their degree in G. Thus, it follows that every vertex of G_2 is of even degree.

Now graph G_2 may not be a connected graph. We show that the number of components of G_2 is less than or equal to two.

Since v_1 and v_3 are the end vertices of the edge e in G_2, it follows that v_1 and v_3 belong to the same component of G_2, say C_1. Now, vertex v_2 may not be in C_1. Let C_2 be the component of G_2 that contains v_2. Let v be a vertex of G_2. Then v is also a vertex of G. Since G is a connected graph, there is a path P from v to v_1 in G.

If P contains one of the edges e_1 or e_2, then P cannot be a path from v to v_1 in G_2. Let P_1 be the path in G_2 that is a portion of the path P starting at v whose edges are also in G_2. Path P_1 may terminate at v_1, v_2, or v_3. If P_1 is a path from v to v_1 in G_1, then v and v_1 belong to the same component, C_1. If P_1 ends at v_3, then (P_1, e, v_1) is a path from v to v_1. Hence in this case, v also belongs to the same component, C_1. Suppose P_1 ends at v_2. Then v belongs to component C_2. Thus, any vertex v of G_2 belongs to either C_1 or C_2. Hence, C_2 has one(if $C_1 = C_2$) or two components.

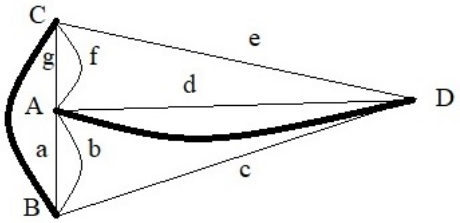
Suppose G_2 has only one component, C_1. Then G_2 is a connected graph with $n - 1$ edges. Thus, by the inductive hypothesis G_2 has an Euler circuit, say T_1. From circuit T_1, we can construct an Euler circuit T in G by simply replacing the subpath (v_1, e, v_3) by the path $(v_1, e_1, v_2, e_2, v_3)$. Hence in this case, we find that G is Eulerian.

Suppose G_2 has two components, C_1 and C_2. Now, each component C_i, $i = 1, 2$ is a connected graph such that each vertex has even degree and the number of edges in C_i is $n_i < n$. Hence, by the inductive hypothesis, C_i has an Euler circuit T_i, $i = 1, 2$. Now, T_1 contains v_1, v_3 and T_2 contains v_2. Hence (v_1, e, v_3) is a subpath of T_1. Moreover, we can assume that T_2 is a circuit from v_2 to v_2.

We now construct an Euler circuit in G by modifying T_1 as follows: In T_1, replace (v_1, e, v_3) by (v_1, e_2, v_2), followed by T_2, followed by (v_2, e_2, v_3). Thus, we find that G has an Euler circuit. The result now follows by induction. □

The above theorem is an effective way of determining when a connected graph is Eulerian.

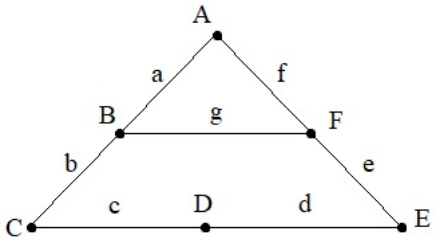
Example 2.3.6. Consider the Königsberg bridge problem. All the vertices in the graph are of odd degree. Then by the preceding two theorems, we can say that there does not exist an Euler circuit for the problem.



But if we add two more edges as shown in the figure, then the resulting graph is Eulerian since every vertex is of even degree.

Definition 2.3.7. An open trail in a graph is called an Euler trail if it contains all the edges.

Example 2.3.8. Consider the following graph. It is a connected graph having two vertices of odd degree. So



it does not have an Euler circuit but the trail $(B, g, F, e, E, d, D, c, C, b, B, a, A, f, F)$ contains all the edges of G. Hence this is an Euler trail.

Theorem 2.3.9. A connected graph G has an open Euler trail if and only if G has only two vertices of odd degree.

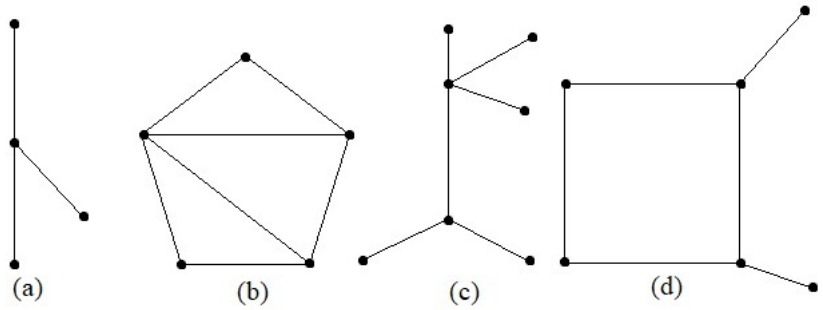
Proof. Suppose G has an open Euler trail P from a vertex u to a vertex v of G. Construct a new graph G_1 by adding a new edge e to G with u and v as the end vertices. In G_1, the trail P with e forms an Euler circuit. Hence every vertex of G_1 is of even degree. In graph G_1, e contributes 1 each to the degree of the vertices u and v. Since G does not contain the edge e, it follows that u and v are the only vertices of odd degree in G.

Conversely assume that a connected graph G has only two vertices u and v of odd degree. Construct a new graph, G_1, by adding a new edge, e, to G with u and v as the end vertices. Then G_1 is a connected graph where every vertex is of even degree. Then G_1 contains an Euler circuit, say P. Now, (u, e, v) is a subpath of P. This subpath is not present in G. Hence, if we delete (u, e, v) from P, then we obtain an open Euler trail P_1 from u to V in G. Hence the theorem is done. □

2.4 Trees

Definition 2.4.1. A graph that is connected and has no cycles is called a tree. Generally, a graph that does not contain any cycles is called an acyclic graph.

Example 2.4.2. Consider the graphs below.



All the graphs are connected. The graphs a and c clearly contains no cycle and hence are trees. Also, the graphs b and d contains cycles and hence are not trees.

Let T be a tree. Then T is a simple connected graph, so T does not have any parallel edges or loops. Let u and v be two vertices in T. It follows that there is at most one edge connecting u and v. Since G is connected,

there is a path from u to v. Let $P = (u, e_1, u_1, e_2, \ldots, u_k, e_k, v)$. If no confusion arises, then we write the path P as $(u, u_1, \ldots, u_k, v)$, that is, when listing the vertices of the path, we will omit the edges.

Theorem 2.4.3. Let u and v be two vertices of a tree T. Then there exists only one path from u to v.

Proof. If $u = v$, then the result is trivial.

Suppose $u \neq v$. Because T is connected, there is at least one path from u to v. Suppose there are distinct paths $P_1 = (u, u_1, \ldots, u_k, v)$ and $P_2 = (u, v_1, \ldots, v_t, v)$ from u to v. Since P_1 and P_2 are distinct, we have the following two cases.

Case 1: $\{u_1, \ldots, u_k\} \cap \{v_1, \ldots, v_t\} = \emptyset$. Then the path P_1 followed by P_2, that is,

$$(u, u_1, \ldots, u_k, v, v_t, \ldots, v_1, u),$$

forms a cycle from u to u, which is a contradiction.

Case 2: $\{u_1, \ldots, u_k\} \cap \{v_1, \ldots, v_t\} \neq \emptyset$. Hence $u_i = v_j$ for some i and j.

Let w_1 be the first common vertex other than u and v, on paths P_1 and P_2. Next, we follow path P_1 until we come to the first vertex w_s, which is again on both P_1 and P_2. This vertex w_s is different from w_1. We must get such a vertex w_s, because P_1 and P_2 meet again at v. Let P_1^* be the portion of the path P_1 from w_1 to w_s and P_2^* be the portion of path P_2 from w_s to w_1. Then, P_1^* followed by P_2^* forms a cycle from w_1 to w_1 in graph T. But this contradicts our assumption that T is a tree, so it has no cycles.

Hence T does not contain two distinct paths between any two distinct vertices u and v.

□

Theorem 2.4.4. In a tree with more than one vertex, there are at least two vertices of degree 1.

Proof. Let T be a tree with more than one vertex. Since T is a connected graph with at least two vertices, there is a path with at least two distinct vertices. Because the number of vertices and the number of edges is finite, the number of paths in T is also finite. Thus we can find a path P of maximal length. Suppose path P is from vertex u to vertex v. We show that $\deg(u) = \deg(v) = 1$.

Suppose $\deg(v) \neq 1$. Let P be the path $(u = v_1, e_1, v_2, e_2, v_3, \ldots, v_{k-1}, e_{k-1}, v)$. Since $\deg(v) \neq 1$, there exists an edge e_k with v as an end vertex such that $e_k \neq e_{k-1}$. Since T has no loops, the other end of e_k can't be v. Suppose the other end is v_k. Suppose $v_k = v_i$ for some i such that $1 \leq i \leq k-1$. Then $(v, e_k, v_i, e_{i+1}, v_{i+1}, \ldots, v_{k-1}, e_{k-1}, v)$ is a cycle from v to v, which contradicts the fact T is a tree. If $v_k \neq v_i$, $1 \leq i \leq k-1$, then we get the path $(v_1, e_1, v_2, e_2, v_3, \ldots, v_{k-1}, e_{k-1}, v, e_k, v_k)$ whose length is greater than that of P. This contradicts the fact that path P is of maximal length in T. It now follows that $\deg(v) = 1$. Similarly, we can show that $\deg(u) = 1$. □

The converse of the above theorem is not true as shown by the following example.

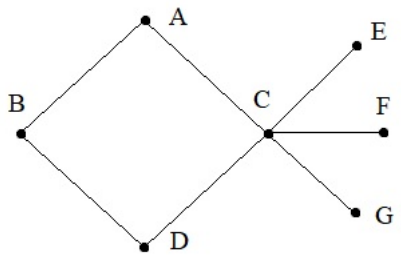
Example 2.4.5. Consider the graph in the given figure.

This is a connected graph and it has at least two vertices of degree 1. But it contains a cycle. Hence it is not a tree.

Theorem 2.4.6. Let T be a tree with n vertices, $n \geq 1$. Then T has exactly $n-1$ edges.

Proof. We prove the result by induction on n.

Basic Step: Let $n = 1$. Since T is a simple graph, it does not contain any loop. Therefore it does not contain any edge and hence the number of edges in T is $0 = 1 - 1$. Hence the theorem is true for $n = 1$.

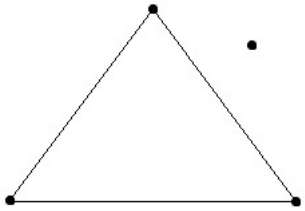


Inductive hypothesis: Let $k \geq 1$ be a positive integer. We assume that the theorem holds for any tree with k vertices.

Inductive step: Let T be a tree with $k+1$ vertices. Since $k+1 \geq 2$, it follows from theorem 2.4.4 that T has at least two vertices of degree 1. Let u be a vertex of degree 1 in T. We construct a new graph G by deleting u from T and also the edge e, which is incident on u. Now, G is still a connected graph and does not contain any cycle. Hence G is a tree with k vertices. By inductive hypothesis, we find that G has exactly $k-1$ edges. This implies that T has k edges. Hence by induction, the theorem holds for any integer n. □

The converse of the above theorem is not true in general. This is proved by the following example.

Example 2.4.7. The given graph is clearly a graph containing 4 vertices and $3 = 4-1$ edges. But this is clearly not a tree. Also, it is not connected.



Theorem 2.4.8. Let T be a graph with n vertices. Then the following are equivalent:

1. T is a tree.
2. T has no loops and if u and v are two distinct vertices in T, then there exists only one path from u to v.
3. T is a connected graph and has $n-1$ edges.
4. T has no cycles and has $n-1$ edges.

2.5 Spanning Tree

We begin with the following definition.

Definition 2.5.1. A tree T is called a spanning tree of a graph G if T is a subgraph of G and T contains all the vertices of G.

Note that the spanning tree of a graph need not be unique. The following theorem gives a necessary and sufficient condition for a graph to have a spanning tree.

Theorem 2.5.2. A graph G has a spanning tree if and only if G is connected.

Proof. Suppose G has a spanning tree, G_1. G_1 contains all the vertices of G. Then between any two vertices, there exists a path in G_1, which is also a path of G. Hence, G is a connected graph.

Conversely, suppose G is a connected graph. If G has no cycles, then G is a tree. Suppose G has cycles. Let C_1 be a cycle in G and e_1 be an edge in C_1. Now construct the graph $G_1 = G \setminus \{e_1\}$, which is obtained by deleting the edge e_1 from G but not removing any vertex from G. Clearly, G_1 is a subgraph of G and it contains all the vertices of G. Because e_1 is an edge of a cycle, G_1 is still a connected graph. If G_1 is acyclic, then G_1 is a tree. If G_1 contains a cycle C_2, then we delete an edge e_2 from C_2 and construct a connected subgraph G_2 that contains all the vertices. If G_2 contains cycles, then we continue this process. Since G has a finite number of edges, it contains only a finite number of cycles. Hence, continuing the process of deleting an edge from a cycle, we eventually obtain a connected subgraph G_k that contains all the vertices of G and is also acyclic. It follows that G_k is a spanning tree of G. □

Exercise 2.5.3.

1. Draw a tree with 9 vertices such that three vertices are of degree 3.
2. How many edges are there in a tree with 16 vertices?
3. How many vertices are there in a tree with 16 edges?
4. Suppose there exists a simple connected graph with 16 vertices that has 15 edges. Does it contain a vertex of degree 1? Justify your answer.

Few Probable Questions

1. Define bipartite graphs. Show that a graph is bipartite if and only if it does not contain any cycle of odd length.
2. Define Euler circuit. Deduce a necessary condition for a connected graph to be Eulerian.
3. Deduce a necessary and sufficient condition for a connected graph G to have an Euler trail.
4. Define a tree. Show that in a tree T, there exits only one path between two vertices of T.
5. Show that in a tree with more than one vertex, there exits at least two vertices of degree 1. Is the converse true? Justify.
6. Show that a tree with n vertices has $n - 1$ edges. Is the converse true? Justify.

Unit 3

Course Structure

- Planar graphs and their properties
- Fundamental cut set and cycles. Matrix representation of graphs

---

3.1 Introduction

The present unit starts with the matrix representation of graphs. We have dealt with two types of matrix representations of graphs, viz., the adjacency matrix and the incidence matrix. The matrix representations are compact and say everything about the graph in a very simple manner as we shall see.

The next topic that is covered is the graph isomorphisms. A graph can exist in different forms having the same number of vertices, edges, and also the same edge connectivity. Such graphs are called isomorphic graphs or “equal" or “same" graphs.

Next we have covered the planar graphs. Such graphs in which the can be drawn in a plane of paper can be thought of as a planar graph and such graphs that don’t satisfy this property is called a non-planar graph. Of particular importance are the connected simple planar graphs from which we can deduce the Kuratowski’s theorem (next unit) that characterises simple non-planar graphs. The proof of this is however excluded.

We have now given a brief idea about all that we are about to study. Let’s explore!

Objectives

After reading this unit, you will be able to

- define incidence matrix and adjacency matrix of a graph
- say when two graphs are said to be same (or, isomorphic)
- define planar graphs and learn related properties
- define planar graphs and related terms like faces, boundaries, etc.
- deduce important results related to planar graphs

3.2 Matrix Representation of a Graph

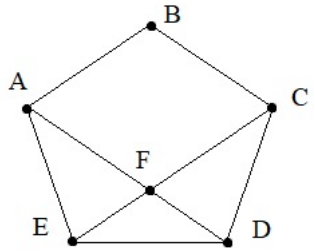
Definition 3.2.1. Let G be a graph with n vertices, where $n > 0$. Let $V(G) = \{v_1, v_2, \ldots, v_n\}$. The **adjacency matrix** A_G with respect to the particular listing, $v_1, v_2, \ldots, v_n$ of n vertices of G is an $n \times n$ matrix $[a_{ij}]$ such that the (i, j)th entry of A_G is the number of edges from v_i to v_j. That is,

$$a_{ij} = \text{number of edges from } v_i \text{ to } v_j.$$

Since a_{ij} is the number of edges from v_i to v_j, the adjacency matrix A_G is a square matrix over the set of non-negative integers.

If G is a digraph, then the adjacency matrix A_G with respect to the particular listing $v_1, v_2, \ldots, v_n$ of n vertices of G is an $n \times n$ matrix $[a_{ij}]$ such that the (i, j)th entry is the number of arcs from a_i to a_j.

Example 3.2.2. Consider the graph G below. The vertices of the graph are $\{A, B, C, D, E, F\}$. Then the



adjacency matrix with respect to this ordering of the vertices is

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

Example 3.2.3. Consider another graph 3.2.1. The vertices of the graph is $\{A, B, C, D\}$. The adjacency

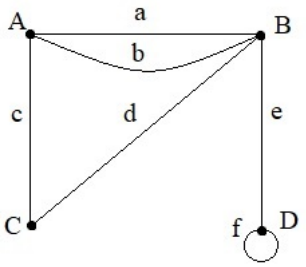


Figure 3.2.1

matrix of the graph with respect to the listing is

$$\begin{bmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

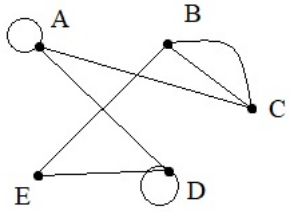
Notice that the matrix A_G is symmetric symmetric since $a_{ij} = a_{ji}$. But, if G is a digraph, then the adjacency matrix need not be symmetric. The adjacency matrix has the following properties:

1. If G does not contain any loops and parallel edges, then each element of A_G is either 0 or 1.
2. If G does not contain any loops, then all the diagonal elements of A_G are 0.

Example 3.2.4. Let A denote 5×5 matrix

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 0 & 1 \\ 1 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

We construct a graph G such that $A_G = A$. For this, we denote the rows by A, B, C, D, E and the columns by A, B, C, D, E. Now we draw a graph with vertices A, B, C, D, E. Since $(1, 1)$ and $(4, 4)$ are the only diagonal elements with entries equal to one, we draw one loop each at the vertices A and D only. Now, we see that $(1, 2)$th element$=$ $(2, 1)$th element $= 0 \Rightarrow$ there is no edge between A and B. Again, $(1, 3)$th element $=$ $(3, 1)$th element $= 1 \Rightarrow$ there exists one edge between A and C. Continuing in this way, we find the following graph



Definition 3.2.5. Let G be a graph with n vertices $v_1, v_2, \ldots, v_n$, where $n > 0$ and m edges $e_1, e_2, \ldots, e_m$. The incidence matrix I_G with respect to the ordering $v_1, v_2, \ldots, v_n$ of vertices and $e_1, e_2, \ldots, e_m$ edges is an $n \times m$ matrix $[a_{ij}]$ such that

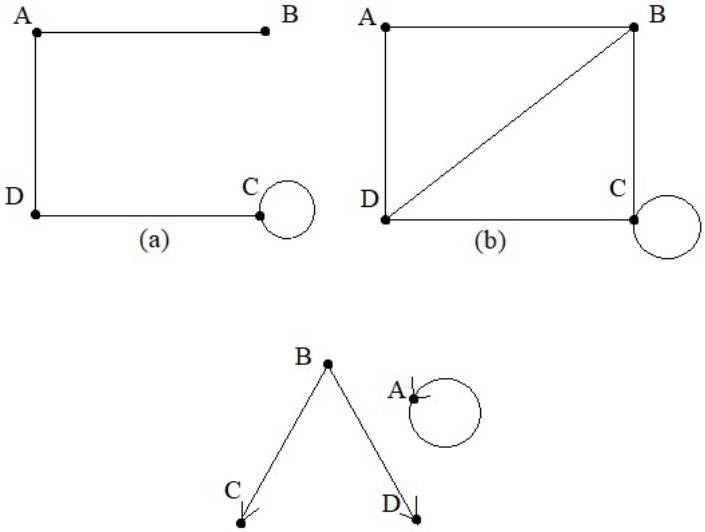
$$\begin{aligned} a_{ij} &= 0; \text{ if } v_i \text{ is not an end vertex of } e_j, \\ &= 1; \text{ if } v_i \text{ is an end vertex of } e_j \text{ but } e_j \text{ is not a loop}, \\ &= 2; \text{ if } e_j \text{ is a loop at } v_i. \end{aligned}$$

Exercise 3.2.6.

1. Find the adjacency matrix of the following graphs with respect to the listing A, B, C, D of the vertices:
2. Draw the graph of G represented by the given adjacency matrix

$$(a)\ A_G = \begin{bmatrix} 0 & 2 & 2 & 0 \\ 2 & 0 & 0 & 1 \\ 2 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad (b)\ A_G = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} \quad (c)\ A_G = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

3. Find the adjacency matrix of the digraph with respect to the listing A, B, C, D:



4. Draw the digraph represented by the given adjacency matrices:

$$(a)\ \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \quad (b)\ \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

5. Find the adjacency matrices of the graphs K_3 and $K_{2,3}$.

3.3 Isomorphism

Definition 3.3.1. Let $G_1 = (V_1, E_1, g_1)$ and $G_2 = (V_2, E_2, g_2)$ be two graphs. G_1 is said to be isomorphic to G_2 if there exists a one-to-one correspondence $f : V_1 \to V_2$ and a one-to-one correspondence $h : E_1 \to E_2$ in such a way that for any edge $e_k \in E_1$, $g_1(e_k) = \{v_i, v_j\}$ in G_1 if and only if $g_2(h(e_k)) = \{f(v_i), f(v_j)\}$ in G_2.

In other words, if $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two graphs, then G_1 is said to be isomorphic to G_2 if there exist a one-to-one correspondence $f : V_1 \to V_2$ and a one-to-one correspondence $h : E_1 \to E_2$ such that for any edge e_k in E_1, vertices v_i, v_j are end vertices of e_k in G_1 if and only if $f(v_i)$, $f(v_j)$ are end vertices of $h(e_k)$ in G_2. When we say two graphs are same, we mean they are isomorphic to each other.

Example 3.3.2. Let G and H be graphs as in figure 3.3.1 Both these graphs have six vertices and six edges. Moreover, both the graphs are simple. The degree sequence of both the graphs is $2, 2, 2, 2, 2, 2$.

Let us define $f : V_1 \to V_2$ and $h : E_1 \to E_2$ by

$$\begin{aligned} f: &\quad A \mapsto U,\ B \mapsto V,\ C \mapsto W,\ D \mapsto X,\ E \mapsto Y,\ F \mapsto Z; \\ h: &\quad a \mapsto u,\ b \mapsto v,\ c \mapsto w,\ d \mapsto x,\ e \mapsto y,\ f \mapsto z. \end{aligned}$$

Then we can check that these maps f and h serve as the one-to-one correspondence maps between the vertex sets and edge sets of the two graphs that satisfies the isomorphism conditions. Thus G and H are isomorphic.

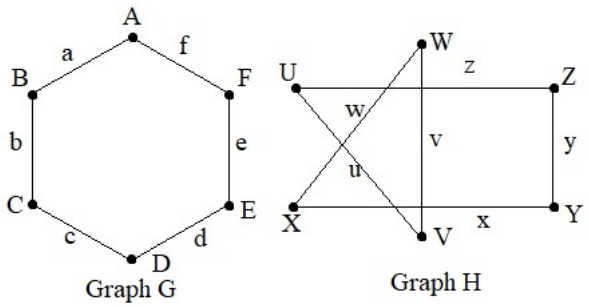


Figure 3.3.1

If two graphs G_1 and G_2 are isomorphic, then it is written as $G_1 \simeq G_2$.
The following theorem is evident

Theorem 3.3.3. Let G, G_1, G_2 and G_3 be graphs. Then the following assertions hold:

(i) $G \simeq G$;

(ii) If $G_1 \simeq G_2$, then $G_2 \simeq G_1$;

(iii) If $G_1 \simeq G_2$, and $G_2 \simeq G_3$, then $G_1 \simeq G_3$.

Proof. Left as an exercise. □

Definition 3.3.4. Two graph G_1 and G_2 are said to be different if G_1 is not isomorphic to G_2.

Let us write few properties of isomorphic graphs.

1. Two graphs G_1 and G_2 are isomorphic if and only if there exists a one-to-one correspondence f between the vertex sets of them such that if v_1, v_2 are adjacent vertices in G_1, then $f(v_1)$ and $f(v_2)$ are adjacent vertices in G_2.

2. Two graphs G_1 and G_2 are isomorphic. Then G_1 has a vertex of degree k if and only if G_2 has a vertex of degree k.

3. Two graphs G_1 and G_2 are isomorphic. Then G_1 has a cycle of length k if and only if G_2 has a cycle of length k.

3.4 Planar Graphs

Consider the graph in figure 3.4.1a. It can be redrawn as in the figure 3.4.1b.

We can also say that the above two graphs are isomorphic or *equal*. In the latter graph, notice that no two edges intersect except at the vertices. Such graphs are called planar graphs as we will formally define now.

Definition 3.4.1. A graph G is called a planar graph if it can be drawn in the plane such that no two edges intersect except at the vertices, which may the common end point of the edges. We can also say that a graph is planar if it is isomorphic to a graph having the property said above.

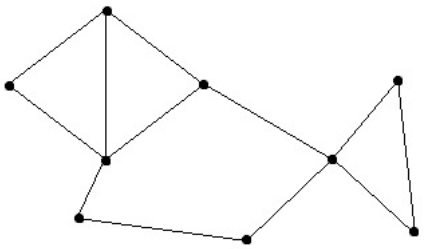
Definition 3.4.2. A graph drawn in the plane (on paper or a chalkboard) is called a plane graph if no two edges meet at any point except the common vertex, if they meet at all.

(a) (b)

Figure 3.4.1: Planar Graphs

From the preceding two definitions, it is clear that a graph is a planar graph if and only if it has a pictorial representation in a plane which is a plane graph. The pictorial representation of a planar graph G as a plane graph is called the planar representation of G.

Consider the planar representation of a planar graph given below



Let G denotes the graph in the above figure. Then G divides the plane into different *regions*, called the **faces** of G. Suppose x is a point in the plane that is not a vertex of G or a point on any edge of G. Then a face of G containing x is the set of all points on the plane that can be reached from x by a straight line or a curved line that does not cross any edge of G or pass through any vertex of G. Thus, it follows that a face is a region produced by a planar graph that is an area of the plane bounded by the edges and that is not further subdivided into sub-areas.

The set of edges that *bound* a region is called its **boundary**. Of course, there exists a region of infinite area in any plane graph G. This is the part of the plane that lies outside the planar representation of G. This region is called the **exterior face**. A face that is not exterior is called an **interior face**. We illustrate these concepts by the following example.

Example 3.4.3. Consider the graph below

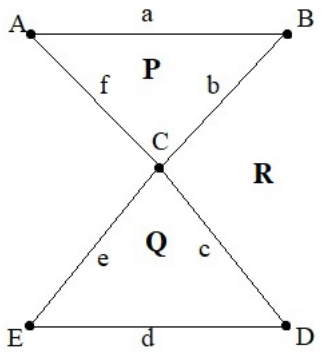
This plane graph divides the plane into three regions:

1: Bounded by the cycle (A, a, B, b, C, f, A). The boundary of P consists of the edges a, b, f.

2: Bounded by the cycle (D, d, E, e, C, c, D). The boundary of Q consists of the edges d, e, c.

3: The part of the plane outside this plane graph. The boundary of the region consists of the edges a, b, c, d, e and f.

It follows that this plane graph contains three faces, namely P, Q and R.



For this plane graph, the number of edges $n_e = 6$, the number of vertices $n_v = 5$, the number of faces $n_f = 3$, and we see that

$$n_v - n_e + n_f = 2.$$

Theorem 3.4.4. Let G be a connected planar graph with n_v vertices, n_e edges and n_f faces. Then $n_v - n_e + n_f = 2$.

Proof. We prove the theorem by induction on n_e.

Basic Step: Let $n_e = 0$. Then it has only one vertex and one region, which is the exterior region. Then, $n_v - n_e + n_f = 1 - 0 + 1 = 2$.

Inductive hypothesis: Let k be a positive integer and assume that $n_v - n_e + n_f = 2$ for any connected planar graph with $n_e = k - 1$.

Inductive Step: Let G be a connected planar graph with $n_e = k$ edges and $n_f = t$ vertices. Suppose G has no cycles. Then G has no interior region, which implies that the exterior region is the only region of the graph. Thus, $n_f = 1$. We now show that G contains a vertex of degree 1. Choose a vertex v in G. If $\deg(v) = 1$, we are done. Suppose $\deg(v) > 1$. Let v_1 be an adjacent vertex of v. Since G has no cycles, G is loop free and hence $v_1 \neq v$. If $\deg(v_1) = 1$, we are done. Suppose $\deg(v_1) > 1$. Let v_2 be an adjacent vertex of v_1. Since G has no cycles, G is loop free and hence v_2 is different from v and v_1. If $\deg(v_2) \neq 1$, we find an adjacent vertex v_3 of v_2 different from v, v_1 and v_2. Because G has finite number of vertices, it follows that G has a vertex u of degree 1. We now delete this vertex and thus from a new connected planar graph H with $k - 1$ edges and $t - 1$ vertices. By the inductive hypothesis, for this graph H, we have, $n_v - n_e + n_f = 2$. Hence, $(t - 1) - (k - 1) + n_f = 2$, which implies that $t - k + n_f = 2$, that is, $n_v - n_e + n_f = 2$ holds for G.

Suppose now that G has a cycle C. Let e be an edge in C. Now construct a new graph $G_1 = G \setminus \{e\}$. This is still a connected planar graph. For this planar graph G_1, we compute n_v, n_e and n_f. Let $n_f = m$. In the construction of G_1, we delete only the edge without deleting any vertex. Therefore, $n_v = t$, $n_e = k - 1$. Now, $C \setminus \{e\}$ will not form a boundary in G_1. Thus in G_1, $n_f = m - 1$. Hence G_1 is a connected planar graph with $n_v = t$ vertices, $n_e = k - 1$ edges, and $n_f = m - 1$ faces. By the inductive hypothesis, it follows that $t - (k - 1) + (m - 1) = 2$. This implies that $t - k + m = 2$. Hence, $n_v - n_e + n_f = 2$.

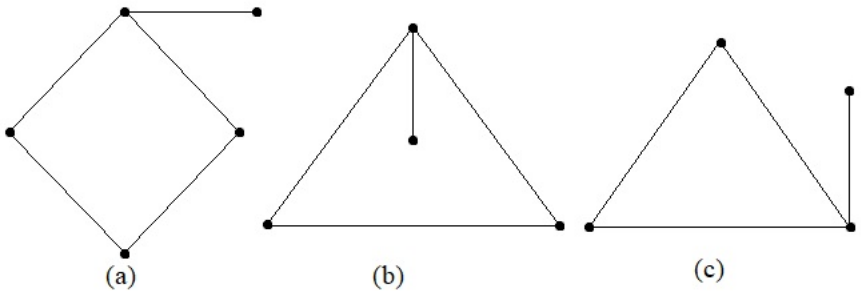
The result now follows by induction. □

Exercise 3.4.5. Verify the above theorem for the following graphs:

Corollary 3.4.6. The graph $K_{3,3}$ is not a planar graph.

Theorem 3.4.7. Let G be a connected simple planar graph with $n_v \geq 3$ vertices and n_e edges. Then

$$n_e \leq 3n_v - 6.$$



Proof. Since G is a planar graph, it has a planar representation. Consider a planar representation of G. Suppose $n_v = 3$. Because G is a simple connected graph with 3 vertices, it follows that $n_e \leq 3$. Then $n_e \leq 3 \cdot 3 - 6 = 3$, which implies that $n_e \leq 3n_v - 6$.

Suppose now $n_v \geq 3$. If G does not contain any cycles then we can show that $n_e = n_v - 1$. Now, $3n_v - 6 = (n_v - 1) + (n_v - 2) + (n_v - 3) > (n_v - 1) = n_e$.

Suppose G contains a cycle. Because G is simple, it may contain a cycle with 3 edges. Thus, the number of edges in the boundary of a face is ≥ 3. Now, there are n_f faces and every edge is a member of some boundary of the planar representation. Hence, the total number of appearances of the edges in boundaries of n_f faces is $\geq n_f \cdot 3$. In counting these appearances, an edge may be counted atmost two times. Thus, the total number of appearances of the n_e edges in boundaries is $\leq 2n_e$. Hence, $n_f \cdot 3 \leq 2n_e$. Now, by Euler's theorem,

$$\begin{aligned} & n_v - n_e + n_f = 2 \\ \Rightarrow \quad & 3n_v - 3n_e + 3n_f = 6 \\ \Rightarrow \quad & 3n_e = 3n_v + 3n_f - 6 \\ \Rightarrow \quad & 3n_e \leq 3n_v + 2n_e - 6 \\ \Rightarrow \quad & n_e \leq 3n_v - 6. \end{aligned}$$

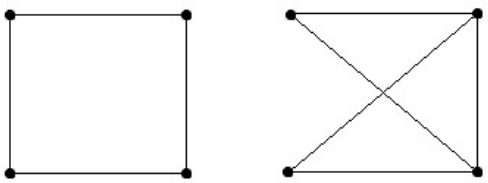
□

Corollary 3.4.8. The graph K_5 is not a planar graph.

Proof. Left as an exercise. □

Few Probable Questions

1. Define isomorphism of graphs. Determine whether the following graphs are isomorphic:



2. Define planar graphs. Show that for a connected simple planar graph G with $n_v \geq 3$, $n_e \leq 3n_v - 6$.
3. Show that for a connected planar graph, $n_v - n_e + n_f = 2$.

Unit 4

Course Structure

- Kuratowski's theorem (statement only) and its use
- Chromatic index, chromatic numbers and stability numbers

4.1 Introduction

In the previous unit, we saw that K_5 and $K_{3,3}$ are not planar. In this unit, we will see that this property of the two above graphs are used in general to characterise planarity of graphs as shown by Kuratowski. Further, we have dealt with the graph coloring. It is nothing but a simple way of labelling graph components such as vertices, edges, and regions under some constraints. Vertex coloring and edge coloring are two common graph coloring problems. The graph coloring problem has huge number of applications, like making schedule or time tables, sudoku, map coloring, etc.

Objectives

After reading this unit, you will be able to

- get to know the Kuratowski's theorem and its consequences;
- define vertex and edge coloring and related terms

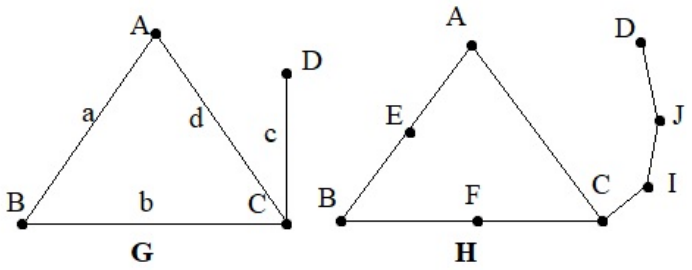
4.2 Kuratowski's Theorem

Let $G = (V, E)$ be a graph. Suppose that e is an edge with v_1, v_2 as end vertices. Construct the subgraph $G_1 = G \setminus \{e\}$. To construct G_1, we have deleted edge e without deleting any vertices from G. We now construct a new graph, $G_2 = (V_2, E_2)$, by taking $V_2 = V \cup \{w\}$, $E_2 = (E \setminus \{e\}) \cup \{f_1, f_2\}$ such that $w \notin V$, $f_1, f_2 \notin E$, v_1, w are end vertices of f_1 and v_2, w are end vertices of f_2. The process of obtaining G_2 from G is called a one-step subdivision of an edge of G.

Definition 4.2.1. A graph H is said to be a subdivision of a graph G if there exist graphs $H_0, H_1, H_2, \ldots, H_n$ such that $H_0 = G$, $H_n = H$, and H_i is obtained from H_{i-1} be a one-step subdivision of an edge of H_{i-1} for $i =, 1, 2, \ldots, n$.

If a graph H is a subdivision of a graph G, then we say that H is obtained from G by subdividing the edges of G.

Example 4.2.2. Consider graphs G and H below.

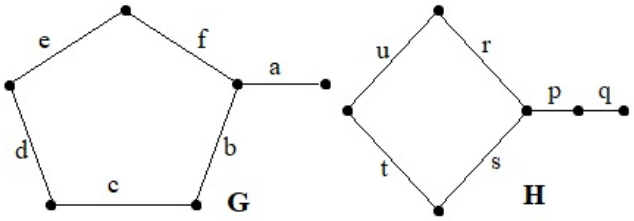


We see that H is obtained from H by a finite sequence of subdivisions of edges. H is obtained from G by dividing the edge a one time, b one time and c twice.

Definition 4.2.3. Two graphs G and H are said to be homeomorphic graphs if there is an isomorphism from a subdivision of G to a subdivision of H.

Consider the following example.

Example 4.2.4. Consider the graphs G and H below. We see that G contains a cycle of length 5, and H



contains a cycle of length 4. Hence these two graphs are not isomorphic. But we find a subdivision G' of G

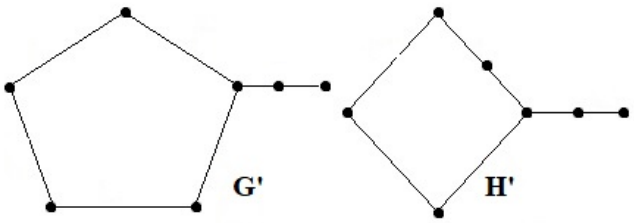


Figure 4.2.1

and H' of H such that G' and H' are isomorphic (see fig. 4.2.1). Hence G and H are homeomorphic.

In 1930, Kuratowski proved the following famous theorem, characterising simple planar graphs in terms of K_5 and $K_{3,3}$.

Theorem 4.2.5. Kuratowski. A simple graph is planar if and only if it does not contain a subgraph homeomorphic to K_5 or $K_{3,3}$.

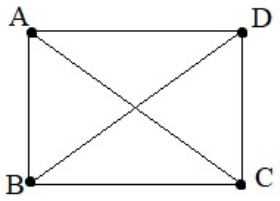
The proof of the above theorem is omitted.

4.3 Graph Coloring

Definition 4.3.1. Let $G = (V, E)$ be a simple graph and $C = \{c_1, c_2, \ldots, c_n\}$ be a set of n colors. A **vertex coloring** of G using the colors of C is a function $f : V \to C$. Let $f : V \to C$ be a vertex coloring of G.If for every adjacent vertices $u, v \in V$, $f(u) \neq f(v)$, then f is called a **proper vertex coloring**.

For each vertex v, its image $f(v)$ is called the **color** of v. It follows that a vertex coloring of a graph G is an assignment of the colors $c_1, c_2, \ldots, c_n$ to the vertices of graph G. Similarly, a proper vertex coloring of G is an assignment of the colors $c_1, c_2, \ldots, c_n$ to the vertices of G such that adjacent vertices have different colors. The following graph is an illustration.

Example 4.3.2. Consider the following graph:



This is a graph with 4 vertices A, B, C and D. Suppose $C = \{r, b, y, g\}$, where r denotes red, b denotes blue, y denotes yellow and g denotes green. Define $f : V \to C$ by

$$\begin{aligned} A &\mapsto r \\ B &\mapsto g \\ C &\mapsto y \\ D &\mapsto b. \end{aligned}$$

Then f is a proper vertex coloring with four colors.

Definition 4.3.3. The smallest number of colors needed to make a proper vertex coloring of a simple graph G is called the **chromatic number** of G and is denoted by $\chi(G)$.

Next we determine the chromatic number of bipartite graphs.

Theorem 4.3.4. Let G be a nontrivial simple graph. Then $\chi(G) = 2$ if and only if G is a bipartite graph.

Proof. Let $G = (V, E)$ be a bipartite graph. Then vertex set V can be partitioned into two non-empty subsets V_1 and V_2 such that each edge of G is incident with one vertex of V_1 and one vertex of V_2. Let $C = \{c_1, c_2\}$ be a set of two colors.

Define a function $f : V \to C$ such that

$$\begin{aligned} f(v) &= c_1; \text{ if } v \in V_1 \\ &= c_2; \text{ if } v \in V_2. \end{aligned}$$

Since $V_1 \cap V_2 = \emptyset$, it follows that f is well-defined. Now, no two vertices of V_1 are adjacent. Therefore, all the vertices can have the same color. Similarly, all the vertices of V_2 can have the same color. From the definition of f, it follows that two adjacent vertices of G have different colors. Thus, $\chi(G) \leq 2$. Also, since G has at least one edge, $\chi(G) < 1$. Hence combining, we get $\chi(G) = 2$.

Conversely suppose that $\chi(G) = 2$. This implies that the graph contains at least one edge. Also, there exists a function $f : V \to C = \{c_1, c_2\}$ such that no two adjacent vertices have the same image.

Let $V_1 = \{v \in V : \ f(v) = c_1\}$ and $V_2 = \{v \in V : \ f(v) = c_2\}$. It follows that $V_1 \cap V_2 = \emptyset$ and $V_1 \cup V_2 = V$. Let e be an edge with end vertices v_1 and v_2. Because v_1 and v_2 can't have the same color, $v_1 \in V_1$ and $v_2 \in V_2$. Thus, G is a bipartite graph. □

Definition 4.3.5. Let G be a graph with vertices $v_1, v_2, \ldots, v_{n-1}, v_n$. The maximum of the integers $\deg(v_i)$, for $i = 1, 2, \ldots, n$ is denoted by $\Delta(G)$.

Theorem 4.3.6. For any simple graph G, $\chi(G) \leq \Delta(G) + 1$.

Proof. We prove this theorem by induction on n, where n is the number of vertices of G.

Basic Step: Let $n = 1$. Then G is a graph with only one vertex and G has no edge. Hence $\chi(G) = 1$ and $\Delta(G) = 0$. This implies that $\chi(G) \leq \Delta(G) + 1$ for $n = 1$.

Inductive hypothesis: Suppose that $k > 1$ is an integer such that for any simple graph G, with $k-1$ vertices, $\chi(G) \leq \Delta(G) + 1$.

Inductive step: Let G be a simple graph with k vertices. Consider a vertex v of G and construct the graph $G_1 = G \setminus \{v\}$. The graph G_1 is obtained by deleting the vertex v and also all the edges incident on v. Clearly, $\Delta(G_1) \leq \Delta(G)$. This is a simple graph with $k - 1$ vertices. Thus, by the inductive hypothesis, $\chi(G_1) \leq \Delta(G_1) + 1$. Then, $\chi(G_1) \leq \Delta(G) + 1$. This implies that G_1 can be properly colored by atmost $\Delta(G_1) + 1$ colors. Now, v has atmost $\Delta(G)$ adjacent vertices. Because $\Delta(G) < \Delta(G) + 1$, it follows that not all the $\Delta(G) + 1$ colors are needed to color these $\Delta(G)$ adjacent vertices. Thus, from these $\Delta(G) + 1$ colors one unused color is definitely available to color vertex v. Hence, $\chi(G) \leq \Delta(G) + 1$. □

Definition 4.3.7. Let $G = (V, E)$ be a simple graph and $C = \{c_1, c_2, \ldots, c_n\}$ be a set of n colors. An **edge coloring** of G using the colors of C is a function $f : E \to C$. Let $f : E \to C$ be an edge coloring of G. If, for any two edges e_1 and e_2 meeting at a common vertex, $f(e_1) \neq f(e_2)$, then f is called a **proper edge coloring**.

For each edge e, its image $f(e)$ is called the color of e. It follows that a proper edge coloring of a graph G is an assignment of the colors $c_1, c_2, \ldots, c_n$ to the edges of graph G such that any two edges meeting at a common vertex have different colors. The following graph is an illustration.

Example 4.3.8. Consider the graph G in fig 4.3.1.

The graph G has six edges a, b, c, d, e, f. Suppose $C = \{R, B, Y, G\}$, where R denotes red, B denotes blue, Y denotes yellow, and G denotes green. Define $f : E \to C$ by

$$\begin{aligned} a &\mapsto R \\ c &\mapsto G \\ d &\mapsto B \\ f &\mapsto Y \\ b &\mapsto B \\ e &\mapsto R. \end{aligned}$$

Then f is a proper edge coloring of the graph G.

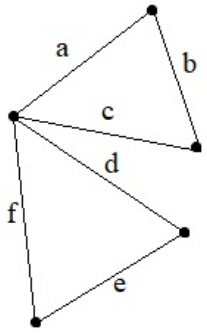


Figure 4.3.1

Definition 4.3.9. The smallest number of colors needed to make a proper coloring of the edges of a simple graph G is called the chromatic index of G, and is denoted by $\chi'(G)$.

For a simple graph, we have the following theorem.

Theorem 4.3.10. For any simple graph G, we have, $\chi'(G) = \Delta(G)$ or $\chi'(G) = \Delta(G) + 1$.

Let us see a few examples.

Example 4.3.11. In a connected simple planar graph, there exists a vertex v such that $\deg(v) \leq 5$.

We know that in a connected simple planar graph, $n_e \leq 3n_v - 6$. Suppose $\deg(v) \geq 6$ for all vertices v. Now, $\sum \deg(v) = 2n_e$. Hence $2n_e \geq 6n_v$. Again, $2n_e \leq 6n_v - 12$. This implies that $6n_v \leq 6n_v - 12$. Thus, we find that $0 \leq -12$, which is absurd. Hence the result.

Example 4.3.12. For the graph $K_{2,3}$, we find $\chi(K_{2,3})$. Let us first draw the graph (fig. 4.3.2). We see that

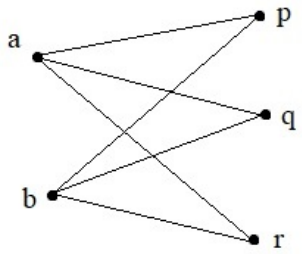


Figure 4.3.2

p, q, r are adjacent vertices of both a and b. Let $C = \{G, R\}$ be the set of two colors. Let us define $f : V \to C$ as follows:

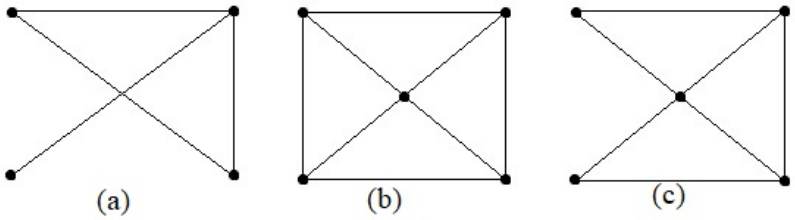
$$\begin{aligned} p &\mapsto R \\ a &\mapsto G \\ b &\mapsto G \\ q &\mapsto R \\ r &\mapsto R. \end{aligned}$$

This is a proper coloring of G. Hence $\chi(K_{2,3}) = 2$.

Example 4.3.13. For the graph K_n, we find $\chi(K_n)$. K_n is a complete graph with n vertices. For any vertex v of K_n, each of the remaining $n-1$ vertices is an adjacent vertex of v. Hence we need n distinct colors for proper coloring of K_n. Then, $\chi(K_n) \geq n$. But K_n has n vertices. So, $\chi(K_n) = n$.

Few Probable Questions

1. Define chromatic number of a graph G. Show that a simple nontrivial graph G has chromatic number 2 if and only if G is bipartite.
2. Show that for any simple graph G, $\chi(G) \leq \Delta(G) + 1$.
3. Find $\chi(G)$ for each of the following graphs:



Unit 5

Course Structure

- Lattices as partial ordered sets. Their properties. Lattices as algebraic system
- Some special Lattices e.g. complete complemented and distributed lattices

---

5.1 Introduction

A lattice is an abstract structure studied in the mathematical subdisciplines of order theory and abstract algebra. It consists of a partially ordered set in which every two elements have a unique supremum (also called a least upper bound or join) and a unique infimum (also called a greatest lower bound or meet). An example is given by the natural numbers, partially ordered by divisibility, for which the unique supremum is the least common multiple and the unique infimum is the greatest common divisor.

Lattices can also be characterized as algebraic structures satisfying certain axiomatic identities. Since the two definitions are equivalent, lattice theory draws on both order theory and universal algebra. Semilattices include lattices, which in turn include Heyting and Boolean algebras. These "lattice-like" structures all admit order-theoretic as well as algebraic descriptions.

Objectives

After reading this unit, you will be able to

- define partial ordered sets and see its examples
- upper and lower bounds of a poset
- define lattice and deduce the algebra of join and meet
- draw the Hasse diagram for posets

5.2 Partially Ordered Sets

Definition 5.2.1. A relation R on a set S is called **antisymmetric** if for all $a, b \in S$, $aRb \in R$ and $bRa \in R$ implies $a = b$.

On the set of all integers, the usual "less than or equal to" relation is an antisymmetric relation since $a \leq b$ and $b \leq a$ implies $a = b$.

Similarly if T is the set of all subsets of a set A, then the inclusion relation "$\subseteq$" is an antisymmetric relation since for any two subsets X and Y of A, we always have $X \subseteq Y$ and $Y \subseteq X$ implies $X = Y$.

Definition 5.2.2. A relation R on a set A is called a **partial order** on A if R is reflexive, antisymmetric and transitive. In other words, if R satisfies the following conditions:

1. aRa for all $a \in A$;

2. For all $a, b \in A$ if aRb and bRa, then $a = b$;

3. For all $a, b, c \in A$, if aRb and bRc, then aRc.

A set A together with a partial order relation R is called a **partially ordered set**, or **poset**, and we denote this poset by (A, R).

Let (A, R) be a poset. If there is no confusion about the partial order, we may refer to the poset simply by A.

Example 5.2.3. The set $\mathbb{Z}$, together with the usual "less than or equal to", $\leq$ relation is a poset. Note that the relation '$<$' is not a partial order relation on $\mathbb{Z}$ since the relation is not reflexive.

Example 5.2.4. Consider $\mathbb{N}$, the set of all natural numbers, and the divisibility relation R on $\mathbb{N}$. That is, for all $a, b \in \mathbb{N}$, aRb if $a|b$, that is, there exists a positive integer c such that $b = ac$. Check that this relation R is partial ordered. Thus, $\mathbb{N}$ with the divisibility relation is a poset.

Though the divisibility relation is a partial order relation on the set of all positive integers, it is not so on the set of all nonzero integers. For example, $5 = (-1)(-5)$ and also, $-5 = (-1)(5)$ and thus, $5|(-5)$ and $(-5)|5$ but $5 \neq -5$.

Let R be a partial order on a set A, that is, (A, R) is a poset. We usually denote R by $\leq_A$, or simple $\leq$. If A is a partially ordered set with a partial order $\leq$, then we denote it has $(A, \leq_A)$ or $(A, \leq)$.

Definition 5.2.5. Let $(S, \leq)$ be a poset and $a, b \in S$. If either $a \leq b$ or $b \leq a$ holds, then we say that a and b are **comparable**. The poset $(S, \leq)$ is called a linearly set, or totally ordered set, or a chain. if for all $a, b \in S$, either $a \leq b$ or $b \leq a$.

Example 5.2.6. Consider the poset $(\mathbb{Z}, \leq)$ with the less equal to relation. For any two integers a and b, either $a < b$, or $a = b$, or $b < a$. Thus, any two integers with respect to the partial order $\leq$ are comparable. Hence $(\mathbb{Z}, \leq)$ is a chain.

Example 5.2.7. Consider the poset $(\mathbb{N}, \leq)$ with respect to the divisibility relation. Notice here that 2 does not divide 5 and 5 does not divide 2. Thus, 2 and 5 are not comparable and hence $(\mathbb{N}, \leq)$ is not a chain.

Theorem 5.2.8. Any subset T of a poset S is itself a poset under the same relation (restricted to T). Any subset of a chain is a chain.

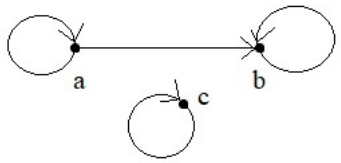
5.2.1 Digraphs of Posets

Because any partial order is also a relation, we can give a digraph representation of partial order.

Example 5.2.9. On the set $S = \{a, b, c\}$, consider the relation

$$R = \{(a, a), (b, b), (c, c), (a, b)\}.$$

The digraph of R is shown below.



From the directed graph it follows that the given relation is reflexive and transitive. This relation is also asymmetric because there is a directed edge from a to b, but there is no directed edge from b to a. Again, in the graph we notice that there are two distinct vertices a and c such that there are no directed edges from a to c and from c to a.

In a digraph of a partial order, one can see that if there is a directed edge from a vertex a to a different vertex b, then there is no directed edge from b to a.

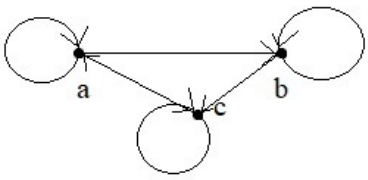
Theorem 5.2.10. A digraph of a partial order relation R cannot contain a closed directed path other than loops. (A path $a_1, a_2, \ldots, a_n$ in a digraph is closed if $a_1Ra_2, a_2Ra_3, \ldots, a_nRa_1$.)

By the above theorem, it follows that if a digraph of a relation contains a closed path other than loops, then the corresponding relation is not a partial order.

Example 5.2.11. On the set $S = \{a, b, c\}$, consider the relation

$$R = \{(a, a), (b, b), (c, c), (a, b), (b, c), (c, a)\}.$$

The digraph of the above relation is given by



In this digraph, a, b, c, a forms a closed path. Hence, the given relation is not a partial order relation.

Hasse Diagram

Posets can also be represented visually by Hasse diagram. First we define a few terms that we will need in the sequel.

Let $(S, \leq)$ be a poset and $x, y \in S$. We say that y **covers** x, if $x \leq y$, $x \neq y$, and there are no element $x \in S$ such that $x < z < y$.

We draw a diagram using the elements of S as follows: We represent the elements of S in the diagram by the elements themselves such that if $x \leq y$, then y is placed above x. We connect x with y by a line segment if and only if y covers x. The resulting diagram is called the **Hasse diagram** of $(S, \leq)$. We see the illustration below.

Example 5.2.12. Let $S = \{1, 2, 3\}$. Then $\mathcal{P}(S) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, S\}$. Now, $(\mathcal{P}(S), \leq$) is a poset, where $\leq$ denotes the set inclusion relation. The poset diagram of $(\mathcal{P}(S), \leq)$ is shown in fig. 5.2.1.

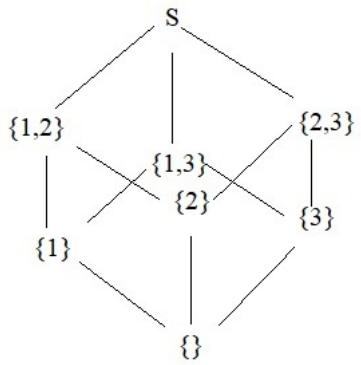


Figure 5.2.1

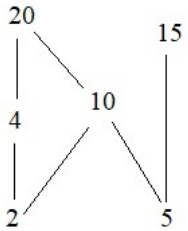
Minimal and Maximal Elements

Definition 5.2.13. Let $(S, \leq)$ be a poset. An element $a \in S$ is called

1. a **minimal element** if there is no element $b \in S$ such that $b < a$,
2. a **maximal element** if there is no element $b \in S$ such that $a < b$,
3. a **greatest element** if $b \leq a$ for all $b \in S$,
4. a **least element** if $a \leq b$ for all $b \in S$.

Let us illustrate this with the following example.

Example 5.2.14. Let $S = \{2, 4, 5, 10, 15, 20\}$. Let $(S, \leq)$ be a poset where $\leq$ denotes the divisibility relation. Then the Hasse diagram becomes



Now, it is clear from the Hasse diagram that there exists no greatest or least element of the poset since no element a satisfies $b \leq a$ for all $b \in S$ (for example, $2 \leq 15$ is not satisfied and also, $15 \leq 20$ is not satisfied),

and also, no element a exists that satisfy $a \leq b$ for every $b \in S$ (if we consider 2 or 5 as the least element then we see that $2 \leq 5$ does not hold and also $5 \leq 2$ does not hold). Now, 5 and 2 are definitely minimal elements since there exist no element $b \in S$ such that $b < 2$ or $b < 5$ (in other words, there is no line segment in the Hasse diagram extending below 2 or 5). Also, 20 and 15 are maximal elements of the poset (verify).

The following lemma ensures the existence of minimal element for every finite poset.

Lemma 5.2.15. Let $(S, \leq)$ be a poset such that S is a finite non-empty set. Then this poset has a minimal element.

Proof. Let a_1 be an element of S. If a_1 is a minimal element, then we are done. Suppose a_1 is not a minimal element. Then there exists $a_2 \in S$ such that $a_2 < a_1$. If a_2 is a minimal element, then we are done, otherwise there exists $a_3 \in S$ such that $a_3 < a_2$. If a_3 is not a minimal element, then we repeat this process. Now, $a_3 < a_2 < a_1$ shows that a_3, a_2, a_1 are distinct elements in S. Since S is finite, after a finite number of steps, we get an element $a_n \in S$, such that a_n is a minimal element. □

We must note that, a poset $(S, \leq)$, where S is a finite non-empty set, has minimal and maximal elements but may not have least or greatest elements. You can take the previous example as a confirmation of this fact.

Definition 5.2.16. Let S be a set and let $\leq_1$ and $\leq_2$ be two partial orders on S. The relation $\leq_2$ is said to be **compatible** with the relation $\leq_1$ if $a \leq_1 b$ implies $a \leq_2 b$.

It should be noted that given a finite non-empty set, say S, we can define a linear order in it as follows.

Since S is non-empty, S has at least one element. Choose an element S, and call it the first element, a_1. Let $S_1 = S \setminus \{a_1\}$. If S_1 is not empty, then from S_1, choose an element a_2. Let $S_2 = S \setminus \{a_1, a_2\}$. If S_2 is not empty, then from S_2, choose an element a_3. Let $S_3 = S\setminus\{a_1, a_2, a_3\}$. If S_3 is not empty, continue this process. Since S is a finite set, this process must stop after a finite number of steps. Hence, there exists a positive integer n such that $S_n = S \setminus \{a_1, \ldots, a_n\}$ is empty, where a_n is the element of $S_{n-1} = S \setminus \{a_1, \ldots, a_{n-1}\}$. We now define a partial order $\leq_1$ on S by $a_1 \leq_1 a_2 \leq_1 a_3 \cdots \leq_1 a_n$. This means that $a_i \leq_1 a_j$ if and only if either $i = j$ or $i < j$, where $i, j \in \{1, 2, \ldots, n\}$. It follows that this is a linear order.

Next suppose that not only S is a finite non-empty set, but S also has a partial order $\leq$. Can we define a linear order $\leq_1$ on S that is compatible with the partial order $\leq$? This following theorem is all about answering this question.

Theorem 5.2.17. Let $(S, \leq)$ be a finite poset. There exists a linear order $\leq_1$ on S which is compatible with the relation $\leq$.

We omit the proof of this theorem and go on to define lattices.

5.3 Lattice

Definition 5.3.1. Let $(S, \leq)$ be a poset and let $\{a, b\}$ be a subset of S. An element $c \in S$ is called an **upper bound** of $\{a, b\}$ if $a \leq c$ and $b \leq c$. Also, if T is any subset of S, then $c \in S$ is called an upper bound of T if $t \leq c$ for all $t \in T$.

An element $d \in S$ is called **least upper bound (lub)** of $\{a, b\}$ if,

1. d is an upper bound of $\{a, b\}$; and
2. if $c \in S$ is an upper bound of $\{a, b\}$, then $d \leq c$.

We can also define the lub of any general subset T of S and denote it by $\sup T$.

Example 5.3.2. Consider the set $\mathbb{N}$ together with the divisibility relation. Consider the subset $\{12, 8\}$. We see that $24, 48, 72$ are all common divisors of 12 and 8. Hence $12 \leq 24$ and $8 \leq 24$; $12 \leq 48$ and $8 \leq 48$; $12 \leq 48$ and $8 \leq 48$; $12 \leq 72$ and $8 \leq 72$. Thus, $24, 48, 72$ are upper bounds of $\{12, 8\}$ and hence 24 is the least upper bound of $\{12, 8\}$. Notice that $24 \notin \{12, 8\}$.

Theorem 5.3.3. In a poset $(S, \leq)$, if a subset $\{a, b\}$ of S has a lub, then it is unique.

Proof. Let $a, b \in S$ and a lub of $\{a, b\}$ exists. Suppose $c, d \in S$ are two lubs of $\{a, b\}$. Then c and d are upper bounds of $\{a, b\}$. Since c is a lub of $\{a, b\}$ and d is an upper bound, so $c \leq d$. Similarly, $d \leq c$. Then we have $c \leq d$ and $d \leq c$. By antisymmetry, we can say that $c = d$. Hence the result. □

The lub of $\{a, b\}$ in $(S, \leq)$, if it exists, is denoted by $a \vee b$, or the "join" of a and b.

Definition 5.3.4. Let $(S, \leq)$ be a poset and let $\{a, b\}$ be a subset of S. An element $c \in S$ is called a lower bound of $\{a, b\}$ if $c \leq a$ and $c \leq b$. Also, if T is any subset of S, then $c \in S$ is called an lower bound of T if $c \leq t$ for all $t \in T$.

An element $d \in S$ is called **greatest lower bound (glb)** of $\{a, b\}$ if,

1. d is a lower bound of $\{a, b\}$; and
2. if $c \in S$ is a lower bound of $\{a, b\}$, then $c \leq d$.

We can also define the glb of any general subset T of S and denote it by $\inf T$.

Then similar to the previous theorem, we can prove the following

Theorem 5.3.5. In a poset $(S, \leq)$, if a subset $\{a, b\}$ of S has a glb, then it is unique.

Proof. Left as an exercise. □

The glb of $\{a, b\}$ in $(S, \leq)$, if it exists, is denoted by $a \wedge b$, or the "meet" of a and b.

Definition 5.3.6. A poset $(L, \leq)$ is called a **lattice** if both $a \vee b$ and $a \wedge b$ exist for all $a, b \in L$. A lattice L is called **complete** if each of its subsets has a lub and glb in L.

Example 5.3.7. Any chain is a lattice in which $a \wedge b$ is simply the smaller of a and b and $a \vee b$ is simply the bigger of the two. Not every lattice is complete; the rational numbers are not complete with respect to the "usual less than or equal to relation", and the real numbers (in their natural order) are also not complete unless $-\infty$ and ∞ are adjoined to it.

Example 5.3.8. Let L be the set of all nonnegative real numbers. Then $(L, \leq)$ is a poset, where $\leq$ denotes the usual "less than or equal to" relation. Let $a, b \in L$. Now, $\max\{a, b\} \in L$ and $\min\{a, b\} \in L$. It is easy to see that $\max\{a, b\}$ is the lub of $\{a, b\}$ and $\min\{a, b\}$ is the glb of $\{a, b\}$. For example, $\max\{2, 5\} = 5 = 2 \vee 5$ and $\min\{2, 5\} = 2 = 2 \wedge 5$. Hence $(L, \leq)$ is a lattice. But it is not complete as we have discussed in the previous example.

Example 5.3.9. Let S be a set. Then $(\mathcal{P}(S), \leq)$ is a poset, where $\leq$ is the set inclusion relation. For $A, B \in \mathcal{P}(S)$, we can show that $A \vee B = A \cup B$ and $A \wedge B = A \cap B$. Hence $(\mathcal{P}(S), \leq)$ is a lattice. This lattice is however, complete and the glb of any family A of subsets of S is simply $\bigcap_A A_\alpha$ and the lub is $\bigcup_A A_\alpha$, both of which belong to $\mathcal{P}(S)$.

Theorem 5.3.10. Let $(L, \leq)$ be a lattice and let $a, b \in L$. Then

L1. $a \vee b = b \vee a$, $a \wedge b = b \wedge a$ (commutative laws),

L2. $a \vee (b \vee c) = (a \vee b) \vee c$, $a \wedge (b \wedge c) = (a \wedge b) \wedge c$ (associative laws),

L3. $a \vee a = a$, $a \wedge a = a$ (idempotent laws),

L4. $a \vee (a \wedge b) = a$, $a \wedge (a \vee b) = a$ (absorption laws).

Proof. Left as an exercise. □

Theorem 5.3.11. Let $(S, \leq)$ be a poset and $a, b \in S$. Then the following conditions are equivalent:

1. $a \leq b$;
2. $a \vee b = b$;
3. $a \wedge b = a$.

This is known as the **consistency** of the poset.

Proof. Left as exercise. □

Theorem 5.3.12. In any lattice $(L, \leq)$, the operations of join and meet are **isotonic**, that is, if $b \leq c$, then

$$a \wedge b \leq a \wedge c \quad \text{and} \quad a \vee b \leq a \vee c.$$

Proof. Let $b \leq c$. Then

$$a \wedge b = (a \wedge a) \wedge (b \wedge c) = (a \wedge b) \wedge (a \wedge c),$$

whence $a \wedge b \leq a \wedge c$ by consistency. Similarly, the other inequality can be shown. □

Theorem 5.3.13. In any lattice $(L, \leq)$, we have the distributive inequalities

$$\begin{array}{lrcl} \text{D} & (a \wedge b) \vee (a \wedge c) & \leq & a \wedge (b \vee c), \\ \text{D'} & a \vee (b \wedge c) & \leq & (a \vee b) \wedge (a \vee c), \end{array}$$

for all $a, b, c \in L$.

Proof. Clearly, $a \wedge b \leq a$ and $a \wedge b \leq b \leq b \vee c$. Hence, $a \wedge b \leq a \wedge (b \vee c)$. Also, $a \wedge c \leq a$, $a \wedge c \leq c \leq b \vee c$. Hence, $a \wedge c \leq a \wedge (b \vee c)$. That is, $a \wedge (b \vee c)$ is an upper bound of $a \wedge b$ and $a \wedge c$, from which, D follows. Similarly, we can prove D'. □

Definition 5.3.14. A lattice $(L, \leq)$ is called **distributive** if it satisfies

$$\text{D1.} \quad a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c),$$

for all $a, b, c \in L$.

The two previous examples of lattices that we discussed earlier, were both distributive lattices. However, it is worth mentioning that all lattices are not distributive as we see in the following example.

Example 5.3.15. Consider the lattice in the figure 5.3.1.
Since $a \wedge (b \vee c) = a \wedge 1 = a \neq 0 = 0 \vee 0 = (a \wedge b) \vee (a \wedge c)$, so this lattice is not distributive.

The next theorem gives a necessary and sufficient condition for a lattice to be distributive.

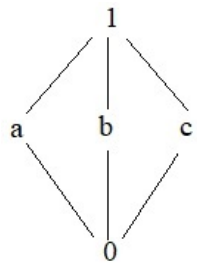


Figure 5.3.1

Theorem 5.3.16. A lattice $(L, \leq)$ is distributive if and only if

$$\text{D2.}\quad a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c),$$

for all $a, b, c \in L$.

Proof. Suppose $(L, \leq)$ is distributive. Let $a, b, c \in L$. Then

$$\begin{aligned}
(a \vee b) \wedge (a \vee c) &= ((a \vee b) \wedge a) \vee ((a \vee b) \wedge c) && \text{by D1}\\
&= (a \wedge (a \vee b)) \vee ((a \vee b) \wedge c) && \text{by L1}\\
&= a \vee ((a \vee b) \wedge c) && \text{by L4}\\
&= a \vee (c \wedge (a \vee b)) && \text{by L1}\\
&= a \vee ((c \wedge a) \vee (c \wedge b)) && \text{by D1}\\
&= (a \vee (c \wedge a)) \vee (c \wedge b) && \text{by L2}\\
&= (a \vee (c \wedge a)) \vee (b \wedge c) && \text{by L1}\\
&= a \vee (b \wedge c) && \text{by L4.}
\end{aligned}$$

Hence, $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$. Similarly, we can show that D2⇒D1. □

Theorem 5.3.17. In a distributive lattice $(L, \leq)$,

$$a \wedge b = a \wedge c \quad \text{and} \quad a \vee b = a \vee c \Rightarrow b = c,$$

for all $a, b, c \in L$.

Proof. Let $(L, \leq)$ be a distributive lattice. Now,

$$\begin{aligned}
b &= b \wedge (a \vee b)\\
&= b \wedge (a \vee c)\\
&= (b \wedge a) \vee (b \wedge c)\\
&= (a \wedge c) \vee (b \wedge c)\\
&= (c \wedge a) \vee (c \wedge b)\\
&= c \wedge (a \vee b)\\
&= c \wedge (a \vee c)\\
&= c.
\end{aligned}$$

□

Note that a poset $(L, \leq)$ may not contain a greatest element, but from the antisymmetric property of $\leq$, it can be shown that if there exists a greatest element in a poset, then it is unique, for if, a and b are two such elements, then $a \leq b$ and by the same argument, $b \leq a$, which implies that $a = b$. Similarly, a poset may contain at most one least element. We denote the greatest element of L by I and the least element by O. The elements O and I, when they exist, are called the **universal bounds** of L, since then $O \leq x$ and $x \leq I$ for all $x \in L$.

Theorem 5.3.18. If $(L, \leq)$ is a poset having O and I, then

$$\begin{aligned} O \wedge x &= O \quad \text{and} \quad O \vee x = x, \\ x \wedge I &= x \quad \text{and} \quad x \vee I = I, \end{aligned}$$

for all $x \in L$.

Proof. Left as exercise. □

Theorem 5.3.19. Let $(L, \leq)$ be a lattice. Then for all $a, b, c \in L$,

$$a \leq c \Rightarrow a \vee (b \wedge c) \leq (a \vee b) \wedge c.$$

This is called modular inequality.

Proof. We have, $a \leq a \vee b$ and $a \leq c$. Hence, $a \leq (a \vee b) \wedge c$. Also, $b \wedge c \leq b \leq a \vee b$ and $b \wedge c \leq c$. Thus, $b \wedge c \leq (a \vee b) \wedge c$. Thus, combining, we get the desired result. □

Definition 5.3.20. Let $(L, \leq)$ be a lattice with I and O. If $a \in L$, then an element $b \in L$ is said to be a **complement** of a if $a \vee b = I$ and $a \wedge b = O$.

Example 5.3.21. Let D_{30} denote the set of all positive divisors of 30. Then

$$D_{30} = \{1, 2, 3, 5, 6, 10, 15, 30\}.$$

Now, $(D_{30}, \leq)$ is a poset, where $a \leq b$ if and only if a divides b. Since 1 divides all the elements of D_{30}, it follows that $1 \leq m$ for all $m \in D_{30}$. Thus, 1 is the least element of this poset. Again, every member of D_{30} divides 30. Thus, $m \leq 30$. Hence, 30 is the greatest element of this poset. The Hasse diagram of this poset is given by fig. 5.3.2.

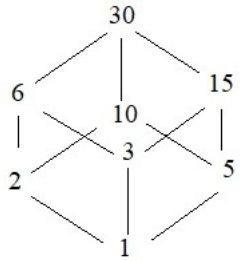


Figure 5.3.2

Let $a, b \in D_{30}$. Let $d = \gcd(a, b)$ and $m = \text{lcm}(a, b)$. Now, $d|a$ and $d|b$. Hence, $d \leq a$ and $d \leq b$. This shows that d is a lower bound of $\{a, b\}$. Let $c \in D_{30}$ and $c \leq a$ and $c \leq b$. Then $c|a$ and $c|b$ and since d is the gcd of a and b, so $c|d$, and hence $c \leq d$. Thus, $d = \gcd(a, b) = \text{glb}\{a, b\}$. Since all the positive divisors

of a, b are also divisors of 30, $d \in D_{30}$, so $d = a \wedge b$. Similarly we can show that $m \in D_{30}$ and $m = a \vee b$. Hence D_{30} is a complete lattice with least element 1 and greatest element 30.

Now, for any $a \in D_{30}$, $\frac{30}{a} \in D_{30}$. Using properties of gcd and lcm, we can show that for any $a \in D_{30}$,

$$a \wedge \frac{30}{a} = 1 \quad \text{and} \quad a \vee \frac{30}{a} = 30.$$

Hence, every element a has a complement $\frac{30}{a}$ in D_{30}.

Note that for any positive integer n, we can construct the lattice $(D_n, \leq)$, where $\leq$ denotes the usual divisibility relation in a similar way as shown in the preceding example.

Theorem 5.3.22. In a distributive lattice $(L, \leq)$ with I and O, every element has at most one complement.

Proof. Let $a \in L$. Suppose b, c are two complements of a in L. Then $a \vee b = I$ and $a \wedge b = O$; $a \vee c = I$ and $a \wedge c = O$. Hence $a \vee b = a \vee c$ and $a \wedge b = a \wedge c$. Then by theorem 5.3.17, it follows that $b = c$. Hence the result. □

A special type of distributive lattice is the **Boolean Algebra**. We will read about it in the upcoming units.

Few Probable Questions

1. Define poset. Show that every non-empty finite set has a minimal element.
2. Define lattice. Deduce the modular inequality of a lattice.
3. Deduce the distributive inequality of a lattice.
4. Deduce the necessary and sufficient condition for a lattice to be distributive.
5. Draw the Hasse diagram of D_{36} with respect to the usual divisibility relation and show that it is a lattice. Also, find the complement of each of the elements of D_{36}, if it exists.

Unit 6

Course Structure

- Sublattices. Direct products and Homomorphism.

6.1 Introduction

This unit is a continuation of the previous one. Here we will be focusing on two procedures to get a new lattice from old ones. One is the notion of sublattice, where a subset of a lattice becomes a lattice in itself. The other is by defining a lattice using Cartesian product of two or more lattices. Further, we will learn about lattice homomorphism. As the name suggests, it is a mapping that preserves the lattice structure between two lattices. Let us explore to know more.

Objectives

After reading this unit, you will be able to

- define sublattice and direct product of lattice
- define and understand lattice homomorphism

6.2 Sublattice

Definition 6.2.1. A **sublattice** of a lattice L is a subset X of L such that $a, b \in X$ imply $a \wedge b \in X$ and $a \vee b \in X$.

A sublattice is a lattice in its own right with the same join and meet operations. The empty set is a sublattice; so is any one-element subset. More generally, given $a \leq b$ in a lattice L, the interval $[a, b]$ of all elements $x \in L$ such that $a \leq x$ and $x \leq b$ is a sublattice.

A subset of a lattice L can be itself under the same (relative) order without being a lattice. Let us check the following example.

Example 6.2.2. Let Σ consist of the subgroups of a group G and let $\leq$ be the usual set inclusion relation. Then Σ is a complete lattice with $H \wedge K = H \cap K$ and $H \vee K$ the least subgroup in Σ containing H and K (which is not their set-theoretic union). Here, the set-union of two non-comparable subgroups is never a subgroup (since we know that the union of two subgroups H and K is a subgroup if and only if either $H \leq K$ or $K \leq H$). Hence this lattice is not a sublattice of the lattice of all subsets of G.

Definition 6.2.3. A property of subsets of a set I is a **closure property** when

1. I has the property, and
2. any intersection of subsets having the given property itself has this property.

Theorem 6.2.4. Let L be any complete lattice and let S be any subset of L such that

1. $I \in S$, and
2. $T \subset S$ implies $\inf T \in S$.

Then S is a complete lattice.

Proof. For any non-empty subset T of S, evidently $\inf T \in L$ is a member of S by 2, and it is the glb of T in S. Also, let U be the set of all upper bounds of T in S. It is non-empty since $I \in S$. Then, $\inf U \in S$ is also an upper bound of T. Moreover, it is the least upper bound since $\inf U \leq u$ for all $u \in U$. This proves that S is a complete lattice. □

Corollary 6.2.5. Those subsets of any set which have a given closure property form a complete lattice, in which the lattice meet of any family of subsets S_α is their intersection, and their lattice join is the intersection of all subsets T_β which contain every S_α.

6.3 Direct Product of Lattices

Let us consider two lattices $(L_1, \vee, \wedge)$ and $(L_2, \vee, \wedge)$. Also let $L = L_1 \times L_2 = \{(a, b) \mid a \in L_1,\ b \in L_2\}$. How can we define join and meet operations on this new set L? Let us take two elements (a_1, b_1) and (a_2, b_2) from L. Let us try and define the join and meet as follows:

$$\begin{aligned}(a_1, b_1) \vee (a_2, b_2) &= (a_1 \vee a_2, b_1 \vee b_2)\\ (a_1, b_1) \wedge (a_2, b_2) &= (a_1 \wedge a_2, b_1 \wedge b_2).\end{aligned}$$

Then $(L, \vee, \wedge)$ is the direct product of $(L_1, \vee, \wedge)$ and $(L_2, \vee, \wedge)$.

Let us consider the following example.

Example 6.3.1. Consider a lattice $(L, \leq)$ as shown in the following figure, where $L = \{1, 2\}$. Determine $(L^2, \leq)$, where $L^2 = L \times L$.

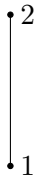


Figure 6.3.1: Hasse diagram of L

First let us start by finding L^2. We see that

$$L^2 = L \times L = \{(1,1),(1,2),(2,1),(2,2)\}.$$

Now take $(1,1),(1,2) \in L^2$. Then

$$(1,1) \vee (1,2) = (1 \vee 1, 1 \vee 2) = (1,2).$$

Also,

$$(1,1) \wedge (1,2) = (1 \wedge 1, 1 \wedge 2) = (1,1).$$

Similarly we can check for every pair of elements of L^2. The Hasse diagram of L^2 is given as follows.

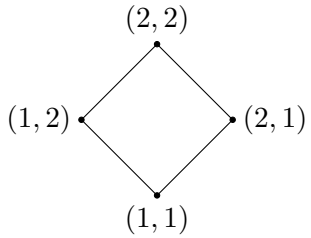


Figure 6.3.2

In the above example we have taken for granted that L^2 forms a lattice. However, we need to show that it is indeed true.

Theorem 6.3.2. The direct product $L_1 \times L_2$ of any two lattices L_1 and L_2 is a lattice.

Proof. For any two elements (a_1, b_1) and (a_2, b_2) in $L = L_1 \times L_2$, the element $(a_1 \vee a_2, b_1 \vee b_2)$ contains both (a_1, b_1) and (a_2, b_2), hence is an upper bound for the pair. Moreover every other upper bound (u, v) of the two satisfies $a_1 \leq u$ and $a_2 \leq u$ and hence by the definition of lub, $a_1 \vee a_2 \leq u$. Similarly, $b_1 \vee b_2 \leq v$, and so, $(a_1 \vee a_2, b_1 \vee b_2) \leq (u, v)$. This shows that

$$(a_1 \vee a_2, b_1 \vee b_2) = (a_1, b_1) \vee (a_2, b_2),$$

if the latter exists. By a similar argument for lower bound, we can show that

$$(a_1 \wedge a_2, b_1 \wedge b_2) = (a_1, b_1) \wedge (a_2, b_2),$$

if the latter exists. This shows that L is a lattice. □

Theorem 6.3.3. The direct product of two distributive lattices is a distributive lattice.

Proof. Let L and M be two distributive lattices and (a_1, b_1), (a_2, b_2) and (a_3, b_3) be elements in their product $L \times M$. Then

$$\begin{aligned}
[(a_1, b_1) \vee (a_2, b_2)] \wedge (a_3, b_3) &= (a_1 \vee a_2, b_1 \vee b_2) \wedge (a_3, b_3) \\
&= ((a_1 \vee a_2) \wedge a_3, (b_1 \vee b_2) \wedge b_3) \\
&= ((a_1 \wedge a_3) \vee (a_2 \wedge a_3), (b_1 \wedge b_3) \vee (b_2 \wedge b_3)) \\
&= ((a_1 \wedge a_3), (b_1 \wedge b_3)) \vee ((a_2 \wedge a_3), (b_2 \wedge b_3)) \\
&= ((a_1, b_1) \wedge (a_3, b_3)) \vee ((a_2, b_2) \wedge (a_3, b_3)).
\end{aligned}$$

We can similarly verify the other distributive law. □

6.4 Lattice Homomorphism

Let $(L, \vee, \wedge)$ and $(M, \vee, \wedge)$ be two lattices. A function $f : L \to M$ is called a lattice homomorphism if

$$f(a \vee b) = f(a) \vee f(b)$$

and

$$f(a \wedge b) = f(a) \wedge f(b)$$

where $a, b \in L$.

This can be easily seen that a lattice homomorphism if it preserves the join and meet operations. We can also say that the mapping f is a lattice homomorphism if it is both a join-homomorphism and a meet-homomorphism. To understand better, let us check the following example.

Example 6.4.1. Consider the lattices $D_6 = \{1, 2, 3, 6\}$ and $D_{30} = \{1, 2, 3, 5, 6, 10, 15, 30\}$ with the divisibility relation. We shall check whether there exist a lattice homomorphism between the two lattices. The Hasse diagram of the two lattices are given below.

(a) D_6 **(b)** D_{30}

Figure 6.4.1: Hasse Diagrams of D_6 and D_{30}

For a function to be homomorphism, we need to check for the two conditions stated in the preceding definition. Let us define a function $f : D_6 \to D_{30}$ as follows.

$$f(1) = 1, \;\; f(2) = 6, \;\; f(3) = 15, \;\; f(6) = 30.$$

Now we see that

$$f(1 \vee 2) = f(2) = 6; \quad f(1) \vee f(2) = 1 \vee 6 = 6$$

and

$$f(1 \wedge 2) = f(1) = 1; \quad f(1) \wedge f(2) = 1 \wedge 6 = 1.$$

Hence, we see that the two conditions of homomorphism are satisfied for $1, 2$ in D_6. In this way, we can check that the two join and meet homomorphism conditions are satisfied for every pair of elements in D_6. Thus, f is a lattice homomorphism.

If further, f is bijective, then f is called a lattice isomorphism. In the above example, we see that f is not surjective. Hence it can't be a lattice isomorphism.

Example 6.4.2. Consider the lattice $\mathbb{N}$ with the usual order. Let $S = \{1, 2, 3\}$ with the usual order. Let $f : \mathbb{N} \to S$ be a homomorphism. If $f(m) = 0$ and $f(n) = 1$, then $m \leq n$, or else, we have

$$f(m \vee n) = f(m) = 0, \;\; f(m) \vee f(n) = 0 \vee 1 = 1.$$

Thus, $f(m \vee n) \neq f(m) \vee f(n)$. So f must have one of the following forms.

1. $f^{-1}(0) = \mathbb{N}$;
2. $f^{-1}(0) = \{1, 2, \ldots, k\}$ and $f^{-1}(1) = \mathbb{N} \setminus \{1, 2, \ldots, k\}$ for some $k \in \mathbb{N}$;
3. $f^{-1}(0) = \{1, 2, \ldots, k\}$, $f^{-1}(1) = \{k+1, k+2, \ldots, k+r\}$ and $f^{-1}(2) = \mathbb{N} \setminus \{1, 2, \ldots, k+r\}$ for some $k, r \in \mathbb{N}$.

Few Probable Questions

1. Define direct product of lattice. Show that the direct product is two lattice is a lattice.

Unit 7

Course Structure

- Boolean Algebra Basic Definitions, Duality, Basic theorems, Boolean algebra as lattices.

7.1 Introduction

In mathematics and mathematical logic, Boolean algebra is the branch of algebra in which the values of the variables are the truth values true and false, usually denoted 1 and 0 respectively. Instead of elementary algebra where the values of the variables are numbers, and the prime operations are addition and multiplication, the main operations of Boolean algebra are the meet (and) denoted as $\wedge$, the join (or) denoted as $\vee$, and the negation (not) denoted as $\neg$. It is thus a formalism for describing logical operations in the same way that elementary algebra describes numerical operations.

Boolean algebra was introduced by George Boole in his first book The Mathematical Analysis of Logic (1847), and set forth more fully in his An Investigation of the Laws of Thought (1854). According to Huntington, the term “Boolean algebra" was first suggested by Sheffer in 1913, although Charles Sanders Peirce in 1880 gave the title “A Boolean Algebra with One Constant" to the first chapter of his “The Simplest Mathematics". Boolean algebra has been fundamental in the development of digital electronics, and is provided for in all modern programming languages. It is also used in set theory and statistics.

Objectives

After reading this unit, you will be able to:

- define Boolean algebra and derive some useful properties of it
- establish a partial order relation on a Boolean algebra
- deduce that an Boolean algebra is a lattice with respect to the partial order defined

7.2 Boolean Algebra

Though we gave a rough idea about Boolean Algebra in the previous unit, we start afresh in this unit to define Boolean Algebra.

Definition 7.2.1. A class of elements B together with two binary operations $(+)$ and $(\cdot)$ (where $a \cdot b$ will be written as ab) is a Boolean algebra if and only if it satisfies the following postulates:

B1. The operations $(+)$ and $(\cdot)$ are commutative.

B2. There exist in B distinct identity elements 0 and 1 relative to the operations $(+)$ and $(\cdot)$ respectively.

B3. Each operation is distributive over the other.

B4. For every $a \in B$, there exists an element a' in B such that

$$a + a' = 1 \quad \text{and} \quad aa' = 0.$$

The symbols " $+$ " and " $\cdot$ " is just a convention. We could use any other symbols in place of these two.

Example 7.2.2. Let S be any set and $\mathcal{P}(S)$ be the set of all subsets of S. Then $\mathcal{P}(S)$ forms a Boolean algebra where the binary operations $(+)$ and $(\cdot)$ are the set-theoretic union and intersections respectively. The corresponding identity elements are S and $\emptyset$ respectively. For every element $T \in \mathcal{P}(S)$, the complement is given by $S \setminus T$.

Theorem 7.2.3. Every statement or algebraic identity deducible from the postulates of a Boolean algebra remains valid if the operations $(+)$ and $(\cdot)$, and the identity elements 0 and 1 are interchanged throughout.

This theorem is called the principle of duality.

Proof. The proof of this theorem follows at once from the symmetry of the postulates with respect to the two operations and the two identities. □

It should be noted that the steps in one proof are dual statements to those in the other, and the justification for each step is the same postulate or theorem in one case as in the other.

Theorem 7.2.4. For every element a in a Boolean algebra B,

$$a + a = a \quad \text{and} \quad aa = a.$$

Proof.

$$\begin{aligned} a &= a + 0 && \text{by B2} \\ &= a + aa' && \text{by B4} \\ &= (a + a)(a + a') && \text{by B3} \\ &= (a + a)(1) && \text{by B4} \\ &= a + a, && \text{by B2} \end{aligned}$$

and similarly,

$$\begin{aligned} a &= a(1) && \text{by B2} \\ &= a(a + a') && \text{by B4} \\ &= aa + aa' && \text{by B3} \\ &= aa + 0 && \text{by B4} \\ &= aa. && \text{by B2} \end{aligned}$$

□

Thus, we can say that $(+)$ and $(\cdot)$ operations are idempotent.

Theorem 7.2.5. For every element a in a Boolean algebra B,

$$a + 1 = 1 \quad \text{and} \quad a0 = 0.$$

Proof.

$$\begin{aligned} 1 &= a + a' && \text{by B4} \\ &= a + a'(1) && \text{by B2} \\ &= (a + a')(a + 1) && \text{by B3} \\ &= 1(a + 1) && \text{by B4} \\ &= a + 1. && \text{by B2} \end{aligned}$$

The other part is left as an exercise. □

Theorem 7.2.6. For each pair of elements a and b in a Boolean algebra B,

$$a + ab = a \quad \text{and} \quad a(a + b) = a.$$

Proof.

$$\begin{aligned} a &= 1a && \text{by B2} \\ &= (1 + b)a && \text{by Theorem 7.2.5} \\ &= 1a + ba && \text{by B3 and B1} \\ &= a + ba && \text{by B2} \\ &= a + ab. && \text{by B1} \end{aligned}$$

The other part is left as exercise. □

Theorem 7.2.7. In every Boolean algebra B, each of the binary operations $(+)$ and $(\cdot)$ is associative. That is, for every a, b, and c in B,

$$a + (b + c) = (a + b) + c \quad \text{and} \quad a(bc) = (ab)c.$$

Proof. First we will show that $a + a(bc) = a + (ab)c$, as follows:

$$\begin{aligned} a + a(bc) &= a && \text{by Theorem 7.2.6} \\ &= a(a + c) && \text{by Theorem 7.2.6} \\ &= (a + ab)(a + c) && \text{by Theorem 7.2.6} \\ &= a + (ab)c. && \text{by B3} \end{aligned}$$

Next we will show that $a' + a(bc) = a' + (ab)c$, as follows:

$$\begin{aligned} a' + a(bc) &= (a' + a)(a' + bc) && \text{by B3} \\ &= 1(a' + bc) && \text{by B4} \\ &= a' + bc && \text{by B2} \\ &= (a' + b)(a' + c) && \text{by B3} \\ &= [1(a' + b)](a' + c) && \text{by B2} \\ &= [(a' + a)(a' + b)](a' + c) && \text{by B4} \\ &= (a' + ab)(a' + c) && \text{by B3} \\ &= a' + (ab)c. && \text{by B3} \end{aligned}$$

Now if we multiply these two equations, we obtain

$$[a + a(bc)][a' + a(bc)] = [a + (ab)c][a' + (ab)c]. \tag{7.2.1}$$

The left side of the above equation may be reduced as follows

$$\begin{aligned} [a + a(bc)][a' + a(bc)] &= [a(bc) + a][a(bc) + a'] && \text{by B1} \\ &= a(bc) + aa' && \text{by B3} \\ &= a(bc) + 0 && \text{by B4} \\ &= a(bc). && \text{by B2} \end{aligned}$$

Similarly, the right side of equation (7.2.1) reduces as follows:

$$\begin{aligned} [a + (ab)c][a' + (ab)c] &= [(ab)c + a][(ab)c + a'] && \text{by B1} \\ &= (ab)c + aa' && \text{by B3} \\ &= (ab)c + 0 && \text{by B4} \\ &= (ab)c. && \text{by B2} \end{aligned}$$

Thus, equation (7.2.1) reduces to

$$a(bc) = (ab)c,$$

which is the required associative law we were to prove. By duality principle, the analogous part for $(+)$ follows. □

From now on, we shall write both $a(bc)$ and $(ab)c$ as abc, and similarly, we shall write both $(a+b)+c$ and $a+(b+c)$ as $a+b+c$.

Theorem 7.2.8. The element a' associated with the element a in a Boolean algebra is unique.

Proof. Suppose that $a + x = 1$, $ax = 0$, and also that $a + y = 1$, $ay = 0$. Then,

$$\begin{aligned} x &= 1.x && \text{by B2} \\ &= (a + y)x && \text{by assumption} \\ &= (ax + yx) && \text{by B3 and B1} \\ &= 0 + yx && \text{by assumption} \\ &= yx && \text{by B2} \\ &= xy && \text{by B1} \\ &= xy + 0 && \text{by B2} \\ &= xy + ay && \text{by assumption} \\ &= (x + a)y && \text{by B3 and B1} \\ &= 1y && \text{by assumption} \\ &= y. && \text{by B2} \end{aligned}$$

□

Thus any two elements associated with a as specified in B4 are equal. In other words, a' is uniquely determined by a. We will refer to a' as the complement of a.

Theorem 7.2.9. For every a in a Boolean algebra B, $(a')' = a$.

Proof. By B4, $a + a' = 1$ and $aa' = 0$. But this is exactly the necessary condition that $(a')'$ is equal to a. By the previous theorem, this is unique and hence the result. □

Theorem 7.2.10. S. In any Boolean algebra, $0' = 1$ and $1' = 0$.

Proof. By theorem 7.2.5, $1 + 0 = 1$, and $1 \cdot 0 = 0$. Since theorem 7.2.8 shows that for each a there is only one element a', these equations imply that $0' = 1$, and $1' = 0$. □

Theorem 7.2.11. For every a and b in a Boolean algebra B,

$$(ab)' = a' + b' \quad \text{and} \quad (a + b)' = a'b'.$$

Proof. First,

$$\begin{aligned} (ab)(a' + b') &= aba' + abb' && \text{by B3} \\ &= 0b + a0 && \text{by B1, B2, B4} \\ &= 0 + 0 = 0. && \text{by theorem 7.2.5} \end{aligned}$$

Further,

$$\begin{aligned} ab + a' + b' &= a' + b' + ab && \text{by B1} \\ &= (a' + b' + a)(a' + b' + b) && \text{by B3} \\ &= (1 + b')(1 + a') && \text{by B4 and B1} \\ &= 1. && \text{by theorem 7.2.5 and B2} \end{aligned}$$

Now, by B4 and theorem 7.2.8, we can show that $(ab)' = a' + b'$. The part can be shown by duality principle. □

This is known as D'Morgan's law.

7.3 Boolean Algebra as Lattices

We now define an order relation on a Boolean algebra B by the following.

Definition 7.3.1. The "order" relation $a \leq b$ is defined by the statement:
For every a and b in a Boolean algebra B, $a \leq b$ if and only if $ab' = 0$.

Let us see certain properties of the relation as follows:

Theorem 7.3.2. The following four properties of $\leq$ are valid in every Boolean algebra for arbitrary elements x, y, and z:

1. $x \leq x$ (reflexive);
2. if $x \leq y$ and $y \leq x$, then $x = y$ (antisymmetry);
3. if $x \leq y$ and $y \leq z$, then $x \leq z$ (transitive);
4. if $x \leq y$ and $x \leq z$, then $x \leq yz$;
5. if $x \leq y$, then $x \leq y + z$, for any z;
6. $x \leq y$ if and only if $y' \leq x'$.

Proof. 1. Left for reader.

2. $x \leq y$ and $y \leq z$ are equivalent to $xy' = 0$ and $yx' = 0$ respectively. Now,

$$\begin{aligned} x &= x(1) && \text{by B2} \\ &= x(y + y') && \text{by B4} \\ &= xy + xy' && \text{by B3} \\ &= xy && \text{by assumption} \\ &= yx && \text{by B1} \\ &= yx + yx' && \text{by B2 and assumption} \\ &= y(x + x') && \text{by B3} \\ &= y(1) = y. && \text{by B4} \end{aligned}$$

3. $x \leq y$ is equivalent to $xy' = 0$. Also, $y \leq z$ is equivalent to $yz' = 0$. Now,

$$\begin{aligned} xz' &= xz'(1) && \text{by B2} \\ &= xz'(y + y') && \text{by B4} \\ &= xyz' + xy'z' && \text{by B1 and associativity} \\ &= 0 + 0. && \text{by assumption} \end{aligned}$$

Thus, $x \leq z$.

4. $x \leq y$ and $x \leq z$ are equivalent to $xy' = 0$ and $xz' = 0$ respectively. Now,

$$\begin{aligned} x(yz)' &= x(y' + z') && \text{by theorem 7.2.11} \\ &= xy' + xz' && \text{by B3} \\ &= 0. && \text{by assumption} \end{aligned}$$

Hence $x \leq yz$.

5. $x \leq y$ is equivalent to $xy' = 0$. Let $z \in B$ be arbitrary. Then

$$\begin{aligned} x(y + z)' &= x(y'z') && \text{by theorem 7.2.11} \\ &= 0. && \text{by associativity and assumption} \end{aligned}$$

Thus, $x \leq y + z$ for any $z \in B$.

6. $x \leq y$ is equivalent to $xy' = 0$. Thus,

$$\begin{aligned} y'(x')' &= y'x && \text{by theorem 7.2.9} \\ &= xy' && \text{by B1} \\ &= 0. && \text{by assumption} \end{aligned}$$

Hence, $y' \leq x'$.

□

The first three points of the above theorem show that B forms a poset with respect to the relation $\leq$ defined above. We will show that Boolean algebra forms lattice with respect to the defined partial order.

Theorem 7.3.3. Let B be a Boolean algebra with respect to the partial order $\leq$ defined as $x \leq y$ if and only if $xy' = 0$. Then B is a lattice with respect to $\leq$.

Proof. We will be done if we show that $\{x, y\}$ has lub and glb in B. We show that $x + y$ is the lub and xy is the glb of the set. Since $x(x + y)' = x(x'y') = xx'y' = 0$ and similarly, $y(x + y)' = 0$ so $x \leq (x + y)$ and $y \leq (x + y)$. Thus, $x + y$ is an upper bound of $\{x, y\}$. Let z be any other upper bound of $\{x, y\}$. Then $x \leq z$ and $y \leq z$ which imply $xz' = 0$ and $yz' = 0$. Now,

$$(x + y)z' = xz' + yz' = 0$$

which shows that $x + y \leq z$. Thus, $x + y$ is the lub of $\{x, y\}$. We can similarly show that xy is the glb of $\{x, y\}$. Thus, $(B, \leq)$ forms a lattice. □

The join and meet are defined as $x \vee y = x + y$ and $x \wedge y = xy$, for any arbitrary $x, y \in B$.

Also, note from the previous theorems that a Boolean algebra is distributive, and each element of it has a complement. Thus, a Boolean algebra is a distributive complemented lattice. Let us see certain examples.

Example 7.3.4. Let B be a Boolean algebra. We simplify the expression $x + (yx)'$, where $x, y \in B$.

We have,

$$\begin{aligned} x + (yx)' &= x + (y' + x') && \text{by theorem 7.2.11} \\ &= (x + x') + y' && \text{by B1} \\ &= 1 + y' && \text{by B4} \\ &= (0y)' = 0' = 1. && \text{by theorems 7.2.11 and 7.2.10} \end{aligned}$$

Example 7.3.5. In a Boolean algebra B, we simplify $(xy)'(x' + y)(y' + y)$, for $x, y \in B$.

We have,

$$\begin{aligned} (xy)'(x' + y)(y' + y) &= (xy)'(x' + y) && \text{by B4 and B2} \\ &= (x' + y')(x' + y) && \text{by theorem 7.2.11} \\ &= x' + y'y && \text{by theorem 7.2.7} \\ &= x'. && \text{by B4} \end{aligned}$$

Example 7.3.6. In a Boolean algebra B, we simplify $(x + z)(xt + xt') + xz + z$, for $x, z, t \in B$.

We have,

$$\begin{aligned} (x + z)(xt + xt') + xz + z &= (x + z)x(t + t') + xz + z \\ &= (x + z)x + xz + z \\ &= x((x + z) + z) + z \\ &= x(x + z) \\ &= xx + xz + z \\ &= x + (x + 1)z \\ &= x + z. \end{aligned}$$

Example 7.3.7. In a Boolean algebra B, we simplify $x'(x + y) + (y + xx)(x + y')$, for $x, y \in B$.

We have,

$$\begin{aligned} x'(x+y)+(y+xx)(x+y') &= x'x+x'y+(y+x)x+(y+x)y' \\ &= x'y+(y+x)x+(y+x)y' \\ &= x'y+yx+xx+yy'+xy' \\ &= x'y+yx+x+xy' \\ &= x'y+x(y+1+y') \\ &= x'y+x \\ &= x+x'y \\ &= (x+x')(x+y) \\ &= x+y. \end{aligned}$$

Few Probable Questions

1. Establish the distributive property of Boolean algebra.
2. Define a partial order relation on a Boolean algebra B. Hence show that it is a lattice with respect to the defined partial order.
3. Deduce the De'Morgan's law for Boolean algebra.
4. Show that the complement of an element in a Boolean algebra is always unique.
5. Show that $0' = 1$ and $1' = 0$ in a Boolean algebra.
6. In a Boolean algebra B, show that for any $a, b \in B$, $a(a+b) = a$.
7. Deduce the idempotent property of both the binary operators $(+)$ and $(\cdot)$ in a Boolean algebra.
8. In a Boolean algebra B, simplify the following:
   (a) $y(x'z+xz')+x(yz+yz')$;
   (b) $xyz+x'+xy'z$;
   (c) $(xy'+x'y)'(x+y)$;
   (d) $(xy)'(x'+y)(y'+y)$.

Unit 8

Course Structure

- Boolean functions, Sum and Product of Boolean algebra, Minimal Boolean Expressions, Prime implicants Propositions and Truth tables.

8.1 Introduction

This unit starts with the dnf and cnf which are normals forms and continuation of the previous unit. Next we move on to logic gates. Logic is an extensive field of study with many special areas of inquiry. In general, logic is concerned with the study and analysis of methods of reasoning or argumentation. Symbolic logic is not precisely defined as distinct from logic in general, but might be described as a study of logic which employs an extensive use of symbols. In any discussion of logic, the treatment centers around the concept of a proposition (statement). The principal tool for treatment of propositions is the algebra of propositions, a Boolean algebra. In talking about propositions, we will also investigate certain logical forms which represent acceptable techniques for constructing precise proofs of theorems. Since statements are formed from words, it is apparent that some consideration must be given to words and their meanings. No logical argument can be based on words that are not precisely described. That part of logic which is concerned with the structure of statements is much more difficult than the areas mentioned previously, and in fact, has not been satisfactorily formalized.

Objectives

After reading this unit, you will be able to

- define disjunctive normal forms and deduce related results
- define conjunctive normal forms and deduce related results
- solve problems related to dnf and cnf
- define propositions and learn to form complex propositions by conjunction, disjunction and negation
- show that the set of all propositions form a Boolean algebra with respect to the conjunction, disjunction and negation so defined
- draw truth tables for complex propositions

8.2 Disjunctive Normal Form

We start assuming that the reader is familiar with the terms monomial, polynomial, terms, factor, variable constants. By a Boolean function we will mean any expression which represents the combination of a finite set of symbols, each representing a constant or a variable, by the operations of $(+)$, $(\cdot)$, or complement. Thus, $(a'+b)'c+ab'x+0$ is a Boolean function provided that each of the symbols a, b, c, x represents an element of a Boolean algebra. Further example such as the equation $x + x' = 1$ represents the statement that a function $x + x'$ of the variable x equals the constant 1.

Among the functions of n variables $x_1, x_2, \ldots, x_n$ which can be written, a particular class of functions is of special interest, namely, those written as a sum of terms in which each term is a product involving all n variables either with or without a prime. Examples of such functions are $x + x'$, xy', $xyz' + x'yzxy'z$ in one, two, and three variables, respectively. The following definition gives a name to such functions.

Definition 8.2.1. A Boolean function is said to be in disjunctive normal form in n variables $x_1x_2, \ldots, x_n$, for $n > 0$, if the function is a sum of terms of the type $f_1(x_1)f_2(x_2)\cdots f_n(x_n)$, where $f_i(x_i)$ is x_i, or x_i' for each $i = 1, 2, \ldots, n$, and no two terms are identical. In addition, 0 and 1 are said to be in disjunctive normal form in n variables for any $n \geq 0$.

Some important properties of the disjunctive normal form are given in the following theorems.

Theorem 8.2.2. Every function in a Boolean algebra which contains no constants is equal to a function in disjunctive normal form.

Proof. Let an arbitrary function (without constants) of the n variables $x_1, x_2, \ldots, x_n$ denoted by f. If f contains an expression of the form $(A+B)'$ or $(AB)'$ for some functions A and B, then D'Morgan's law may be applied to yield $A'B'$ and $A' + B'$ respectively. This process may be continued until each prime which appears applies only to a single variable x_i.

Next, by applying the distributive law of $(\cdot)$ over $(+)$, f can be reduced to a polynomial.

Now suppose some term t does not contain either x_i or x_i' for some variable x_i. This term may be multiplied by $x_i + x_i'$ without changing the function. Continuing this process for each missing variable in each of the terms in f will give an equivalent function whose terms contain x_j or x_j' for each $j = 1, 2, \ldots, n$.

Finally by idempotent property, duplicate terms are eliminated and this completes the proof. □

The following is an illustration.

Example 8.2.3. Write the function $f = (xy' + xz)' + x'$ in disjunctive normal form.

We have,

$$\begin{aligned}
(xy' + xz)' + x' &= (xy')'(xz)' + x' \\
&= (x' + y)(x' + z') + x' \\
&= x' + x'y + yz' + x' \\
&= x'(x + y')(z + z') + yz'(x + x') \\
&= x'yz + x'yz' + x'y'z + x'y'z' + xyz' + x'yz' \\
&= x'y'z + xyz' + x'yz' + x'y'z + x'y'z'.
\end{aligned}$$

The usefulness of the normal form lies primarily in the fact that each function uniquely determines a normal form in a given number of variables, as we shall see in later theorems. However, any function may be placed in normal form in more than one way by changing the number of variables. For example, $f = xy$ is in normal form in x and y, but if xy is multiplied by $z + z'$, then $f = xyz + xyz'$ also in normal form in the variables x, y, and z. Similarly, $g = x'yz + xyz + x'yz' - xyz'$ is in normal form in x, y, and z, but reduces, on

factoring, to $g = x'y - xy$, which is in normal form in x and y. From now on we shall assume that unless stated otherwise, disjunctive normal form refers to that disjunctive normal form which contains the smallest possible number of variables. With this exception, we will be able to show that the normal form of a function is uniquely determined by the function.

Suppose that we desire to select a single term out of the possible terms in a disjunctive normal form in n variables. This corresponds to selecting either x_i or x_i', for each of the n variables x_i, $i = 1, 2, \ldots, n$. Thus there are exactly 2^n distinct terms which may occur in a normal form in n variables.

Theorem 8.2.4. That disjunctive normal form in n variables which contains 2^n terms is called the complete disjunctive normal form in n variables.

It will be a consequence of the following theorems that the complete disjunctive normal form is identically 1. A simple argument to prove this directly is to note that for any variable x_j, the coefficients of x_j and x_j' must be identical in a complete normal form, namely, these coefficients are each the complete normal form in the remaining $n - 1$ variables. Factoring serves to eliminate x_j, and this process may be repeated to eliminate each variable in succession, thus reducing the expression to 1.

Theorem 8.2.5. If each of n variables is assigned the value 0 or 1 in an arbitrary, but fixed manner, then exactly one term of the complete disjunctive normal form in the n variables will have the value 1 and all other terms will have the value 0.

Proof. Let $a_1, a_2, \ldots, a_n$ represent the values assigned to $x_1, x_2, \ldots, x_n$ in that order, where each a_i is 0 or 1. Select a term from the complete normal form as follows: use x_i if $a_i = 1$, and use x_i' if $a_i = 0$ for each x_i, $i = 1, 2, \ldots, n$. The term so selected is then a product of n ones, and hence is 1. All other terms in the complete normal form will contain at least one factor 0 and hence will be 0. □

Corollary 8.2.6. Two functions are equal if and only if their respective disjunctive normal forms contain the same terms.

Proof. Two functions with the same terms are obviously equal. Conversely, if two functions are equal, then they must have the same value for every choice of value for each variable. In particular, they assume the same value for each set of values 0 and 1 which may be assigned to the variables. By idempotent property, the combinations of values of 0 and 1 which, when assigned to the variables, make the function assume the value 1 uniquely determine the terms which are present in the normal form for the function. Hence both normal forms contain the same terms. □

Corollary 8.2.7. To establish any identity in Boolean algebra, it is sufficient to check the value of each function for all combinations of 0 and 1 which may be assigned to the variables.

We have seen in the preceding theorems that a function is completely determined by the values it assumes for each possible assignment of 0 and 1 to the respective variables. This suggests that functions could be conveniently specified by giving a table to represent such properties. In applications, particularly to the design of circuits, this is precisely the way in which Boolean functions are constructed. If such a table has been given, then the function, in disjunctive normal form, may be written down by inspection. For each set of conditions for which the function is to be 1, a corresponding term is included in the disjunctive normal form selected, as indicated in the proof of the idempotent property in the previous unit. The sum of these terms gives the function, although not necessarily in simplest form. The following example indicates this method.

Example 8.2.8. Find and simplify the function $f(x, y, z)$ specified by table 8.1

Note that the table shows the value of f for each of the $2^3 = 8$ possible assignments of 0 and 1 to x, y, and z.

Row	x	y	z	$f(x,y,z)$
1	1	1	1	0
2	1	1	0	1
3	1	0	1	1
4	1	0	0	0
5	0	1	1	0
6	0	1	0	0
7	0	0	1	1
8	0	0	0	0

Table 8.1

We observe that for the combinations represented by rows 2, 3 and 7 of the table, the function will have the value 1. Thus the disjunctive normal form of f will contain three terms. For 2, since the x variable is 1, y variable is 1 and z variable is zero, the term in f corresponding to this combination will be xyz' (note that the value if 1 by idempotent property). Similarly, for the terms in 3 and 7th rows, we get $xy'z$ and $x'y'z$ respectively (each giving values 1). Thus, summing these terms over, we get $f(x,y,z) = xyz' + xy'z + x'y'z$. We have

$$\begin{aligned} f(x,y,z) &= xyz' + xy'z + x'y'z \\ &= xyz' + (x+x')y'z \\ &= xyz' + y'z. \end{aligned}$$

Exercise 8.2.9. 1. Express the following in disjunctive normal form in the smallest possible number of variables:

(a) $x'yz + xy'z' + x'y'z + x'yz' + xy'z + x'y'z'$

(b) $(x+y')(y+z')(z+x')(x'+y')$

(c) $(u+v+w)(uv+u'w)'$

(d) $xy' + xz + xy$

(e) $xyz + (x+y)(x+z)$

(f) $x + x'y$

(g) $(x+y)(x+y')(x'+z)$

2. Write separately, and simplify, the three functions f_1, f_2 and f_3 as given in the table 8.3.

8.3 Conjunctive Normal Form

There are other normal forms, besides the disjunctive normal form, which are equally useful. One of these represents each function as a product of sums, rather than as a sum of products. If each statement in the preceding section were replaced by its dual, the resulting discussion would be a corresponding treatment of this second form called the conjunctive normal form. To make this clear, the definition and theorems are repeated here in their dual forms. No proofs are needed, of course, because of the principle of duality.

Row	x	y	z	f_1	f_2	f_3
1	1	1	1	0	0	1
2	1	1	0	1	1	1
3	1	0	1	0	1	0
4	1	0	0	1	0	0
5	0	1	1	0	0	0
6	0	1	0	0	1	0
7	0	0	1	0	1	1
8	0	0	0	0	0	1

Table 8.2

Definition 8.3.1. A Boolean function is said to be in conjunctive normal form in n variables $x_1, x_2, \ldots, x_n$ for $n > 0$, if the function is a product of factors of the type $f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n)$, where $f + i(x_i)$ is x_i or x_i' for each $i = 1, 2, \ldots, n$, and no two factors are identical. In addition, 0 and 1 are said to be in conjunctive normal form in n variables for $n \geq 0$.

Theorem 8.3.2. Every function in a Boolean algebra which contains no constants is equal to a function in conjunctive normal form.

Example 8.3.3. Write the function $(xy' + xz)' + x'$ in conjunctive normal form.

The procedure is essentially dual to that of the disjunctive normal form that we saw in the previous section, although, depending on the initial form of the function, it may require more steps to perform the reduction in one case than in another. Here, after primes are removed from parentheses, the function is factored into linear factors and then extra variables are introduced as needed by adding, within each factor, products of the form ww'. The final step is to expand into linear factors again and remove like factors. The solution for this example is given by the steps below.

$$\begin{aligned}
(xy' + xz)' + x' &= (x' + y)(x' + z') + x' \\
&= (x' + x' + y)(x' + x' + z') \\
&= (x' + y)(x' + z') \\
&= (x' + y + zz')(x + z' + yy') \\
&= (x' + y + z)(x' + y + z')(x' + y + z')(x' + y' + z') \\
&= (x' + y + z)(x' + y + z')(x' + y' + z').
\end{aligned}$$

Definition 8.3.4. That conjunctive normal form in n variables which contains 2^n factors is called the complete conjunctive normal form in n variables.

Theorem 8.3.5. If each of n variables is assigned the value 0 or 1 in an arbitrary, but fixed manner, then exactly one factor of the complete conjunctive normal form in the n variables will have the value 0 and all other factors will have the value 1.

Note that to select the factor which will be 0 when a set of values $a_1, a_2, \ldots, a_n$ are assigned to $x_1, x_2, \ldots, x_n$ in that order, where each a_i is 0 or 1, we simply apply duality principle of that described in the previous section. x_i is selected if $a_i = 0$ and x_i' is selected if $a_i = 1$ for each $i = 1, 2, \ldots, n$. The proper factor is then the sum of these letters, each of which has value 0. All other factors have the value 1.

Corollary 8.3.6. Two functions, each expressed in conjunctive normal form in n variables, are equal if and only if they contain identical factors.

Row	x	y	z	$f(x,y,z)$
1	1	1	1	1
2	1	1	0	1
3	1	0	1	0
4	1	0	0	1
5	0	1	1	1
6	0	1	0	1
7	0	0	1	0
8	0	0	0	1

Example 8.3.7. Find and simplify the function $f(x, y, z)$ specified in the table above.

Observe that only two rows of the table show the value 0 for f. Corresponding to the third row, we see that x is 1, y is 0 and z is 1. So the corresponding factor will be $x' + y + z'$. Similarly, for the 7th row, we have the term as $x + y + z'$. Thus, we would have $f(x, y, z) = (x' + y + z')(x + y + z')$ which gives us,

$$\begin{aligned} f(x,y,z) &= (x'+y+z')(x+y+z') \\ &= x'y + x'z' + y + yz' + z'x + z'y \\ &= (x'+1)y + z'(x'+y+x+y) \\ &= y + z'(x'+x+y) \\ &= y + z'(1+y) \\ &= y + z'. \end{aligned}$$

In problems of this type, the disjunctive normal form would normally be used if the number of 1's is were less than the number of 0's in the f column, and the conjunctive normal form would be used if the number of 0's were less than the number of l's.

Again, as in the previous section, we can use the conjunctive normal form to find complements of functions written in this form by inspection. The complement of any function written in conjunctive normal form is that function whose factors are exactly those factors of the complete conjunctive normal form which are missing from the given function. For example, the complement of $(x + y')(x' + y)$ is $(x + y)(x' + y')$.

It may be desirable to change a function from one normal form to the other. This can be done more readily than by following the general procedure for converting a function to a particular form. An example will illustrate the method, which is based on the fact that $(f')' = f$.

Example 8.3.8. Find the conjunctive normal form for the function

$$f = xyz + x'yz + xy'z' + x'yz'.$$

We have,

$$\begin{aligned} f &= xyz + x'yz + xy'z' + x'yz' \\ &= [(xyz + x'yz + xy'z' + x'yz')']' \\ &= [(x'+y'+z')(x+y'+,z')(x'+y+z)(x+y'+z)]' \\ &= (x+y+z)(x'+y+z')(x+y+z')(x'+y'+z). \end{aligned}$$

Here, the first complement was taken with the aid of D'Morgan's law and the second complement was taken by the method discussed above. These steps could have been reversed, with the same results. A similar procedure will change a function from conjunctive normal form to disjunctive normal form.

Exercise 8.3.9. 1. Express each of the following in conjunctive normal form in the smallest possible number of variables:

(a) $xyz + (x + y)(x + z)$

(b) $(x'y + xyz' + xy'z + x'y'z't + t')'$

(c) $x'yz + xy'z' + x'y'z + x'yz' + xy'z + x'y'z'$

(d) $(x + y')(y + z')(z + x')(x' + y')$

(e) $(u + v + w)(uv + u'w)'$

(f) $xy' + xz + xy$

(g) $xyz + (x + y)(x + z)$

2. Change each of the following from disjunctive normal form to conjunctive normal form:

(a) $uv + u'v + u'v'$

(b) $abc + ab'c' + a'bc' + a'b'c + a'b'c'$

3. Change each of the following from conjunctive normal form to disjunctive normal form:

(a) $(x + y')(x' + y)(x' + y')$

(b) $(u + v + w)(u + v + w')(u + v + w)(u' + v + w')(u' + v' + w)(u' + v' + w')$

4. Write separately, and simplify, the four functions f_1, f_2, f_3 and f_4 as given in the table below. Use whichever normal form seems easier.

Row	x	y	z	f_1	f_2	f_3	f_4
1	1	1	1	1	0	0	1
2	1	1	0	0	1	1	1
3	1	0	1	1	0	0	1
4	1	0	0	1	0	1	0
5	0	1	1	1	0	1	1
6	0	1	0	1	0	1	1
7	0	0	1	0	1	0	1
8	0	0	0	1	0	0	0

Table 8.3

8.4 Propositions and definitions of symbols

In the algebra of sets, it is necessary to start with certain primitive concepts in the form of undefined terms. This is typical of any formal system and is true of the algebra of propositions as well. The terms **true**, **false**, and proposition will be taken here as undefined. Without any attempt to investigate the philosophical meaning of truth and falsehood, we will assume that the words true and false are attributes which apply to propositions. By a proposition, we will infer the content of meaning of any declarative sentence which is free of ambiguity

and which has the property that it is either true or false, but not both. The following examples are typical propositions:

3 is a prime number;

living creatures exist on the planet Venus.

Note that of these propositions, the first is known to be true, while the second is either true or false. In contrast to these, the following is not a proposition:

this statement you are reading is false.

We shall use lower case italic letters to represent propositions. Where no specific proposition is given, these will be called propositional variables and used to represent arbitrary propositions.

From any proposition, or set of propositions, other propositions may be formed. The simplest example is that of forming from the proposition p, the negation of p, denoted by $\neg p$ or p'. or example, suppose that p is the proposition

sleeping is pleasant.

has negation

sleeping is unpleasant.

Any two propositions p and q may be combined in various ways to form new propositions. To illustrate, let p be the proposition

ice is cold,

and let q be the proposition

blood is green.

These propositions may be combined by the connective **and** to form the proposition

ice is cold and blood is green.

This proposition is referred to as the **conjunction** of p and q. We will denote the conjunction of p and q by pq, and we will require that the proposition be true in those cases in which both p and q are true, and false in cases in which either one or both of p and q are false.

Another way in which the propositions in the preceding paragraph may be combined is indicated in the proposition

either ice is cold or blood is green.

This proposition is referred to as the **disjunction** of p and q. We will denote the disjunction of p and q by $p+q$ and is the proposition "either p or q or both". We will require that this proposition be true whenever either one of p and q or both are true, and false only when both are false.

It follows from our definitions that the negation of "p or q" is the proposition "not p and not q," which can also be stated "neither p nor q." Likewise, the negation of "p and q" is "either not p or not q. " That is, the laws of D'Morgan hold for propositions just as they do for sets. In symbolic form we have the following laws for propositions:

$$\begin{aligned}(p+q)' &= p'q' \\ (pq)' &= p'+q'.\end{aligned}$$

Example 8.4.1. Let p be the proposition "missiles are costly" and q be the proposition "Grandpa chews gum". Write in English the propositions represented by the symbols

1. $p + q'$ 2. $p'q'$ 3. $pq' + p'q$

We have

$$\begin{aligned} p &: \text{missiles are costly;} \\ q &: \text{Grandpa chews gum;} \\ p' &: \text{missiles are not costly;} \\ q' &: \text{Grandpa does not chew gum;} \end{aligned}$$

Then,

1. $p + q'$: Either missiles are costly or Grandpa does not chew gum.
2. $p'q'$: Missiles are not costly and Grandpa does not chew gum.
3. $pq' + p'q$: Either missiles are costly and Grandpa does not chew gum, or missiles are not costly and Grandpa chews gum.

Exercise 8.4.2. 1. Which of the following sentences, or phrases, represent propositions?

(a) Grass is yellow.

(b) Beautiful white roses.

(c) All mathematics is difficult, and some mathematics is impossible.

2. Let p be the proposition "mathematics is easy," and let q be the proposition "two is less than three." Write out, in reasonable English, the propositions represented by

i. $p + q$ ii. $pq' + p'q$

8.5 Truth tables

To show that the set of propositions and the operations of conjunction, disjunction, and negation form a Boolean algebra, it is necessary first to define the concept of equality. Two propositional functions g and h, each functions of the n propositional variables $p_1, p_2, \ldots, p_n$, are said to be equal if and only if they have the same truth value for every possible way of assigning truth values to each of the n variables. To complete our algebra, we will create two new propositions represented by 0 and 1, respectively. We define 0 to be a proposition that is always false, and 1 to be a proposition that is always true. The equation $p = 0$ is equivalent to the statement that p is false. Similarly, $q = 1$ is equivalent to saying that q is true.

The definition we have given for equality makes it possible to represent a function with a table of values exactly as was done previously. The only difference is that now we have a special meaning attached to the symbols which appear in the table. These symbols stand for propositions rather than for abstract elements of an arbitrary Boolean algebra. Such a table will be termed a **truth table**. We give an example of such a table below.

This table represents the truth values of the propositions pq and $p + q$ for two simple propositions p and q according to their truth values.

The construction of a truth table for a complicated propositional function can best be carried out in steps, using at each step the basic truth table for one of the operations $(+)$, $(\cdot)$ or complement.

If it happens that the truth table for a function contains only 1's (in the function column), we call the corresponding proposition a **tautology**. Both $p + p'$ and $pq + pq' + p'q + p'q'$ are examples of tautologies for any propositions p and q.

Row	p	q	pq	$p+q$
1	1	1	1	1
2	1	0	0	1
3	0	1	0	1
4	0	0	0	0

We again attempt to draw the truth table of $(r' + pq)'$, for three propositions p, q and r. Since we have three propositions and two truth values, viz., 0 and 1, so we have $2^3 = 8$ possible combinations of truth values for the propositions. The table is given in table 8.4.

Row	p	q	r	r'	pq	$r' + pq$	$(r' + pq)'$
1	1	1	1	0	1	1	0
2	1	1	0	1	1	1	0
3	1	0	1	0	0	0	1
4	1	0	0	1	0	1	0
5	0	1	1	0	0	0	1
6	0	1	0	1	0	1	0
7	0	0	1	0	0	0	1
8	0	0	0	1	0	1	0

Table 8.4

An illustration of the usefulness of truth tables occurs in the proof of the following theorem. From the definition of equality, it follows that two functions are equal if and only if their truth tables are identical. This fact is used in the third part of the proof below.

Theorem 8.5.1. The algebra of propositions is a Boolean algebra.

Proof. In order to prove that the set of propositions forms a Boolean algebra, we will have to show that the four postulates hold which we stated in the beginning of the previous unit. We begin with them one by one.

(a) From the definition of disjunction and conjunction of propositions (denoted as $(\cdot)$ and $(+)$), it follows that they are commutative and hence the first postulate holds true.

(b) 0 is the identity element for the operation $(+)$ since $0 + p$ has the same truth value as p and hence equals p. Similarly, $(1)(q)$ has the same truth value as q and hence equals q, showing that 1 is the identity for the operation of conjunction.

(c) Each operation is distributive over the other as is shown by the table below (table 8.5).

From table 8.5, it can be seen that the truth values of $p + qr$ and $(p + q)(p + r)$ are same, and hence they are equal. Also, the truth values of $pq + pr$ and $p(q + r)$ are same and hence they are equal.

(d) For each proposition p, there is a second proposition p', the negation of p, which satisfies the relations $pp' = 0$ and $p + p' = 1$ as can be verified by the truth table 8.6.

Thus, p' is the complement of p.

Hence the theorem. □

p	q	r	pq	pr	qr	$p+qr$	$pq+pr$	$p+q$	$p+r$	$q+r$	$p(q+r)$	$(p+q)(p+r)$
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	0	1	0	0	1	1	1	1	1	1	1
1	0	1	0	1	0	1	1	1	1	1	1	1
1	0	0	0	0	0	1	0	1	1	0	0	1
0	1	1	0	0	1	1	0	1	1	1	0	1
0	1	0	0	0	0	0	0	1	0	1	0	0
0	0	1	0	0	0	0	0	0	1	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0

Table 8.5

p	p'	pp'	$p+p'$
1	0	0	1
0	1	0	1

Table 8.6

Exercise 8.5.2. 1. Determine which of the following are tautologies by constructing the truth table for each.

(a) $pq+p'+q'$ 2. $p+q+p'$

2. Construct a truth table for each of the following functions.

(a) $pqr+p'qr'+p'q'r'$ 2. $(p'+qr)'(pq+q'r)$ 3. $pq'+p'(qr+q'r)'$

Few Probable Questions

1. Define disjunctive normal form. Find the disjunctive normal form for the function $f(x,y,z)=(x+y')(y+z')(z+x')(x'+y')$, in the smallest possible number of variables.
2. Define conjunctive normal form. Find the conjunctive normal form for the function $f(x,y,z)=xy'+xz+xy$, in the smallest possible number of variables.
3. Convert f from cnf to dnf where $f(x,y)=(x+y')(x'+y)(x'+y')$.
4. Let p be the proposition "x is an even number," and let q be the proposition "x is the product of two integers." Translate into symbols each of the following propositions.

   (a) Either x is an even number, or x is a product of two integers.

(b) Either x is an even number and a product of integers, or x is an odd number and is not a product of integers.

(c) x is neither an even number nor a product of integers.

5. Write, in reasonable English, the negation of each of the following propositions.

(a) Either good health is desirable, or I have been misinformed.

(b) Oranges are not suitable for use in vegetable salads.

(c) There is a number which, when added to 6, gives a sum of 13.

6. Construct the truth table for $(p' + qr)'(pq + q'r)$.

Unit 9

Course Structure

- Logic gates and circuits, Applications of Boolean Algebra to Switching theory (using AND, OR, and NOT gates),
- Karnaugh Map method.

9.1 Introduction

In this unit, we will introduce a third important application of Boolean algebra, the algebra of circuits, involving two-state (bistable) devices. The simplest example of such a device is a switch or contact. The theory introduced holds equally well for such two-state devices as rectifying diodes, magnetic cores, transistors, various types of electron tubes, etc. The nature of the two states varies with the device and includes conducting versus nonconducting, closed versus open, charged versus discharged, magnetized versus nonmagnetized, high-potential versus low-potential, and others. The algebra of circuits is receiving more attention at present, both from mathematicians and from engineers, than either of the two applications of Boolean algebra which we considered in the previous chapters. The importance of the subject is reflected in the use of Boolean algebra in the design and simplification of complex circuits involved in electronic computers, dial telephone switching systems, and many varied kinds of electronic control devices. The algebra of circuits fits into the general picture of Boolean algebra as an algebra with two elements 0 and 1. This means that except for the terminology and meaning connecting it with circuits, it is identical with the algebra of propositions considered as an abstract system. Either of these Boolean algebras is much more restricted than an algebra of sets.

Objectives

After reading this unit, you will be able to

- learn basic elements of a switching circuit
- learn to minimize a switching circuit using Boolean function
- define the logical circuit elements
- learn to simplify functions using Karnaugh maps

9.2 Switching Circuits

For the present, we will limit our discussion to the simplest kinds of circuits, those involving only switches. We will designate a switch by a single letter $a, b, c, x, y, \ldots$. If two switches operate so that they open and close simultaneously, we designate them by the same letter. If they operate so that the first is always open when the second is closed, and closed when the second is open, we denote the first by a letter, say x, and the second by x' (or, equally well, the first by x' and the second by x).

A circuit consisting of two switches x and y connected in parallel is denoted by $x+y$, and a circuit consisting of x and y connected in series is denoted by xy. Thus to each series-parallel circuit, there corresponds an algebraic expression; and conversely to each algebraic expression involving only $(+)$, $(\cdot)$ and negation, there corresponds a circuit (fig 9.2.1). We will speak of this relationship by saying that the function represents the circuit, and the circuit realizes the function. We will agree to assign the value 1 to a letter if it represents a

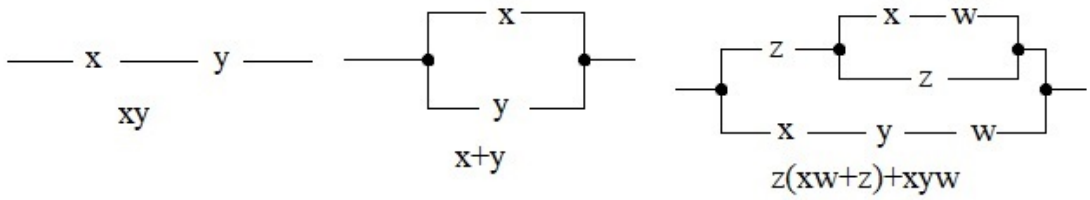


Figure 9.2.1

closed switch, and the value 0 if it represents an open switch. If a and a' both appear, then a is 1 if and only if a' is 0. A switch that is always closed is represented by 1, one that is always open by 0. Letters play the role of variables which take on the value 0 or 1, and we note the close analogy to proposition variables, which have the same possible values, although the meaning attached to these values has changed.

Two circuits involving switches $a, b, \ldots$ are said to be equivalent if the closure conditions of the two circuits are the same for any given position of the switches involved (values of the variables $a, b, \ldots$). That is, they are equivalent if for every position of the switches, current may either pass through both (both closed) or not pass through either (both open). Two algebraic expressions are defined to be equal if and only if they represent equivalent circuits.

It is now possible, by drawing the appropriate circuits and enumerating the possible positions of the switches involved, to check that each of the laws of Boolean algebra is valid when interpreted in terms of switching circuits. For example, consider the circuits that realize the functions on each side of the identity stating the distributive law for $(+)$ over $(\cdot)$, shown in figure 9.2.2. By inspection, it is apparent that the circuit

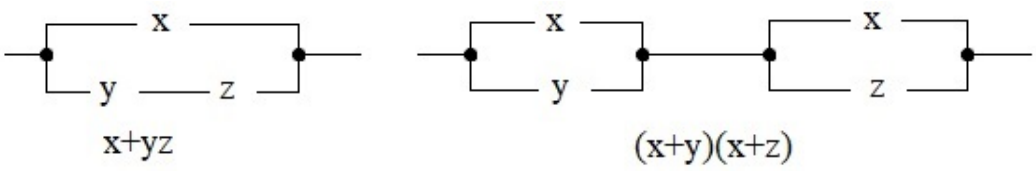


Figure 9.2.2

is closed (current can pass) if switch x is closed, or if both y and z are closed, and that the circuit is open (current cannot pass) if x and either y or z are open. Hence the circuits are equivalent, and this distributive law holds.

A simpler procedure for checking the validity of the fundamental laws is to note that numerical values of the switching functions a', ab, and $a+b$ are identical to the truth tables for the corresponding propositional functions (table 9.1).

Row	a	b	a'	ab	$a+b$
1	1	1	0	1	1
2	1	0	0	0	1
3	0	1	1	0	1
4	0	0	1	0	0

Table 9.1: Closure Properties of switching functions a', ab and $a+b$

Example 9.2.1. We want to find a circuit which realizes the Boolean function $xyz' + x'(y+z')$.

This expression indicates a series connection of x, y, and z' in parallel with a circuit corresponding to $x'(y'+z')$. This latter circuit consists of x' in series with a parallel connection of y and z. Hence the circuit diagram is that shown in fig. 9.2.3.

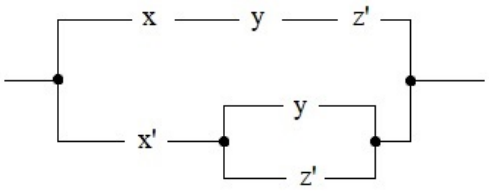


Figure 9.2.3

Example 9.2.2. We want to find the Boolean function which represents the circuit shown in fig. 9.2.4.

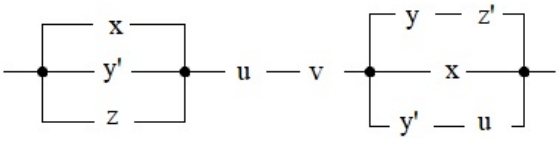


Figure 9.2.4

By inspection, the function is $(x+y'+z)uv(yz'+x+y'u)$.

Example 9.2.3. Construct the table of closure properties for the function $f(x,y,z) = x'y + z(x+y')$.

A table of closure properties for a function is identical, except for interpretation, to a truth table for a propositional function. This function has the closure properties listed in table 9.2.

Exercise 9.2.4.

1. Draw circuits which realize each of the following expressions, without first simplifying the expressions.
   (a) $abc + ab(dc + ef)$
   (b) $a + b(c + de) + fg$
2. Find the function which represents the circuits in the figure 9.2.5.
3. Find circuit which realize the function given in table 9.3.

Row	x	y	z	$x'y$	$x+y'$	$z(x+y')$	$x'y+z(x+y')$
1	1	1	1	0	1	1	1
2	1	1	0	0	1	0	0
3	1	0	1	0	1	1	1
4	1	0	0	0	1	0	0
5	0	1	1	1	0	0	1
6	0	1	0	1	0	0	1
7	0	0	1	0	1	1	1
8	0	0	0	0	1	0	0

Table 9.2

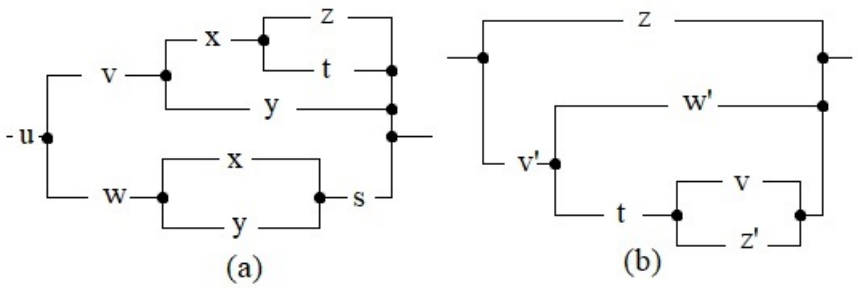


Figure 9.2.5

9.2.1 Simplification of circuits

In the previous section, we showed that the algebra of circuits is a Boolean algebra, and hence all the results proved earlier for Boolean algebras hold. In particular, theorems and rules relating to simplification of Boolean functions apply in the algebra of circuits.

Two basic problems that arise in connection with applications of Boolean algebra to switching circuits are (a) simplification of a given circuit which is known to have the desired closure properties, and (b) the design of circuits with given properties. The design problem will be discussed in later sections, and in this section we will consider the problem of simplifying a given circuit. This problem has often been solved in specific cases by trial-and-error methods. There are several known methods, based on the theory of Boolean functions, for writing schematic charts for simplifying functions. We will emphasize instead a straightforward approach using the properties of Boolean algebras directly to effect reasonable simplifications.

A general method of simplifying a circuit is first to find the Boolean function which represents the circuit, then to simplify the function as we have done repeatedly in earlier sections, and finally to draw a new circuit diagram realizing the simplified function. We give a simple illustration below.

Example 9.2.5. Simplify the circuit in fig. 9.2.6.

This circuit is represented by the Boolean function $(xy+abc)(xy+a'+b'+c')$, which simplifies to xy. Hence the given circuit is equivalent to the series connection of the two switches x and y, with the diagram given in fig. 9.2.7.

In using the basic laws of Boolean algebra, it often happens that a possible simplification is overlooked. It may happen that a certain step is easier to recognize if stated in terms of one of the dual laws rather than in terms of the other. This suggests another method of simplification which may help. To simplify a function f,

Row	x	y	z	f_1
1	1	1	1	0
2	1	1	0	1
3	1	0	1	1
4	1	0	0	0
5	0	1	1	0
6	0	1	0	0
7	0	0	1	0
8	0	0	0	1

Table 9.3

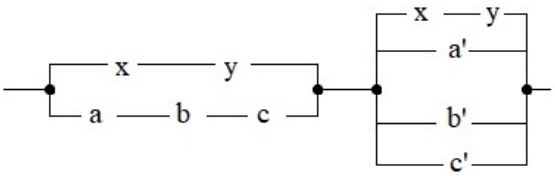


Figure 9.2.6

the dual of f may be taken and the resulting expression simplified. If the dual is taken again, the function f is obtained in a different form. This will usually be simpler than the original.

Example 9.2.6. Simplify the circuit in fig. 9.2.8. The circuit is represented by the function $f = cb + ab'cd + cd' + ac' + a'bc' + b'c'd'$. Consider the first three terms as the function g, and the last three terms as the function h. Then $g = cb + ab'cd + cd'$. The dual of g, which we write as $d(g)$ is then

$$d(g) = (c + b)(a + b' + c + d)(c + d') = c + abd'.$$

Taking the dual again, we find

$$g = c(a + b + d').$$

Similarly,

$$\begin{aligned} h &= ac' + a'bc' + b'c'd' \\ d(h) &= (a + c')(a' + b + c')(b' + c' + d') = c' + abd', \end{aligned}$$

Combining g and h yields

$$f = (c + c')(a + b + d') = a + b + d',$$

which corresponds to the circuit given in fig. 9.2.9.

Exercise 9.2.7. Simplify the circuits given in fig. 9.2.10.

9.3 Logical Circuit elements

Circuit elements involving diodes or vacuum tubes are very common. Rather than discuss the many types of electronic apparatus that may be used, we will introduce the idea of a logical circuit element. It will be

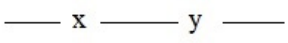


Figure 9.2.7

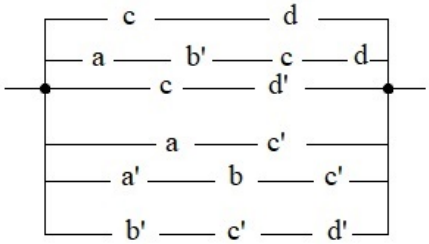


Figure 9.2.8

enough to know that these elements can be constructed; in fact, commercially packaged elements of these types, suitable for use in many types of equipment, can be purchased directly. We will conceive of a logical circuit element as a little box or package with one or more input leads (wire connections) and one or more output leads. These leads will carry signals in the form of positive voltage corresponding to a value 1, or zero voltage corresponding to a value 0. We will use a single letter, say x, to stand for the condition of the lead. When the lead carries a signal, we will say that x takes on the value 1. When the lead does not carry a signal, we say that x has the value 0. This represents only a slight modification of our earlier point of view, where 1 and 0 meant closed or open circuits, since we can think of a closed circuit as one carrying a signal, and of an open circuit as one incapable of carrying a signal. Other signals than that of a positive voltage could be used equally well, and in fact the signal used will in general depend on the type of components used in circuit construction. We will use just this one type of signal for simplicity, and we will adapt all our circuits to its use.

We will draw a circuit element as a circle with a letter inside to designate the type of element, and with lines indicating inputs and outputs. Arrows on these lines will indicate the difference between input and output, an arrow pointing toward the circle being used on each input.

The first logical circuit element we will consider has a single input and a single output. The function of this element is to obtain the complement of a given signal; that is, the output is 0 when the input is 1, and conversely. Fig. 9.3.1 shows the notation we will use, a circle with C in the center. The input is designated x, so the output is x'.

The next two logical circuit elements correspond to the logical connections "and" and "or." Each may have two or more inputs and only a single output. The "and" element is shown in diagrams as a circle with A in the center. This element produces an output signal (output has value 1) if and only if every input carries a signal (has value 1). If the inputs to an "and" element are x, y, and z, for example, the output function may be written as xyz, where the notation is that of Boolean algebra. The "or" element, represented graphically by a circle with O in the center, produces an output signal whenever one or more inputs carry a signal. If the inputs to an "or" element are x, y, and z, for example, the output is the Boolean function $x + y + z$. Fig. 9.3.2 shows the symbolic notations for these elements. Each is shown with only two inputs.

9.4 Karnaugh Maps

A Karnaugh map provides a pictorial method of grouping together expressions with common factors and therefore eliminating unwanted variables. The Karnaugh map can also be described as a special arrangement

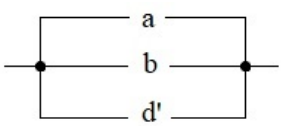


Figure 9.2.9

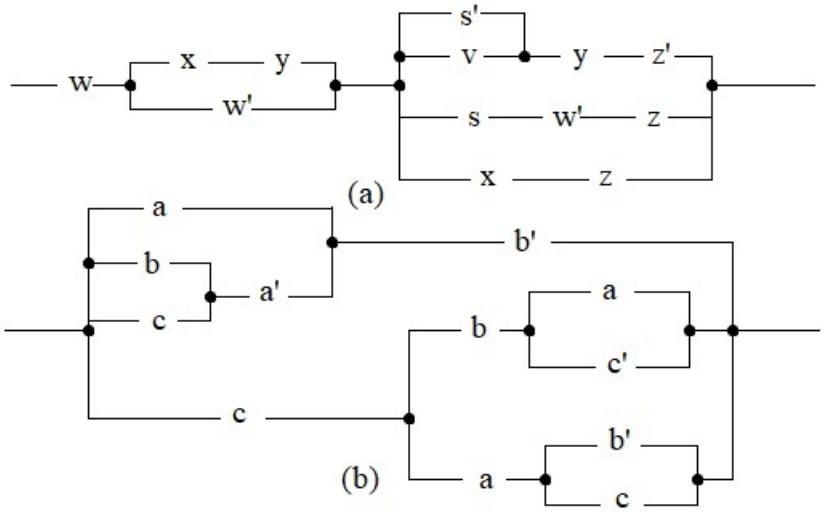


Figure 9.2.10

of a truth table.

The diagram below illustrates the correspondence between the Karnaugh map and the truth table for the general case of a two variable problem (fig. 9.4.1).

The values inside the squares are copied from the output column of the truth table, therefore there is one square in the map for every row in the truth table. Around the edge of the Karnaugh map are the values of the two input variable. x is along the top and y is down the left hand side. The diagram 9.4.2 explains this:

The values around the edge of the map can be thought of as coordinates. So as an example, the square on the top right hand corner of the map in the above diagram has coordinates $x = 1$ and $y = 0$. This square corresponds to the row in the truth table where $x = 1$ and $y = 0$ and $f = 1$. Note that the value in the f column represents a particular function to which the Karnaugh map corresponds.

Example 9.4.1. Consider the following map (fig. 9.4.3). The function plotted is:

$$f(x, y) = xy' + xy.$$

Note that values of the input variables form the rows and columns. That is the logic values of the variables x and y (with one denoting true form and zero denoting false form) form the head of the rows and columns respectively. Bear in mind that the above map is a one dimensional type which can be used to simplify an expression in two variables. There is a two-dimensional map that can be used for up to four variables, and a three-dimensional map for up to six variables.

Using algebraic simplification,

$$f = xy' + xy = x(y' + y) = x.$$

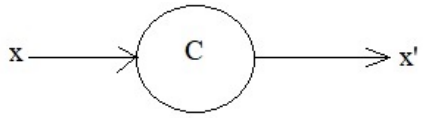


Figure 9.3.1

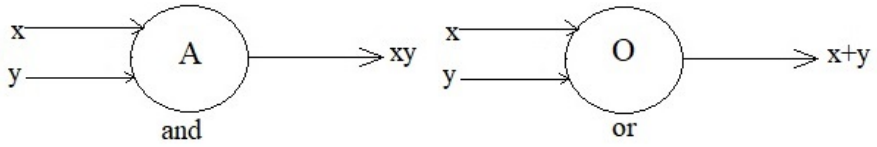


Figure 9.3.2

x	y	f
0	0	a
0	1	b
1	0	c
1	1	d

y \ x	0	1
0	a	b
1	c	d

Figure 9.4.1

x	y	f
0	0	0
0	1	1
1	0	**1**
1	1	1

y \ x	0	1
0	0	**1**
1	1	1

Figure 9.4.2

Variable B becomes redundant due to B4. Referring to the map 9.4.3, the two adjacent 1's are grouped together. Through inspection it can be seen that variable y has its true and false form within the group. This eliminates variable y leaving only variable x which only has its true form. The minimised answer therefore is f.

Example 9.4.2. Consider the expression $f(x, y) = x'y' + xy' + x'y$ plotted on the Karnaugh map 9.4.4. Pairs of 1's are grouped as shown in the figure, and the simplified answer is obtained by using the following steps:

Note that two groups can be formed for the example given above, bearing in mind that the largest rectangular clusters that can be made consist of two 1's. Notice that a 1 can belong to more than one group. The first group labelled I, consists of two 1's which correspond to $x = 0,\ \ y = 0$ and $x = 1,\ \ y = 0$. Put in another way, all squares in this example that correspond to the area of the map where $y = 0$ contains 1's, independent of the value of x. So when $y = 0$, the output is 1. The expression of the output will contain the term y'.

For group labelled II corresponds to the area of the map where $x = 0$. The group can therefore be defined as x'. This implies that when $x = 0$ the output is 1. The output is therefore 1 whenever $y = 0$ and $x = 0$. Hence the simplified answer is $f = x' + y'$.

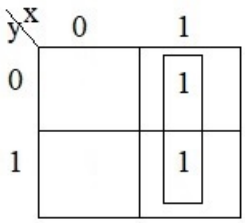


Figure 9.4.3

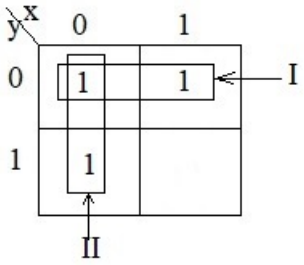


Figure 9.4.4

Few Probable Questions

1. Construct a table of closure properties and draw circuits which realize the function $(a + b' + c)(a + bc') + c'd + d(b' + c)$.

2. Find the function which represents the circuit in fig. 9.4.5.

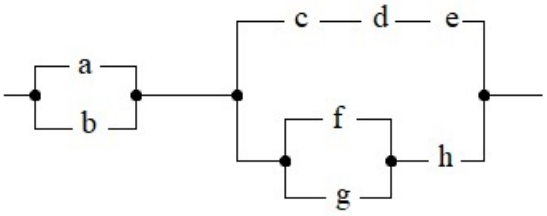


Figure 9.4.5

3. Minimise the following problems using the Karnaugh maps method.

   (a) $f = x'y'z' + x'y + xyz' + xz$;

   (b) $x'y + yz' + yz + xy'z'$.

4. Find circuits which realize each of the functions given in table 9.6.

Row	x	y	z	f_1	f_2
1	1	1	1	1	1
2	1	1	0	0	1
3	1	0	1	0	0
4	1	0	0	1	1
5	0	1	1	1	1
6	0	1	0	1	0
7	0	0	1	0	1
8	0	0	0	1	1

Table 9.6

5. Simplify each of the circuits given in fig. 9.4.6.

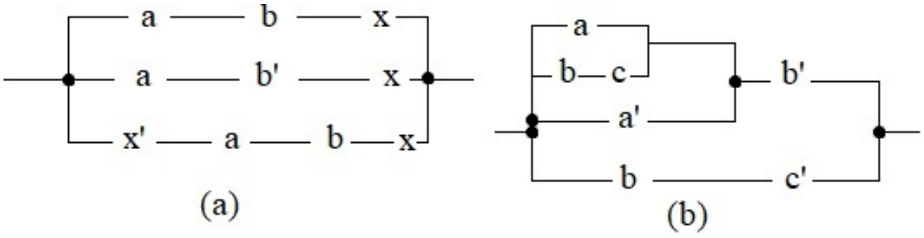


Figure 9.4.6

Unit 10

Course Structure

- Combinatorics: Introduction, Basic counting principles, Permutation and combination, pigeonhole principle, Recurrence relations and generating functions.

10.1 Introduction

Combinatorics studies the way in which discrete structures can be combined or arranged. Enumerative combinatorics concentrates on counting the number of certain combinatorial objects - e.g. the twelvefold way provides a unified framework for counting permutations, combinations and partitions. Analytic combinatorics concerns the enumeration (i.e., determining the number) of combinatorial structures using tools from complex analysis and probability theory. In contrast with enumerative combinatorics which uses explicit combinatorial formulae and generating functions to describe the results, analytic combinatorics aims at obtaining asymptotic formulae. Design theory is a study of combinatorial designs, which are collections of subsets with certain intersection properties. Partition theory studies various enumeration and asymptotic problems related to integer partitions, and is closely related to q-series, special functions and orthogonal polynomials. Originally a part of number theory and analysis, partition theory is now considered a part of combinatorics or an independent field. Order theory is the study of partially ordered sets, both finite and infinite.

Objectives

After reading this unit, you will be able to

- learn the sum rule and product rule principles and solve examples related to them
- learn various mathematical functions such as factorial function, and solve examples related to them
- define permutation and combination and solve related problems
- learn pigeonhole and generalized pigeonhole principle and solve related sums
- learn Inclusion-Exclusion principle and solve related sums
- define tree diagrams and solve sums related to these

10.2 Basic Counting principles

There are two basic counting principles used throughout this chapter. The first one involves addition and the second one multiplication.

1. **Sum Rule Principle:** Suppose some event E can occur in m ways and a second event F can occur in n ways, and suppose both events cannot occur simultaneously. Then E or F can occur in $m + n$ ways.

2. **Product Rule Principle:** Suppose there is an event E which can occur in m ways and, independent of this event, there is a second event F which can occur in n ways. Then combinations of E and F can occur in mn ways.

The above principles can be extended to three or more events. That is, suppose an event E_1 can occur in n_1 ways, a second event E_2 can occur in n_2 ways, and, following E_2; a third event E_3 can occur in n_3 ways, and so on.

Sum Rule: If no two events can occur at the same time, then one of the events can occur in:

$$n_1 + n_2 + \cdots \text{ ways.}$$

Product Rule: If the events occur one after the other, then all the events can occur in the order indicated in:

$$n_1 \cdot n_2 \cdots \text{ ways.}$$

Example 10.2.1. Suppose a college has 3 different history courses, 4 different literature courses, and 2 different sociology courses.

1. The number m of ways a student can choose one of each kind of courses is $m = 3(4)(2) = 24$.

2. The number n of ways a student can choose just one of the courses is $n = 3 + 4 + 2 = 9$.

There is a set theoretical interpretation of the above two principles. Specifically, suppose $n(A)$ denotes the number of elements in a set A. Then:

1. **Sum Rule Principle:** Suppose A and B are disjoint sets. Then

$$n(A \cup B) = n(A) + n(B).$$

2. **Product Rule Principle:** Let $A \times B$ be the Cartesian product of sets A and B. Then

$$n(A \times B) = n(A) \cdot n(B).$$

Example 10.2.2. There are four bus lines between A and B, and three bus lines between B and C. Find the number m of ways that a man can travel by bus: (a) from A to C by way of B; (b) roundtrip from A to C by way of B; (c) roundtrip from A to C by way of B but without using a bus line more than once.

(a) There are 4 ways to go from A to B and 3 ways from B to C; hence $n = 4 \cdot 3 = 12$.

(b) There are 12 ways to go from A to C by way of B, and 12 ways to return. Thus $n = 12 \cdot 12 = 144$.

(c) The man will travel from A to B to C to B to A. Enter these letters with connecting arrows as follows:

$$A \to B \to C \to B \to A.$$

The man can travel four ways from A to B and three ways from B to C, but he can only travel two ways from C to B and three ways from B to A since he does not want to use a bus line more than once. Enter these numbers above the corresponding arrows as follows:

$$A \overset{4}{\to} B \overset{3}{\to} C \overset{2}{\to} B \overset{3}{\to} A.$$

Thus, by the Product Rule, $n = 4 \cdot 3 \cdot 2 \cdot 3 = 72$.

10.3 Mathematical Functions

We discuss two important mathematical functions frequently used in combinatorics.

10.3.1 Factorial Function

The product of the positive integers from 1 to n inclusive is denoted by $n!$, read "n factorial." Namely:

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-2)(n-1)n = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1.$$

Accordingly, $1! = 1$, $n! = n(n-1)!$. It is also convenient to define $0! = 1$.

10.3.2 Binomial Coefficients

The symbol $\binom{n}{r}$, read nCr, or "n Choose r", where r and n are positive integers with $r \leq b$, is defined as follows

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Note that $n - (n-r) = r$. This yields the following lemma.

Lemma 10.3.1. $\binom{n}{n-r} = \binom{n}{r}$ or equivalently, $\binom{n}{a} = \binom{n}{b}$, where $a + b = n$.

Motivated by that fact that we defined $0! = 1$, we define:

$$\binom{n}{0} = \frac{n!}{0!n!} = 1 \quad \text{and} \quad \binom{0}{0} = \frac{0!}{0!0!} = 1.$$

Binomial Coefficients and Pascal's Triangle

The numbers $\binom{n}{r}$ are called binomial coefficients, since they appear as the coefficients in the expansion of $(a+b)^n$. Specifically:

Theorem 10.3.2. (Binomial Theorem)

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

The coefficients of the successive powers of $a + b$ can be arranged in a triangular array of numbers, called Pascal's triangle, as pictured in fig. 10.3.1. The numbers in Pascal's triangle have the following interesting properties:

$$(a+b)^0 = 1$$
$$(a+b)^1 = a + b$$
$$(a+b)^2 = a^2 + 2ab + b^2$$
$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$
$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$
$$(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$$
$$(a+b)^6 = a^6 + 6a^5b + 15a^4b^2 + 20a^3b^3 + 15a^2b^4 + 6ab^5 + b^6$$

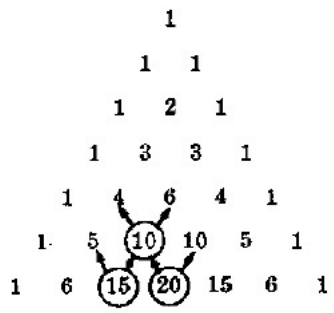


Figure 10.3.1

1. The first and last number in each row is 1.
2. Every other number can be obtained by adding the two numbers appearing above it.

Since these numbers are binomial coefficients, we state the above property formally.

Theorem 10.3.3.
$$\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}.$$

Exercise 10.3.4. 1. Compute: (a) $4!$, $5!$; (b) $6!, 7!, 8!, 9!$; (c) $50!$ [Hint: For large n, use Sterling's approximation: $n! = \sqrt{2\pi n} n^{\pi}\, \mathrm{e}^{-n}$, where $e \approx 2.718$].

2. Compute: (a) $\binom{18}{5}$, (b) $\binom{12}{4}$

3. Prove
$$\binom{17}{6} = \binom{16}{5} + \binom{16}{6}.$$

10.4 Permutations

Definition 10.4.1. Any arrangement of a set of n objects in a given order is called a permutation of the object (taken all at a time). Any arrangement of any $r \leq n$ of these objects in a given order is called an "r-permutation" or "a permutation of the n objects taken r at a time."

Consider, for example, the set of letters A, B, C, D. Then:

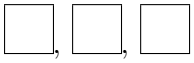
- BDCA, DCBA, and ACDB are permutations of the four letters (taken all at a time).
- BAD, ACB, DBC are permutations of the four letters taken three at a time.
- AD, BC, CA are permutations of the four letters taken two at a time.

We usually are interested in the number of such permutations without listing them. The number of permutations of n objects taken r at a time will be denoted by $P(n, r)$. We have the following theorem.

Theorem 10.4.2.
$$P(n, r) = \frac{n!}{(n-r)!}.$$

We emphasize that there are r factors in $n(n-1)(n-2)\cdots(n-r+1)$.

Example 10.4.3. Find the number m of permutations of six objects, say, A, B, C, D, E, F, taken three at a time. In other words, find the number of "three-letter words" using only the given six letters without repetition. Let us represent the general three-letter word by the following three positions:



The first letter can be chosen in 6 ways; following this the second letter can be chosen in 5 ways; and, finally, the third letter can be chosen in 4 ways. Write each number in its appropriate position as follows:

By the Product Rule there are $m = 6 \cdot 5 \cdot 4 = 120$ possible three-letter words without repetition from the six letters. Namely, there are 120 permutations of 6 objects taken 3 at a time. This agrees with the formula in the previous theorem.

$$P(6, 30) = 6 \cdot 5 \cdot 4 = 120.$$

Consider now the special case of $P(n, r)$ when $r = n$. We get the following result.

Corollary 10.4.4. There are $n!$ permutations of n objects (taken all at a time).

For example, there are $3! = 6$ permutations of the three letters A, B, C. These are:

$$ABC,\ ACB,\ BAC,\ BCA,\ CAB,\ CBA.$$

10.4.1 Permutations with Repetitions

Frequently we want to know the number of permutations of a multiset, that is, a set of objects some of which are alike. We will let

$$P(n; n_1, n_2, \ldots, n_t)$$

denote the number of permutations of n objects of which n_1 are alike, n_2 are alike, ..., n_t are alike.

Theorem 10.4.5. We have,

$$P(n; n_1, n_2, \ldots, n_t) = \frac{n!}{n_1!n_2!\cdots n_t!}.$$

We indicate the proof of the above theorem by a particular example. Suppose we want to form all possible five-letter "words" using the letters from the word "BABBY." Now there are $5! = 120$ permutations of the objects B_1, A, B_2, B_3, Y, where the three B's are distinguished. Observe that the following six permutations

$$B_1B_2B_3AY,\ \ B_2B_1B_3AY,\ \ B_3B_1B_2AY,\ \ B_1B_3B_2AY,\ \ B_2B_3B_1AY,\ \ B_3B_2B_1AY$$

produce the same word when the subscripts are removed. The 6 comes from the fact that there are $3! = 3 \cdot 2 \cdot 1 = 6$ different ways of placing the three B's in the first three positions in the permutation. This is true for each set of three positions in which the B's can appear. Accordingly, the number of different five-letter words that can be formed using the letters from the word "BABBY" is:

$$P(5; 3) = \frac{5!}{3!} = 20$$

Example 10.4.6. Find the number m of seven-letter words that can be formed using the letters of the word "BENZENE."

We seek the number of permutations of 7 objects of which 3 are alike (the three E's), and 2 are alike (the two N's). Thus,

$$m = P(7; 3, 2) = \frac{7!}{3!2!} = 420.$$

Ordered Samples

Definition 10.4.7. Many problems are concerned with choosing an element from a set S, say, with n elements. When we choose one element after another, say, r times, we call the choice an ordered sample of size r.

We consider two cases.

1. **Sampling with replacement:** Here the element is replaced in the set S before the next element is chosen. Thus, each time there are n ways to choose an element (repetitions are allowed). The Product rule tells us that the number of such samples is:

$$n \cdot n \cdot n \cdots n \ (r \ \text{ factors}) = n^r.$$

2. **Sampling without replacement:** Here the element is not replaced in the set S before the next element is chosen. Thus, there is no repetition in the ordered sample. Such a sample is simply an r-permutation. Thus the number of such samples is:

$$P(n, r) = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}.$$

Example 10.4.8. Three cards are chosen one after the other from a 52-card deck. Find the number m of ways this can be done: (a) with replacement; (b) without replacement.

(a) Each card can be chosen in 52 ways. Thus $m = 52(52)(52) = 140608$.

(b) Here there is no replacement. Thus the first card can be chosen in 52 ways, the second in 51 ways, and the third in 50 ways. Therefore, $m = P(52, 3) = 52(51)(50) = 132600$.

Exercise 10.4.9.

1. Find the number n of distinct permutations that can be formed from all the letters of each word: (a) THOSE; (b) UNUSUAL; (c) SOCIOLOGICAL.
2. Find n if $P(n, 2) = 72$.
3. A class contains 8 students. Find the number n of samples of size 3: (a)With replacement; (b)Without replacement.

10.5 Combinations

Definition 10.5.1. Let S be a set with n elements. A combination of these n elements taken r at a time is any selection of r of the elements where order does not count. Such a selection is called an r-combination; it is simply a subset of S with r elements. The number of such combinations will be denoted by $C(n, r)$.

Before we give the general formula for $C(n, r)$, we consider a special case.

Example 10.5.2. Find the number of combinations of 4 objects, A, B, C, D, taken 3 at a time.
Each combination of three objects determines $3! = 6$ permutations of the objects as follows:

$$\begin{aligned} ABC &: \ ABC,\ ACB,\ BAC,\ BCA,\ CAB,\ CBA \\ ABD &: \ ABD,\ ADB,\ BAD,\ BDA,\ DAB,\ DBA \\ ACD &: \ ACD,\ ADC,\ CAD,\ CDA,\ DAC,\ DCA \\ BCD &: \ BDC,\ BDC,\ CBD,\ CDB,\ DBC,\ DCB. \end{aligned}$$

Thus the number of combinations multiplied by 3! gives us the number of permutations; that is,

$$C(4,3)\cdot 3! = P(4,3) \quad \text{or} \quad C(4,3) = \frac{P(4,3)}{3!}.$$

But $P(4,3) = 4\cdot 3\cdot 2 = 24$ and $3! = 6$; hence $C(4,3) = 4$ as noted above.

As indicated above, any combination of n objects taken r at a time determines $r!$ permutations of the objects in the combination; that is,

$$P(n,r) = r!C(n,r).$$

Accordingly, we obtain the following formula for $C(n,r)$ which we formally state as a theorem.

Theorem 10.5.3. We have,

$$C(n,r) = \frac{P(n,r)}{r!} = \frac{n!}{r!(n-r)!}.$$

Recall that the binomial coefficient $\binom{n}{r}$ was defined as

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Hence,

$$C(n,r) = \binom{n}{r}.$$

We shall use $C(n,r)$ and $\binom{n}{r}$ interchangeably.

Example 10.5.4. A farmer buys 3 cows, 2 pigs, and 4 hens from a man who has 6 cows, 5 pigs, and 8 hens. Find the number m of choices that the farmer has.

The farmer can choose the cows in $C(6,3)$ ways, the pigs in $C(5,2)$ ways, and the hens in $C(8,4)$ ways. Thus the number m of choices follows:

$$m = \binom{6}{3}\binom{5}{2}\binom{8}{4} = 20\cdot 10\cdot 70 = 14000.$$

Example 10.5.5. A class contains 10 students with 6 men and 4 women. We want to find the number n of ways to:

(a) select a 4-member committee from the students. This concerns combinations, not permutations, since order does not count in a committee. There are "10 choose 4" such committees. That is:

$$n = C(10,4) = \binom{10}{4} = 210.$$

(b) select a 4-member committee with 2 men and 2 women. The 2 men can be chosen from the 6 men in $C(6,2)$ ways, and the 2 women can be chosen from the 4 women in $C(4,2)$ ways. Thus, by the Product Rule:

$$n = \binom{6}{2}\binom{4}{2} = 15\cdot 6 = 90.$$

(c) elect a president, vice president, and treasurer. This concerns permutations, not combinations, since order does count. Thus, $n = P(6,3) = 6\cdot 5\cdot 4 = 120$.

Exercise 10.5.6. 1. A box contains 8 blue socks and 6 red socks. Find the number of ways two socks can be drawn from the box if: (a) They can be any color. (b) They must be the same color.

2. Find the number m of committees of 5 with a given chairperson that can be selected from 12 people.

10.6 Pigeonhole Principle

Many results in combinational theory come from the following almost obvious statement.

Theorem 10.6.1. (Pigeonhole Principle) If n pigeonholes are occupied by $n + 1$ or more pigeons, then at least one pigeonhole is occupied by more than one pigeon.

This principle can be applied to many problems where we want to show that a given situation can occur.

Example 10.6.2.

1. Suppose a department contains 13 professors, then two of the professors (pigeons) were born in the same month (pigeonholes).
2. Find the minimum number of elements that one needs to take from the set $S = \{1, 2, \ldots, 9\}$ to be sure that two of the numbers add up to 10.

   Here the pigeonholes are the five sets $\{1, 9\}, \{2, 8\}, \{3, 7\}, \{4, 6\}, \{5\}$. Thus any choice of six elements (pigeons) of S will guarantee that two of the numbers add up to ten.

The Pigeonhole Principle is generalized as follows.

Theorem 10.6.3. (Generalized Pigeonhole Principle) If n pigeonholes are occupied by $kn + 1$ or more pigeons, where k is a positive integer, then at least one pigeonhole is occupied by $k + 1$ or more pigeons.

Example 10.6.4. Find the minimum number of students in a class to be sure that three of them are born in the same month.

Here $n = 12$ months are the pigeonholes, and $k+1 = 3$, so $k = 2$. Hence among any $kn+1 = 25$ students (pigeons), three of them are born in the same month.

Exercise 10.6.5.

1. Find the minimum number of students needed to guarantee that five of them belong to the same class (Freshman, Sophomore, Junior, Senior).
2. Let L be a list (not necessarily in alphabetical order) of the 26 letters in the English alphabet (which consists of 5 vowels, A, E, I, O, U, and 21 consonants).
   (a) Show that L has a sublist consisting of four or more consecutive consonants.
   (b) Assuming L begins with a vowel, say A, show that L has a sublist consisting of five or more consecutive consonants.

10.7 Inclusion-Exclusion Principle

Let A and B be any finite sets. Then we know that

$$n(A \cup B) = n(A) + n(B) - n(A \cap B).$$

In other words, to find the number $n(A \cup B)$ of elements in the union of A and B, we add $n(A)$ and $n(B)$ and then we subtract $n(A \cap B)$. This follows from the fact that, when we add $n(A)$ and $n(B)$, we have counted the elements of $n(A \cap B)$ twice. The principle in fact holds for any finite number of sets. We state it for three sets.

Theorem 10.7.1. For any finite sets A, B, C, we have

$$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C).$$

Example 10.7.2. Find the number of mathematics students at a college taking at least one of the languages French, German, and Russian, given the following data:

65 study French,	20 study French and German,	
45 study German,	25 study French and Russian,	8 study all three languages.
42 study Russian,	15 study German and Russian.	

We want to find $n(F \cup G \cup R)$, where F, G, and R denote the sets of students studying French, German, and Russian, respectively.

By the Inclusion–Exclusion Principle,

$$\begin{aligned} n(F \cup G \cup R) &= n(F) + n(G) + n(R) - n(F \cap G) - n(F \cap R) - n(G \cap R) + n(F \cap G \cap R) \\ &= 65 + 45 + 42 - 20 - 25 - 15 + 8 = 100. \end{aligned}$$

Namely, 100 students study at least one of the three languages.

Now, suppose we have any finite number of finite sets, say $A_1, A_2, \ldots, A_m$. Let s_k be the sum of the cardinalities

$$n(A_{i1} \cap A_{i2} \cap \ldots A_{ik})$$

of all possible k-tuple intersections of the given m sets. Then we have the following general Inclusion–Exclusion Principle.

Theorem 10.7.3. We have

$$n(A_1 \cup A_2 \cup \ldots \cup A_m) = s_1 - s_2 + s_3 - \cdots + (-1)^{m-1} s_m.$$

Exercise 10.7.4.

1. Suppose among 32 people who save paper or bottles (or both) for recycling, there are 30 who save paper and 14 who save bottles. Find the number m of people who: (a) save both; (b) save only paper; (c) save only bottles.

2. Let A, B, C, D denote, respectively, art, biology, chemistry, and drama courses. Find the number N of students in a dormitory given the data:

   2 take A, 5 take A and B, 4 take B and D, 2 take B, C, D,
   20 take B, 7 take A and C, 3 take C and D, 3 take A, C, D,
   20 take C, 4 take A and D, 3 take A, B, C, 2 take all four,
   8 take D, 16 take B and C, 2 take A, B, D, 71 take none.

10.8 Tree Diagrams

Definition 10.8.1. A tree diagram is a device used to enumerate all the possible outcomes of a sequence of events where each event can occur in a finite number of ways.

The construction of tree diagrams is illustrated in the following example

Example 10.8.2. (a) We want to find the product set $A \times B \times C$, where $A = \{1, 2\}$, $B = \{a, b, c\}$ and $C = \{x, y\}$. The tree diagram for $A \times B \times C$ is shown in fig. 10.8.1 (a). Here the tree is constructed from left to right, and the number of branches at each point corresponds to the possible outcomes of the next event. Each endpoint (leaf) of the tree is labelled by the corresponding element of $A \times B \times C$. As noted previously, $A \times B \times C$ has $n = 2 \cdot 3 \cdot 2 = 12$ elements.

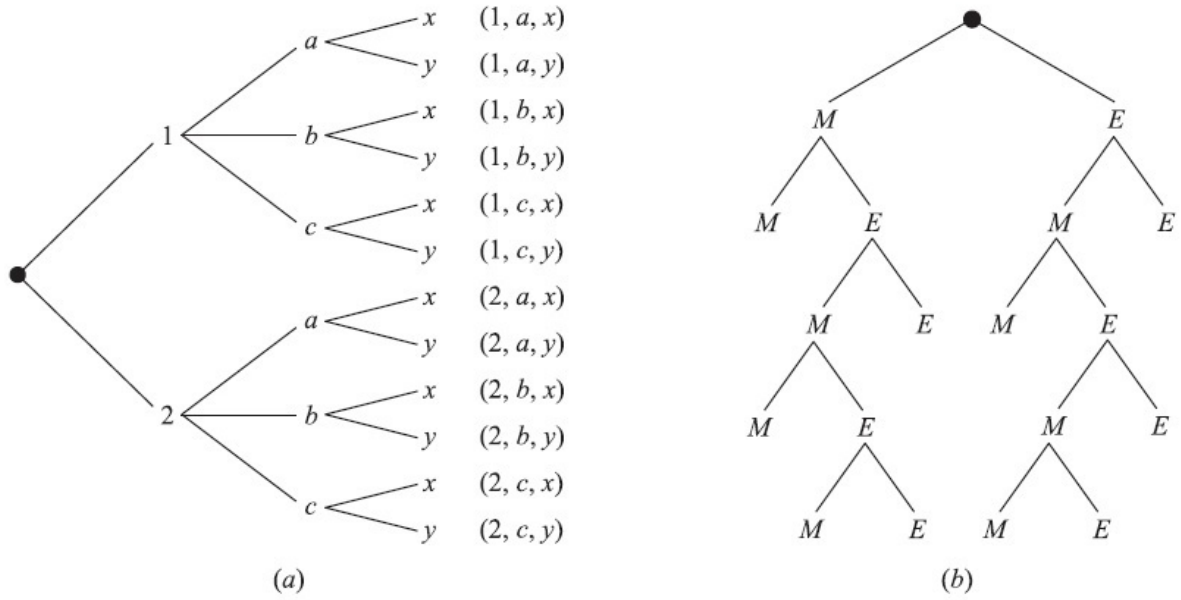


Figure 10.8.1

(b) Mark and Erik are to play a tennis tournament. The first person to win two games in a row or who wins a total of three games wins the tournament. We want to find the number of ways the tournament can occur.

The tree diagram showing the possible outcomes of the tournament appears in fig. 10.8.1 (b). Here the tree is constructed from top-down rather than from left-right. (That is, the "root" is on the top of the tree.) Note that there are 10 endpoints, and the endpoints correspond to the following 10 ways the tournament can occur:

$$MM,\ MEMM,\ MEMEM,\ MEMEE,\ MEE,\ EMM,\ EMEMM,\ EMEME,\ EMEE,\ EE.$$

The path from the beginning (top) of the tree to the endpoint describes who won which game in the tournament.

Exercise 10.8.3. 1. Teams A and B play in a tournament. The first team to win three games wins the tournament. Find the number n of possible ways the tournament can occur. Construct the appropriate tree diagram.

2. Construct the tree diagram that gives the permutations of $\{a, b, c\}$.

10.9 Few Probable Questions

1. State sum rule principle. A store sells clothes for men. It has 3 kinds of jackets, 7 kinds of shirts, and 5 kinds of pants. Find the number of ways a person can buy: (a) one of the items; (b) one of each of the three kinds of clothes.

2. State product rule principle. Suppose a code consists of five characters, two letters followed by three digits. Find the number of: (a) codes; (b) codes with distinct letter; (c) codes with the same letters.

3. Find n if: (a) $P(n, 4) = 42P(n, 2)$; (b) $2P(n, 2) + 50 = P(2n, 2)$.

4. Consider all positive integers with three different digits. (Note that zero cannot be the first digit.) Find the number of them which are: (a) greater than 700; (b) odd; (c) divisible by 5.

5. A class contains l0 students. Find the number n of ordered samples of size 4: (a) with replacement; (b) without replacement.

6. A women student is to answer 10 out of 13 questions. Find the number of her choices where she must answer:

   (a) the first two questions;

   (b) the first or second question but not both;

   (c) exactly 3 out of the first 5 questions;

   (d) at least 3 of the first 5 questions.

7. Consider all integers from 1 up to and including 300. Find the number of them that are divisible by:

   (a) at least one of 3, 5, 7;

   (b) 3 and 5 but not by 7;

   (c) by 5, but by neither 3 nor 7;

   (d) by none of the numbers 3, 5, 7.

8. Find the number m of elements in the union of sets A, B, C, D where:

   (a) A, B, C, D have 50, 60, 70, 80 elements, respectively.

   (b) Each pair of sets has 20 elements in common.

   (c) Each three of the sets has 10 elements in common.

   (d) All four of the sets have 5 elements in common.

9. State pigeonhole principle. Suppose 5 points are chosen at random in the interior of an equilateral triangle T where each side has length two inches. Show that the distance between two of the points must be less than one inch.

Unit 11

Course Structure

- Grammar and Language: Introduction, Alphabet, Words, Free semi group, Languages, Regular expression and regular languages, Grammars.

11.1 Introduction

Words in the English language can be combined in various ways.The grammar of English tells us whether a combination of words in a valid sentence. For example “the cat is reading book" is a valid sentence because it is formed by a noun phrase the cat, followed by a verb phrase: is reading a book. We do not care that it is meaningless. Since we are concerned only with the syntax and not with its semantics i.e. meaning. Research in the automatic translation of one language to another has led to the concept of formal language. It is specified by a well-defined set of rules of syntax. Rules of syntax are important not only in linguistic but also in the study of programming languages.

So, a formal language is a language that is specified by a well-defined set of rules of syntax and a formal grammar is any compact, precise definition of a language. A grammar implies an algorithm that would generate all legal sentences of the language. Let us learn more in this unit.

Objectives

After reading this unit, you will be able to

- define alphabets, words, concatenation of words, subwords
- see that the set of all words form a semigroup with respect to the concatenation of words
- define language, regular expression and regular languages
- define grammar find its relation with languages

11.2 Alphabet, Words, Free Semigroup

Definition 11.2.1. Consider a non-empty set A of symbols. A word or string w on the set A is a finite sequence of its elements.

For example, suppose $A = \{a, b, c\}$. Then the following sequences are words on A:

$$u = ababb, \quad \text{and} \quad v = accbaaa.$$

When discussing words on A, we frequently call A the alphabet, and its elements are called letters. We will also abbreviate our notation and write a^2 for aa, a^3 for aaa, and so on. Thus, for the above words, $u = ab^2$ and $v = ac^2ba^3$.

The empty sequence of letters, denoted by λ, or ϵ, or 1 is also considered to be a word on A, called the **empty word**. The set of all words on A is denoted by A^*.

Definition 11.2.2. The length of a word u, written $|u|$ or $l(u)$, is the number of elements in its sequence of letters.

For the above words u and v, we have $l(u) = 5$ and $l(v) = 7$. Also, $l(\lambda) = 0$.

Unless otherwise stated, the alphabet A will be finite, the symbols u, v, w will be reserved for words on A, and the elements of A will come from the letters a, b, c.

Definition 11.2.3. (**Concatenation**) Consider two words u and v on the alphabet A. The concatenation of u and v, written uv, is the word obtained by writing down the letters of u followed by the letters of v.

For the above words u and v, we have

$$uv = ababbaccbaaa = ab^2ac^2ba^3$$

As with letters, for any word u, we define $u^2 = uu$, $u^3 = uuu$, and in general, $u^{n+1} = uu^n$.

Clearly, for any words u, v, w, the words $(uv)w$ and $u(vw)$ are identical, they simply consist of the letters of u, v, w written down one after the other. Also, adjoining the empty word before or after a word u does not change the word u. That is:

Theorem 11.2.4. The concatenation operation for words on an alphabet A is associative. The empty word λ is an identity element for the operation.

(Generally speaking, the operation is not commutative, e.g., $uv \neq vu$ for the above words u and v.)

Definition 11.2.5. (Subwords, Initial Segments) Consider any word $u = a_1a_2 \dots a_n$ on an alphabet A. Any sequence $w = a_ja_{j+1} \dots a_k$ is called a subword of u. In particular, the subword $w = a_1a_2 \dots a_k$ beginning with the first letter of u, is called an initial segment of u. In other words, w is a subword of u if $u = v_1wv_2$ and w is an initial segment of u if $u = wv$. Observe that λ and u are both subwords or uv since $u = \lambda u$.

Consider the word $u = abca$. The subwords and initial segments of u are as follows:

1. Subwords: $\lambda, a, b, c, ab, bc, ca, abc, bca, abca = u$.
2. Initial segments: $\lambda, a, ab, abc, abca = u$.

Observe that the subword $w = a$ appears in two places in u. The word ac is not a subword of u even though all its letters belong to u.

Definition 11.2.6. Let F denote the set of all non-empty words from an alphabet A with the operation of concatenation. As noted above, the operation is associative. Thus F is a semigroup; it is called the **free semigroup** over A or the free semigroup generated by A.

One can easily show that F satisfies the right and left cancellation laws. However, F is not commutative when A has more than one element. We will write F_A for the free semigroup over A when we want to specify the set A.

Now let $M = A^*$ be the set of all words from A including the empty word λ. Since λ is an identity element for the operation of concatenation, M is a monoid, called the **free monoid** over A.

11.3 Languages

Definition 11.3.1. A **language** L over an alphabet A is a collection of words on A. Recall that A^* denotes the set of all words on A. Thus a language L is simply a subset of A^*.

Example 11.3.2. Let $A = \{a, b\}$. The following are languages on A.

1. $L_1 = \{a, ab, ab^2, \ldots\}$, consisting of all words beginning with an a and followed by zero or more b's.
2. $L_2 = \{b^m a b^n : \ m \geq 0, \ n \geq 0\}$, consisting of all words with exactly one a.

11.3.1 Operations on Languages

Suppose L and M are languages over an alphabet A. Then the "concatenation" of L and M, denoted by LM, is the language defined as follows:

$$LM = \{uv : \ u \in L, \ v \in V\}$$

That is, LM denotes the set of all words which come from the concatenation of a word from L with a word from M. For example, for the languages L_1 and L_2 described in the preceding example, we would have

$$L_1 L_2 = \{ab^m a b^n : \ m \geq 0, \ n \geq 0\}$$

Clearly, the concatenation of languages is associative since the concatenation of words is associative.

Powers of a language L are defined as follows:

$$L^0 = \lambda, \ L^1 = L, \ L^2 = LL, \ \ L^{m+1} = L^m L, \ m > 1.$$

The unary operation L^* of a language L, called the **Kleene closure** of L is defined as

$$L^* = \bigcup_{k=0}^{\infty} L^k.$$

11.4 Regular Expressions and Regular Languages

Let A be a (nonempty) alphabet. This section defines a regular expression r over A and a language $L(r)$ over A associated with the regular expression r. The expression r and its corresponding language $L(r)$ are defined inductively as follows.

Definition 11.4.1. Each of the following is a regular expression over an alphabet A.

1. The symbol λ and the pair "()"(empty expression) are regular expressions;

2. Each letter a in A is a regular expression;
3. If r is a regular expression, then r^* is a regular expression;
4. If r_1 and r_2 are regular expressions, then $(r_1 \vee r_2)$ is a regular expression;
5. If r_1 and r_2 are regular expressions, then $(r_1 r_2)$ is a regular expression.

All regular expressions are formed in this way.

Definition 11.4.2. The Language $L(r)$ over A defined by a regular expression r over A is as follows:

1. $L(\lambda) = \{\lambda\}$ and $L(()) = \emptyset$;
2. $L(a) = \{a\}$, where a is a letter in A;
3. $L(r^*) = (L(r))^*$ (the Kleene closure of $L(r)$);
4. $L(r_1 \vee r_2) = L(r_1) \cup L(r_2)$ (union of the languages);
5. $L(r_1 r_2) = L(r_1)L(r_2)$ (concatenation of the languages).

And finally,

Definition 11.4.3. Let L be a language over A. Then L is called a regular language over A if there exists a regular expression r over A such that $L = L(r)$.

Example 11.4.4. Let $A = \{a, b\}$. Each of the following is an expression r and its corresponding language $L(r)$:

1. Let $r = a^*$. Then $L(r)$ consists of all powers of a including the empty word.
2. Let $r = aa^*$. Then $L(r)$ consists of all positive powers of a excluding the empty word.
3. Let $r = a \vee b^*$. Then $L(r)$ consists of a or any word in b, that is, $L(r) = \{a, \lambda, b, b^2, \ldots\}$.
4. Let $r = (a \vee b)^*$. Then $L(r) = \{a\} \cup \{b\} = A$. Hence, $L(r) = A^*$.

Example 11.4.5. Consider the following languages over $A = \{a, b\}$. Find a regular expression r over A such that $L_i = L(r)$, for $i = 1, 2$.

1. $L_1 = \{a^m b^n : \ m > 0, \ n > 0\}$. L_1 consists of those words beginning with one or more a's followed by one or more b's. Thus we can set $r = aa^*bb^*$. Note that this r is not unique. We could also take $r = a^*abb^*$.
2. $L_2 = \{a^m b^m : \ m > 0\}$. L_2 consists of all words beginning with one or more a's followed by the same number of b's. There exists no regular expression r such that $L_2 = L(r)$; that is, L_2 is not a regular language.

Exercise 11.4.6.
1. Let $u = a^2b$ and $v = b^3ab$. Find (a) uvu; (b)λu, $u\lambda$.
2. State the difference between the free semigroup on an alphabet A and the free monoid on A.
3. Let $A = \{a, b, c\}$. Find where (a) $L = \{b^2\}$; (b) $L = \{a, b, c^3\}$
4. Let $A = \{a, b, c\}$. State whether w belongs to $L(r)$ or not, where (a) $r = a^* \vee (b \vee c)^*$; (b) $r = a^*(b \vee c)^*$.

11.5 Grammars

Fig. 11.5.1 shows the grammatical construction of a specific sentence. Observe that there are:

1. various variables, for example, (sentence), (noun phrase), etc.;
2. various terminal words, example, "The", "boy",etc.;
3. a beginning variable (sentence);
4. various substitutions or productions, for example,

$$\begin{aligned} \langle sentence \rangle &\rightarrow \langle noun\ phrase \rangle \langle verb\ phrase \rangle \\ \langle object\ phrase \rangle &\rightarrow \langle article \rangle \langle noun \rangle \\ \langle noun \rangle &\rightarrow apple \end{aligned}$$

The final sentence only contains terminals, although both variables and terminals appear in its construction by the productions. This intuitive description is given in order to motivate the following definition of a grammar and the language it generates.

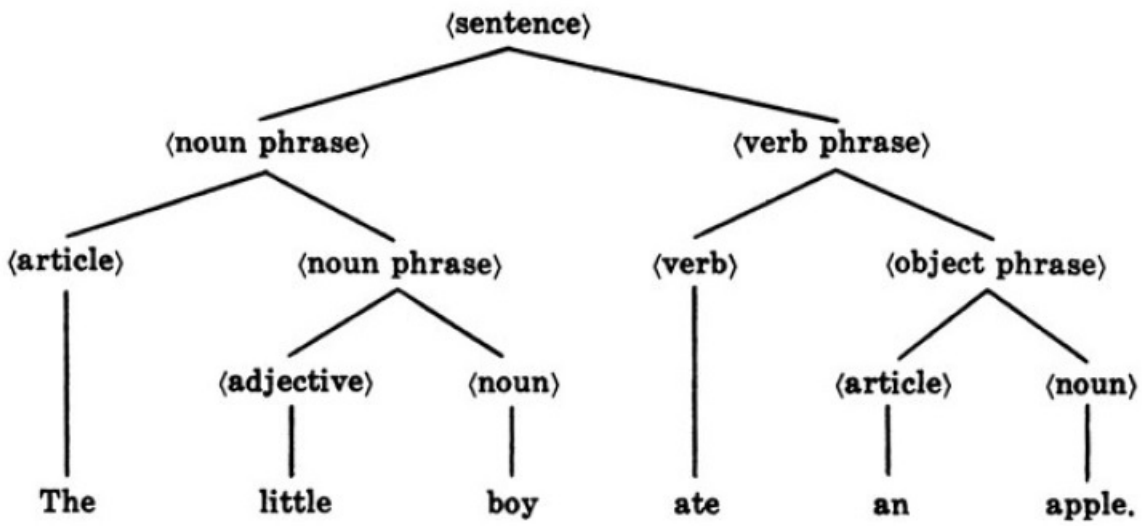


Figure 11.5.1

Definition 11.5.1. A phrase structure grammar or, simply, a grammar G consists of four parts:

1. A finite set (vocabulary) V;
2. A subset T of V whose elements are called **terminals**; the elements of $N = V \setminus T$ are called **non-terminals** or **variables**;
3. A non-terminal symbol S called the **start symbol**;
4. A finite set P of productions. (A production is an ordered pair (α, β), usually written $\alpha \rightarrow \beta$, where α and β are words in V, and the production must contain at least one non-terminal on its left side α.)

Such a grammar G is denoted by $G = G(V, T, S, P)$.

The following notation, unless otherwise stated or implied, will be used for our grammars. Terminals will be denoted by italic lower case Latin letters $a, b, c \ldots$, and non-terminals will be denoted by italic capital Latin letters $A, B, C, \ldots$, with S as the start symbol. Also, Greek letters, $\alpha, \beta, \ldots$, will denote words in V, that is, words in terminals and non-terminals. Furthermore, we will write $\alpha \to (\beta_1, \beta_2, \cdots, \beta_k)$.

11.5.1 Language $L(G)$ of a Grammar G

Suppose w and w' are words over the vocabulary set V of a grammar G. We write

$$w \Rightarrow w'$$

if w' can be obtained from w by using one of the productions; that is, if there exist words u and v such that $w = u\alpha v$ and $w' = u\beta v$ and there is a production $\alpha \to \beta$. Furthermore, we write

$$w \Rightarrow\Rightarrow w'$$

if w' can be obtained from w using a finite number of productions.

Now let G be a grammar with terminal set T . The language of G, denoted by $L(G)$, consists of all words in T that can be obtained from the start symbol S by the above process; that is,

$$L(G) = \{w \in T^* : S \Rightarrow\Rightarrow w\}$$

Example 11.5.2. The following defines a grammar G with S as the start symbol:

$$V = \{A, B, S, a, b\}, \quad T = \{a, b\}, \quad P = \{S \xrightarrow{1} AB, A \xrightarrow{2} Aa, B \xrightarrow{3} Bb, A \xrightarrow{4} a, B \xrightarrow{5} b\}$$

Now, $w = a^2b^4$ can be obtained from the start symbol S as follows:

$$S \Rightarrow AB \Rightarrow AaB \Rightarrow aaB \Rightarrow aaBb \Rightarrow aaBbb \Rightarrow aaBbbb \Rightarrow aabbbb = a^2b^4.$$

Here we used the productions 1, 2, 4, 3, 3, 3, 5, respectively. Thus, we write $S \Rightarrow\Rightarrow a^2b^4$ belongs to $L(G)$. More generally, the production sequence:

$$1,\ 2(r \text{ times}),\ 4,\ 3(s \text{ times}),\ 5$$

will produce the word $w = a^r ab^s b$, where r and s are non-negative integers. On the other hand, no sequence of productions can produce an a after a b. Accordingly,

$$L(G) = \{a^m b^n :\ m > 0,\ n > 0\}$$

That is, the language $L(G)$ of the grammar G consists of all words which begin with one or more a's followed by one or more b's.

Few Probable Questions

1. Define Language. Let $A = \{a, b\}$. Find a regular expression r such that $L(r)$ consists of all words w where:
   (a) w begins with a^2 and ends with b^2;
   (b) w contains an even number of a's.
2. Find the language $L(G)$ generated by the grammar G with variables S, A, B, terminals a, b, and productions $S \to aB, B \to b, B \to bA, A \to aB$.

Unit 12

Course Structure

- Finite Automata (FA).

12.1 Introduction

Automata theory is the study of abstract machines and automata, as well as the computational problems that can be solved using them. It is a theory in theoretical computer science and discrete mathematics (a subject of study in both mathematics and computer science). The word automata (the plural of automaton) comes from a Greek word, which means "self-making".

Automata theory is closely related to formal language theory. An automaton is a finite representation of a formal language that may be an infinite set. Automata are often classified by the class of formal languages they can recognize, typically illustrated by the Chomsky hierarchy, which describes the relations between various languages and kinds of formalized logics.

Automata play a major role in theory of computation, compiler construction, artificial intelligence, parsing and formal verification.

Objectives

After reading this unit, you will be able to

- define finite state automata and related terms and find its relation with languages

12.2 Finite State Automata

Definition 12.2.1. A finite state automaton (FSA) or, simply, an automaton M, consists of five parts:

1. A finite set (alphabet) A of inputs.
2. A finite set S of (internal) states.
3. A subset Y of S (called accepting or "yes" states).
4. An initial state s_0 in S.

5. A next-state function $F : S \times A \to A$.

Such an automaton M is denoted by $M = (A, S, Y, s_0, F)$ (The plural of automaton is automata).

Example 12.2.2. The following defines an automaton M with two input symbols and three states:

1. $A = \{a, b\}$, input symbols.
2. $S = \{s_0, s_1, s_2\}$, internal states.
3. $Y = \{s_0, s_1\}$, "yes" states.
4. s_0, initial state,
5. Next state function $F : S \times A \to S$ defined explicitly as follows:

$$F(s_0, a) = s_0,\ F(s_1, a) = s_0,\ F(s_2, a) = s_2$$
$$F(s_0, b) = s_1,\ F(s_1, b) = s_2,\ F(s_2, b) = s_2.$$

12.2.1 State Diagram of an Automaton M

An automaton M is usually defined by means of its state diagram $D = D(M)$ rather than by listing its five parts. The state diagram $D = D(M)$ is a labelled directed graph as follows.

1. The vertices of $D(M)$ are the states in S and an accepting state is denoted by means of a double circle.
2. There is an arrow (directed edge) in $D(M)$ from state s_j to state s_k labelled by an input a if $F(s_j, a) = s_k$.
3. The initial state s_0 is indicated by means of a special arrow which terminates at s_0 but has no initial vertex.

For each vertex s_j and each letter a in the alphabet A, there will be an arrow leaving s_j, which is labelled by a; hence the outdegree of each vertex is equal to number of elements in A. For notational convenience, we label a single arrow by all the inputs which cause the same change of state rather than having an arrow for each such input.

The state diagram $D = D(M)$ of the automaton M in the preceding Example is shown in fig. 12.2.1.

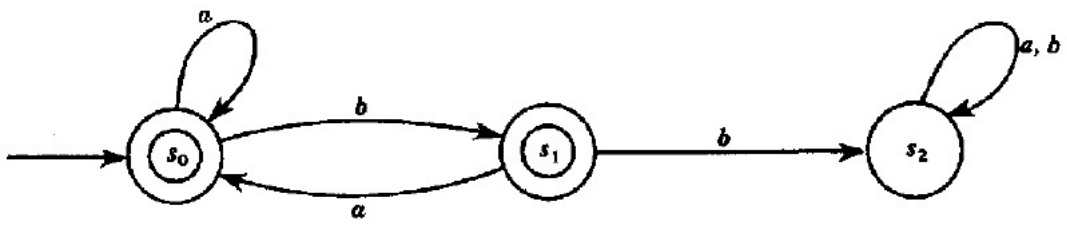


Figure 12.2.1

Language $L(M)$ Determined by an Automaton M

Each automaton M with input alphabet A defines a language over A, denoted by $L(M)$, as follows.

Let $w = a_1a_2\cdots a_m$ be a word on A. Then w determines the following path in the state diagram graph $D(M)$ where s_0 is the initial state and $F(s_{i-1}, a_i) = s_i$ for $i \geq 1$

$$P = (s_0, a_1, s_1, a_2, s_2, \ldots, a_m, s_m).$$

We say that M recognizes the word w if the final state s_m is an accepting state in Y. The language $L(M)$ of M is the collection of all words from A which are accepted by M.

Example 12.2.3. We determine whether or not the automaton M in fig. 12.2.1accepts the words

$$w_1 = ababba, \quad w_2 = baab, \quad w_3 = \lambda.$$

Using fig. 12.2.1 and the words w_1 and w_2, we obtain the respective paths:

$$P_1 = s_0 \xrightarrow{a} s_0 \xrightarrow{b} s_1 \xrightarrow{a} s_0 \xrightarrow{b} s_1 \xrightarrow{b} s_2 \xrightarrow{a} s_2, \quad \text{and} \quad P_2 = s_0 \xrightarrow{b} s_1 \xrightarrow{a} s_0 \xrightarrow{a} s_0 \xrightarrow{b} s_1$$

The final state in P_1 is s_2 which is not in Y. Hence w_1 is not accepted by M. Also, the final state of P_2 is s_1 which is in Y so w_2 is accepted by M. The final state determined by w_3 is the initial state s_0 which is in Y. Thus w_3 is also accepted by M.

We also describe the language $L(M)$. $L(M)$ will consist of all words w on A which do not have two successive b's. This comes from the following facts:

1. We can enter the state s_2 if and only if there are two successive b's.
2. We can never leave s_2.
3. The state s_2 is the only rejecting (non-accepting) state.

Example 12.2.4. Consider the automaton M in fig. 12.2.2. We want to find the words w in language L that are accepted by M. The system can reach the accepting state s_2 only when there exists an a in w which follows a b.

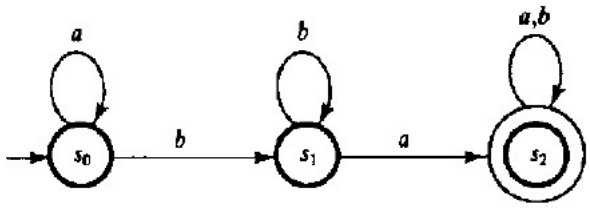


Figure 12.2.2

The fundamental relationship between regular languages and automata is contained in the following theorem.

Theorem 12.2.5. (Kleene): A language L over an alphabet A is regular if and only if there is a finite state automaton M such that $L = L(M)$.

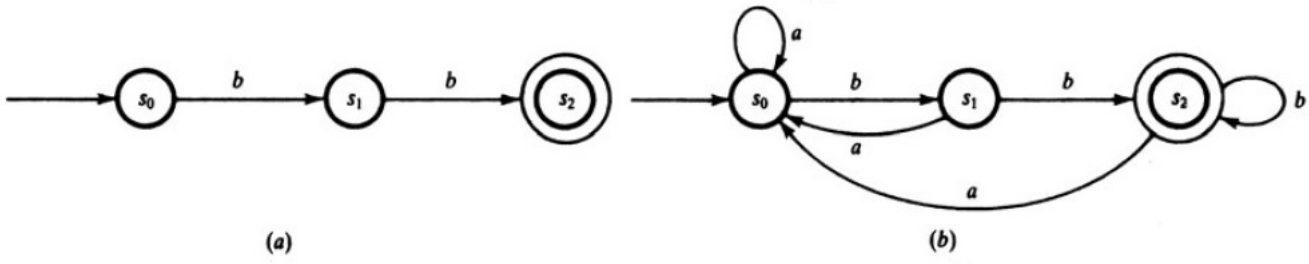


Figure 12.2.3

Example 12.2.6. Let $A = \{a, b\}$. We construct an automaton M which will accept precisely those words from A which end in two b's. Since b^2 is accepted, but not λ or b, we need three states, s_0, the initial state, and s_1 and s_2 with an arrow labelled b going from s_0 to s_1 and one from s_1 to s_2. Also, s_2 is an accepting state, but not s_0 nor s_1. This gives the graph in fig. 12.2.3(a). On the other hand, if there is an a, then we want to go back to s_0, and if we are in s_2 and there is a b, then we want to stay in s_2. These additional conditions give the required automaton M which is shown in fig. 12.2.3(b).

Example 12.2.7. Let $A = \{a, b\}$. We construct an automaton M which will accept those words from A which begin with an a followed by (zero or more) b's in fig. 12.2.4.

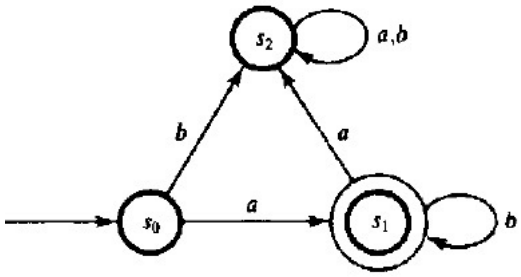


Figure 12.2.4

Pumping Lemma

Let M be an automaton over A with k states. Suppose $w = a_1 a_2 \cdots a_n$ is a word over A accepted by M and suppose $|w| = n > k$, the number of states. Let $P = (s_0, s_1, \ldots, s_n)$ be the corresponding sequence of states determined by the word w. Since $n > k$, two of the states in P must be equal, say $s_i = s_j$ where $i < j$. Let w be divided into subwords x, y, z as follows:

$$x = a_1 a_2 \cdots a_i, \quad y = a_{i+1} \cdots a_j, \quad z = a_{j+1} \cdots a_n.$$

As shown in fig. 12.2.5, xy ends in $s_i = s_j$; hence xy^m also ends in s_i. Thus, for every m, $w_m = xy^m z$ ends in s_n, which is an accepting state.

The above discussion proves the following important result.

Theorem 12.2.8. (Pumping Lemma): Suppose M is an automaton over A such that:

1. M has k states.

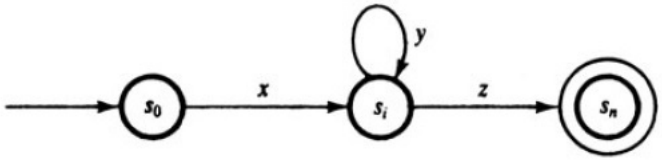


Figure 12.2.5

2. M accepts a word w from A where $|w| > k$.

Then $w = xyz$ where, for every positive m, $w_m = xy^m z$ is accepted by M.

The next example gives an application of the Pumping Lemma.

Example 12.2.9. We want to show that the language $L = \{a^m b^m : m > 0\}$ is not regular.

Suppose L is regular. Then by theorem 12.2.5, there exists a finite state automaton M which accepts L. Suppose M has k states. Let $w = a^k b^k$. Then $|w| > k$. By theorem 12.2.8, $w = xyz$ where y is not empty and $w_2 = xy^2z$ is also accepted by M. If y consists of only a's or only b's, then w_2 will not have the same number of a's as b's. If y contains both a's and b's, then w_2 will have a's following b's. In either case, w_2 does not belong to L, which is a contradiction. Hence L is not regular.

Few Probable Questions

1. Define finite state automaton. Let M be the automaton with the following input set A, state set S with initial state s_0 and accepting set Y:

   $$A = \{a, b\}, \quad S = \{s_0, s_1, s_2\}, \quad Y = \{s_2\}.$$

   Also, the next-state function is given by

   $$F(s_0, a) = s_0, \ F(s_1, a) = s_1, \ F(s_2, a) = s_2$$
   $$F(s_0, b) = s_1, \ F(s_1, b) = s_2, \ F(s_2, b) = s_2.$$

   Draw the State diagram $D(M)$ of M. Also, describe the language $L = L(M)$ accepted by M.

2. Let $A = \{a, b\}$. Construct an automaton M which will accept precisely those words from A which have an even number of a's.

Unit 13

Course Structure

- Finite State Machine. Non-deterministic and deterministic FA. Push Down Automation (PDA), Equivalence of PDAs and Context Free Languages (CFLs).

13.1 Introduction

This unit discusses two types of "machines." The first is a finite state machine (FSM) which is similar to a finite state automaton (FSA) except that the finite state machine “prints" an output using an output alphabet which may be distinct from the input alphabet.

Objectives

After reading this unit, you will be able to

- define finite state machines and draw their state tables and diagrams

13.2 Finite State Machines

Definition 13.2.1. A finite state machine (or complete sequential machine) M consists of six parts:

1. A finite set A of input symbols;
2. A finite set S of "internal" states;
3. A finite set Z of output symbols;
4. An initial state s_0 in S;
5. A next-state function $f : S \times A \to S$;
6. An output function g from $S \times A$ into Z.

Such a machine M is denoted by $M = M(A, S, Z, s_0, f, g)$.

Example 13.2.2. The following defines a finite state machine M with two input symbols, three internal states, and three output symbols:

$$A = \{a, b\}, \quad S = \{s_0, s_1, s_2\}, \quad Z = \{x, y, z\}, \quad \text{Initial state } s_0,$$

the next-state function $f : S \times A \to S$ defined by

$$f(s_0, a) = s_1, \quad f(s_1, a) = s_2, \quad f(s_2, a) = s_0$$
$$f(s_0, b) = s_2, \quad f(s_1, b) = s_1, \quad f(s_2, b) = s_1.$$

and the output function $g : S \times A \to Z$ defined by

$$g(s_0, a) = x, \quad g(s_1, a) = x, \quad g(s_2, a) = z$$
$$g(s_0, b) = y, \quad g(s_1, b) = z, \quad g(s_2, b) = y.$$

13.2.1 State Table and State Diagram of a Finite State Machine

There are two ways of representing a finite state machine M in compact form. One way is by a table called the state table of the machine M, and the other way is by a labelled directed graph called the state diagram of the machine M.

The state table combines the next-state function f and the output function g into a single table which represent the function $F : S \times A \to S \times Z$ defined as follows

$$F(s_j, a_j) = [f(s_i, a_j), g(s_i, a_j)]$$

For instance, the state table of the machine M in the preceding example is given in table 13.1. The states are

F	a	b
s_0	s_1, x	s_2, y
s_1	s_2, x	s_1, z
s_2	s_0, z	s_1, y

Table 13.1

listed on the left of the table with the initial state first, and the input symbols are listed on the top of the table. The entry in the table is a pair (s_k, z_r) where $s_k = f(s_i, a_j)$ is the next state and $z_r = g(s_i, a_j)$ is the output symbol. The corresponding state diagram is given in fig. 13.2.1.

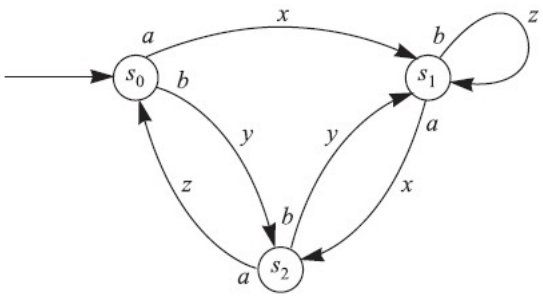


Figure 13.2.1

The state diagram $D = D(M)$ of a finite state machine M is a labelled digraph the vertices of which are the states of M. Moreover, if

$$F(s_i, a_j) = (s_k, z_r) \quad \text{or equivalently,} \quad s_k = f(s_i, a_j), \;\; z_r = g(s_i, a_j)$$

then there is an arc (arrow) from s_i to s_k which is labelled with the pair a_j, z_r. We usually put the input symbol a_i near the base of the arrow (near s_i) and the output symbol z_r near the center of the arrow. We also label the initial state s_0 by drawing an extra arrow into s_0. See fig. 13.2.1.

Input and Output Tapes

The above discussion of a finite state machine M does not show the dynamic quality of M. Suppose M is given a string (word) of input symbols, say

$$u = a_1 a_2 \cdots a_m$$

We visualize these symbols on an "input tape." The machine M "reads" these input symbols one by one and, simultaneously, changes through a sequence of states

$$v = s_0 s_1 s_2 \cdots s_m$$

where s_0 is the initial state, while printing a string(word) of output symbols

$$w = z_1 z_2 \cdots z_m$$

on an "output tape." Formally, the initial state s_0 and the input string u determine the strings v and w as follows, where $i = 1, 2, \ldots, m$:

$$s_i = f(s_{i-1} a_i), \quad \text{and} \quad z_i = g(s_{i-1}, a_i).$$

Example 13.2.3. Consider the machine M of fig. 13.2.1. Suppose the input is the word $u = abaab$. We calculate the sequence v of states and the output word w from the state diagram as follows. Beginning at the initial state s_0, we follow the arrows which are labelled by the input symbols as follows:

$$s_0 \xrightarrow{a,x} s_1 \xrightarrow{b,z} s_1 \xrightarrow{a,x} s_2 \xrightarrow{a,z} s_0 \xrightarrow{b,y} s_2$$

This yields the following sequence v of states and output word w:

$$v = s_0 s_1 s_1 s_2 s_0 s_2 \quad \text{and} \quad w = xzxy.$$

Binary Addition

This subsection describes a finite state machine M which can do binary addition. By adding 0's at the beginning of our numbers, we can assume that our numbers have the same number of digits. If the machine is given the input $1101011 + 0111011$ then we want the output to be the binary sum 10100110. Specifically, the input is the string of pairs of digits to be added:

$$11, \;\; 11, \;\; 00, \;\; 11, \;\; 01, \;\; 11, \;\; 10, \;\; b$$

where b denotes blank spaces, and the output should be the string:

$$0, \;\; 1, \;\; 1, \;\; 0, \;\; 0, \;\; 1, \;\; 0, \;\; 1$$

We also want the machine to enter a state called "stop" when the machine finishes the addition.

The input symbols and output symbols are, respectively, as follows:

$$A = \{00, 01, 10, 11, b\} \quad \text{and} \quad Z = \{0, 1, b\}.$$

The machine M that we "construct" will have three states:

$$S = \{\text{carry}(c), \text{no carry}(n), \text{stop}(s)\}$$

Here n is the initial state. The machine is shown in fig. 13.2.2. In order to show the limitations of our

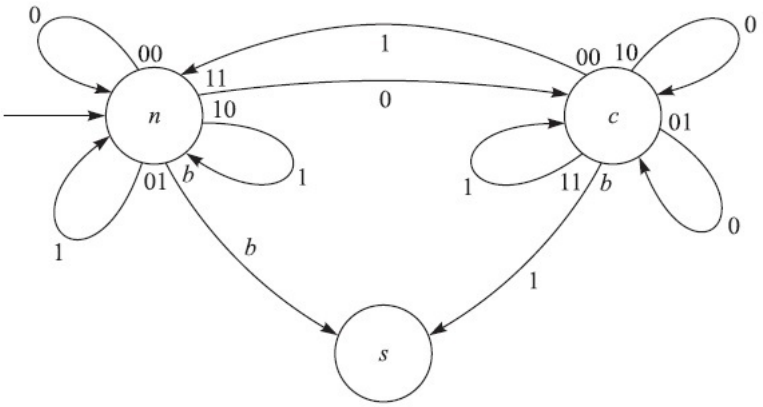


Figure 13.2.2

machines, we state the following theorem.

Theorem 13.2.4. There is no finite state machine M which can do binary multiplication.

If we limit the size of the numbers that we multiply, then such machines do exist. Computers are important examples of finite state machines which multiply numbers, but the numbers are limited as to their size.

Few Probable Questions

1. Define finite state machine. Let M be a FSM with state table 13.2.

F	a	b
s_0	s_1, x	s_2, y
s_1	s_3, y	s_1, z
s_2	s_1, z	s_0, x
s_3	s_0, z	s_2, z

Table 13.2

   (a) Find the input set A, the state set S, the output set Z, and the initial state.
   (b) Draw the state diagram $D = D(M)$ of M
   (c) Suppose $w = aababaabbab$ is an input word (string). Find the corresponding output word v.

Unit 14

Course Structure

- Computable Functions.

14.1 Introduction

Computable functions are the formalized analogue of the intuitive notion of algorithms, in the sense that a function is computable if there exists an algorithm that can do the job of the function, i.e. given an input of the function domain it can return the corresponding output. Computable functions are used to discuss computability without referring to any concrete model of computation such as Turing machines or register machines. Any definition, however, must make reference to some specific model of computation but all valid definitions yield the same class of functions. Particular models of computability that give rise to the set of computable functions are the Turing-computable functions and the general recursive functions.

Objectives

After reading this unit, you will be able to

- learn about Turing machines and how to work with them
- define computable functions and solve related problems

14.2 Turing Machines

There are a number of equivalent ways to formally define a "computable" function. We do it by means of a Turing machine M. This section formally defines a Turing machine M, and the next section defines a computable function.

Our definition of a Turing machine uses an infinite two-way tape, quintuples, and three halt states. Other definitions use a one-way infinite tape and/or quadruples, and one halt state. However, all the definitions are equivalent.

A Turing machine M involves three disjoint non-empty sets:

1. A finite tape set where $B = a_0$ is the blank symbol:

$$A = \{a_1, a_2, \ldots, a_m\} \cup \{B\}$$

2. A finite state set where s_0 is the initial state:

$$S = \{s_1, s_2, \ldots, s_n\} \cup \{s_H, s_Y, s_N\}$$

where s_H(HALT) is the halting state, s_Y(YES) is the accepting state, and s_N(NO) is the non-accepting state.

3. A direction set where L denotes "left" and R denotes "right:"

$$d = \{L, R\}$$

Definition 14.2.1. An expression is a finite (possibly empty) sequence of elements from $A \cup S \cup d$. In other words, an expression is a word whose letters (symbols) come from the sets A, S, and d.

Definition 14.2.2. A tape expression is an expression using only elements from the tape set A.

The Turing machine M may be viewed as a read/write tape head which moves back and forth along an infinite tape. The tape is divided lengthwise into squares (cells), and each square may be blank or hold one tape symbol. At each step in time, the Turing machine M is in a certain internal state s_i scanning one of the tape symbols a_j on the tape. We assume that only a finite number of non-blank symbols appear on the tape.

Fig. 14.2.1(a) is a picture of a Turing machine M in state s_2 scanning the second symbol where $a_1a_3Ba_1a_1$ is printed on the tape. (Note again that B is the blank symbol.) This picture may be represented by the expression $\alpha = a_1s_2a_3Ba_1a_1$ where we write the state s_2 of M before the tape symbol a_3 that M is scanning. Observe that α is an expression using only the tape alphabet A except for the state symbol s_2 which is not at the end of the expression since it appears before the tape symbol a_3 that M is scanning. Fig. 14.2.1 shows two other informal pictures and their corresponding picture expressions.

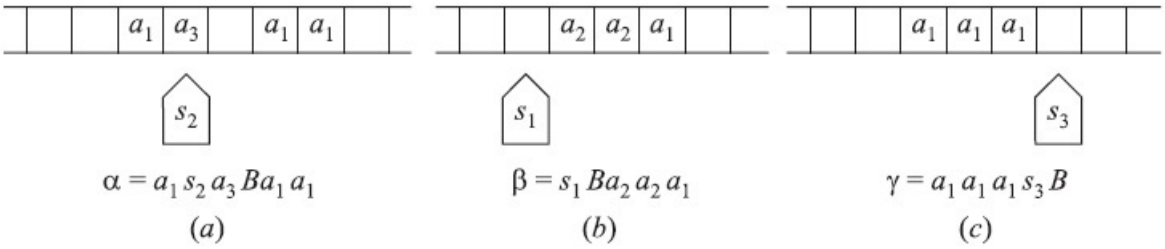


Figure 14.2.1

Definition 14.2.3. A picture α is an expression as follows where P and Q are tape expressions (possibly empty):

$$\alpha = Ps_ia_kQ$$

Definition 14.2.4. Let $\alpha = Ps_ia_kQ$ be a picture. We say that the Turing machine M is in state s_i scanning the letter a_k and that the expression on the tape is the expression Pa_kQ, that is, without its state symbol s_i.

As mentioned above, at each step in time the Turing machine M is in a certain state s_i and is scanning a tape symbol a_k. The Turing machine M is able to do the following three things simultaneously:

1. M erases the scanned symbol a_k and writes in its place a tape symbol a_l(where we permit $a_l = a_k$);
2. M changes its internal states s_i to a state s_j (where we permit $s_j = s_j$).
3. M moves one square to the left or moves one square to the right.

The above action by M may be described by a five-letter expression called a quintuple which we define below.

Definition 14.2.5. A quintuple q is a five-letter expression of the following form:

$$q = \left(s_i, a_k, a_l, s_j, \left\{ \begin{matrix} L \\ R \end{matrix} \right\} \right)$$

That is, the first letter of q is a state symbol, the second is a tape symbol, the third is a tape symbol, the fourth is a state symbol, and the last is a direction symbol L or R.

Next we give a formal definition of a Turing machine.

Definition 14.2.6. A Turing machine M is a finite set of quintuples such that:

1. No two quintuples begin with the same first two letters.
2. No quintuple begins with s_H, s_Y, or s_N.

First condition guarantees that the machine M cannot do more than one thing at any given step, and second condition guarantees that M halts in state s_H, s_Y, or s_N.

The following is an alternative equivalent definition.

Definition 14.2.7. Turing machine M is a partial function from

$$S \setminus \{s_H, s_Y, s_N\} \times A \text{ into } A \times S \times d$$

The term partial function simply means that domain of M is a subset of $S \setminus \{s_H, s_Y, s_N\} \times A$.

The action of the Turing machine described above can now be formally defined.

Definition 14.2.8. Let α and β be pictures. We write $\alpha \to \beta$ if one of the following holds where a, b, c are tape letters and P and Q are tape expressions (possibly empty):

1. $\alpha = Ps_iacQ$, $\beta = Pbs_jcQ$ and M contains the quintuple $q = s_iabs_jR$;
2. $\alpha = Pcs_iaQ$, $\beta = Ps_jcbQ$ and M contains the quintuple $q = s_iabs_jL$;
3. $\alpha = Ps_ia$, $\beta = Pbs_jB$ and M contains the quintuple $q = s_iabs_jR$;
4. $\alpha = s_iaQ$, $\beta = s_jBbQ$ and M contains the quintuple $q = s_iabs_jL$.

Observe that, in all four cases, M replaces a on the tape by b (where we permit $b = a$), and M changes its state from s_i to s_j (where we permit $s_j = s_i$). Furthermore:

1. Here M moves to the right.
2. Here M moves to the left.
3. Here M moves to the right; however, since M is scanning the rightmost letter, it must add the blank symbol B on the right.
4. Here M moves to the left; however, since M is scanning the leftmost letter, it must add the blank symbol B on the left.

Definition 14.2.9. A picture α is said to be terminal if there is no picture β such that $\alpha \to \beta$.

14.2.1 Computing with a Turing Machine

Definition 14.2.10. A computation of a Turing machine M is a sequence of pictures $\alpha_1, \alpha_2, \ldots, \alpha_m$ such that $\alpha_{i-1} \to \alpha_i$ for $i = 1, 2, \ldots, m$ and α_m is a terminal picture.

Turing Machines with Input

Definition 14.2.11. An input for a Turing machine M is a tape expression W. The initial picture for an input W is $\alpha(W)$, where $\alpha(W) = s_0(W)$.

Observe that the initial picture $\alpha(W)$ of the input W is obtained by placing the initial state s_0 in front of the input tape expression W. In other words, the Turing machine M begins in its initial state s_0 and it is scanning the first letter of W.

Definition 14.2.12. Let M be a Turing machine and let W be an input. We say M halts on W if there is a computation beginning with the initial picture $\alpha(W)$.

That is, given an input W, we can form the initial picture $\alpha(W) = s_0(W)$ and apply M to obtain the sequence

$$\alpha(W) \to \alpha_1 \to \alpha_2 \to \ldots$$

Two things can happen:

1. M halts on W. That is, the sequence ends with some terminal Picture α_r.
2. M does not halt on W. That is, the sequence never ends.

Grammars and Turing Machines

Turing machines may be used to recognize languages. Specifically, suppose M is a Turing machine with tape set A. Let L be the set of words W in A such that M halts in the accepting state s_Y when W is the input. We will then write $L = L(M)$, and we will say that M recognizes the language L. Thus an input W does not belong to $L(M)$ if M does not halt on W or if M halts on W but not in the accepting state s_Y.

Theorem 14.2.13. A language L is recognizable by a Turing machine M if and only if L is a type 0 language.

14.3 Computable Functions

Computable functions are defined on the set of non-negative integers. We denote the set of non-negative integers by N_0. Throughout this section, the terms number, integer, and nonnegative integer are used synonymously. The preceding section described the way a Turing machine M manipulates and recognizes character data. Here we show how M manipulates numerical data. First, however, we need to be able to represent our numbers by our tape set A. We will write 1 for the tape symbol a_1 and 1^n for $111\ldots1$, where 1 occur n times.

Definition 14.3.1. Each number n will be represented by the tape expression $\langle n \rangle$ where $\langle n \rangle = 1^{n+1}$.

Thus, $\langle 4 \rangle = 11111 = 1^5$, $\langle 0 \rangle = 1$.

Definition 14.3.2. Let E be an expression. Then $[E]$ will denote the number of times 1 occurs in E.

Then $[11Bs_2a_3111ba_4] = 5$.

Definition 14.3.3. A function $f : N_0 \to N_0$ is computable if there exists a Turing machine M such that, for every integer n, M halts on $\langle n \rangle$ and

$$f(n) = [\text{term}(\alpha(\langle n \rangle)].$$

We then say that M computes f.

That is, given a function f and an integer n, we input $\langle n \rangle$ and apply M. If M always halts on $\langle n \rangle$ and the number of 1's in the final picture is equal to $f(n)$, then f is a computable function and we say that M computes f.

Example 14.3.4. The function $f(n) = n + 3$ is computable. The input is $W = 1^{n+1}$. Thus we need only add two 1's to the input. A Turing machine M which computes f follows:

$$M = \{q_1, q_2, q_3\} = \{s_0 1 s_0 L, s_0 B 1 s_1 L, s_1 B 1 s_H L\}.$$

Observe that:

1. q_1 moves the machine M to the left.
2. q_2 writes 1 in the blank square B, and moves M to the left.
3. q_3 writes 1 in the blank square B, and halts M.

Accordingly, for any positive integer n,

$$s_0 1^{n+1} \to s_0 B 1^{n+1} \to s_1 B 1^{n+2} \to s_H B 1^{n+3}$$

Thus M computes $f(n) = n + 3$. It is clear that, for any positive integer k, the function $f(n) = n + k$ is computable.

Theorem 14.3.5. Suppose $f : N_0 \to N_0$ and $g : N_0 \to N_0$ are computable. Then the composition function $h = g \circ f$ is computable.

14.3.1 Functions of Several Variables

This subsection defines a computable function $f(n_1, n_2, \ldots, n_k)$ of k variables. First we need to represent the list $m = (n_1, n_2, \ldots, n_k)$ in our alphabet A.

Definition 14.3.6. Each list $m = (n_1, n_2, \ldots, n_k)$ of k integers is represented by the tape expression

$$\langle m \rangle = \langle n_1 \rangle B \langle n_2 \rangle B \cdots B \langle n_k \rangle$$

For example, $\langle (2, 0, 4) \rangle = 111B1B11111 = 1^3 B 1^1 B 1^5$.

Definition 14.3.7. The function $f(n_1, n_2, \ldots, n_k)$ of k variables is computable if there is a Turing machine M such that, for every list $m = (n_1, n_2, \ldots, n_k)$, M halts on $\langle m \rangle$ and

$$f(m) = [\text{term}(\alpha(\langle m \rangle))]$$

We then say that M computes f.

Example 14.3.8. The addition function $f(m, n) = m + n$ is computable. The input is $W = 1^{m+1} B 1^{n+1}$. Thus we need only erase two of the 1's. A Turing machine M which computes f follows:

$$M = \{q_1, q_2, q_3, q_4\} = \{s_0 1 B s_1 R, s_1 1 B s_H R, s_1 B B s_2 R, s_2 1 B s_H R\}$$

Observe that:

1. q_1 erases the first 1 and moves M to the right.
2. If $m \neq 0$, then q_2 erases the second 1 and halts M.
3. If $m = 0$, q_3 moves M to the right past the blank square B.
4. q_4 erases the 1 and halts M.

Accordingly, if $m \neq 0$, we have,

$$s_0 1^{m+1} B 1^{n+1} \to s_1 1^m B 1^{n+1} \to s_H 1^{m-1} B 1^{n+1}$$

but if $m = 0$ and $m + n = n$, we have

$$s_0 1 B 1^{n+1} \to s_1 B 1^{n+1} \to s_2 1^{n+1} \to s_H 1^n$$

Thus, M computes $f(m, n) = m + n$.

Few Probable Questions

1. Define Turing machine. Suppose $\alpha = aas_2ba$ is a picture. Find β such that $\alpha \to \beta$ if the Turing machine M has the quintuple q where: (a) $q = s_2bas_1L$; (b) $q = s_2bbs_3R$.
2. Define computable functions. Show that the function f is computable where:
   (a) $f(n) = n - 1$, when $n > 0$, and $f(0) = 0$.
   (b) $f(x, y) = y$.

Unit 15

Course Structure

- Fields and σ-fields of events. Probability as a measure. Random variables. Probability distribution.

15.1 Introduction

The theory of probability had its origin in gambling and games of chance. It owes much to the curiosity of gamblers who pestered their friends in the mathematical world with all sorts of questions. A random (or statistical) experiment is an experiment in which

- All outcomes of the experiment are known in advance.
- Any performance of the experiment results in an outcome that is not known in advance.
- The experiment can be repeated under identical conditions.

In probability theory we study this uncertainty of a random experiment. It is convenient to associate with each such experiment a set Ω, the set of all possible outcomes of the experiment. To engage in any meaningful discussion about the experiment, we associate with Ω a σ-field S of subsets of Ω. We recall that a σ-field is a non-empty class of subsets of Ω that is closed under the formation of countable unions and complements and contains the null set ϕ.

The sample space of a statistical experiment is a pair (Ω, S), where

- Ω is the set of all possible outcomes of the experiment.
- S is a σ-field of subsets of Ω.

The elements of Ω are called *sample points.* Any set $A \in S$ is known as an *event*. Clearly, A is a collection of sample points. We say that an event A happens if the outcome of the experiment corresponds to a point in A. Each one point set is known as a *simple* or *elementary event*. If the set Ω contains only a finite number of points, we say that (Ω, S) is a *finite sample space*. If Ω contains at most a countable number of points, we call (Ω, S) a *discrete sample space*. If, however, Ω contains uncountably many points, we say that (Ω, S) is an *uncountable sample space*. In particular, if $\Omega = R_k$ or some rectangle in R_k, we call it a *continuous sample*

space.

Let us toss a coin. The set Ω is the set of symbols H and T, where H denotes head and T represents tail. Also, S is the class of all subsets of Ω, namely $\{\{H\}, \{T\}, \{H, T\}, \phi\}$. If the coin is tossed two times, then

$$\Omega = \{(H, H),\ (H, T),\ (T, H),\ (T, T)\},$$

and

$$\begin{aligned} S = \Big\{ & \phi,\ \{(H, H)\},\ \{(H, T)\},\ \{(T, H)\},\ \{(T, T)\},\ \{(H, H)\} \\ & \{(H, H), (H, T)\},\ \{(H, H), (T, H)\},\ \{(H, H), (T, T)\},\ \{(H, T), (T, H)\}, \\ & \{(T, T), (T, H)\},\ \{(T, T), (H, T)\},\ \{(H, H), (H, T), (T, H)\}, \\ & \{(H, H), (H, T), (T, T)\},\ \{(H, H), (T, H), (T, T)\},\ \{(H, T), (T, H), (T, T)\}, \Omega \Big\} \end{aligned} \quad (15.1.1)$$

where the first element of a pair denotes the outcome of the first toss, and the second element, the outcome of the second toss. The event *at least one head* consists of sample points $(H, H), (H, T), (T, H)$. The event *at most one head* is the collection of sample points $(H, T), (T, H), (T, T)$.

15.2 Random Variables

Suppose that to each point of a sample space we assign a number. We then have a function defined on the sample space. This function is called a random variable (or stochastic variable) or more precisely a random function (stochastic function). It is usually denoted by a capital letter such as X or Y. In general, a random variable has some specified physical, geometrical, or other significance.

Assume Ω is the sample space of an experiment. A function that takes real values, $X : \Omega \to \mathbb{R}$, is called a random variable if for any interval $I \subset \mathbb{R}$, the set $\{\omega \in \Omega : X(\omega) \in I\}$ is an event of the sample space Ω. The probability that X takes values on I will be simply denoted by $P(X \in I)$.

Example 15.2.1. Suppose that a coin is tossed twice so that the sample space is S = {HH, HT, TH, TT}. Let X represent the number of heads that can come up. With each sample point we can associate a number for X as shown in Table 19.4.1. Thus, for example, in the case of HH (i.e., 2 heads), X = 2 while for TH (1 head), X = 1. It follows that X is a random variable.

Sample Point	*HH*	*HT*	*TH*	*TT*
X	2	1	1	0

Figure 15.2.1

A random variable that takes on a finite or countably infinite number of values is called a discrete random variable while one which takes on a non-countably infinite number of values is called a non-discrete random variable.

If X is a random variable wich can take a finite number or countably infinite number of values X is called discrete random variable. When random variable is discrete, the possible values of X may be assumed as x_1, x_2,..., x_n, ...

If X is a random variable which can take all values in an interval, then X is called a continuous random variable.

Example 15.2.2. The time it takes to complete an exam for a 60 min test. Possible values = all real numbers on the interval $[0, 60]$

15.3 Probability Distribution

While dealing with random variables and their probabilities it is often found that there exists a functional relationship between the value taken by the random variable and the corresponding probability. This initiates to express the relation between random variables and their probabilities whit the help of mathematical functions. These functions are called as probability distributions. Depending on the nature of the random variable distributions can be either discrete or continuous. If the random variable X takes discrete values only, then its probability distribution is called a discrete probability distribution or probability mass function (pmf). However if the random variable X, is such that it can take any value within a given interval them the corresponding distribution is called as continuous probability distribution or probability density function (pdf). Binomial distribution, Poisson distribution, geometric distribution and negative binomial distribution are some examples of discrete random variable. Examples of continuous distribution are normal distribution, beta distribution, gamma distribution etc.

For any discrete variable X, a real-valued function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = P(X = x), \quad x \in \mathbb{R} \tag{15.3.1}$$

is called the probability function of X (sometimes also called the probability mass function associated with X).

or,

If X is a discrete random variable which can take the values $x_1, x_2, x_3, \ldots$ such that $P(X = x_i) = p_i$, then p_i is called the probability function or probability mass function or point probability function, provided p_i $(i = 1, 2, 3, \ldots)$ satisfies the following conditions:

1. $p_i \geq 0$ for all i, and
2. $\sum_i p_i = 1$.

The collection of pairs (x_i, p_i), $i = 1, 2, 3, \ldots$, is called the probability distribution of the random variable X, which is sometimes displayed in the form of a table as given below:

$X = x_i$	$P(X = x_i)$
x_1	p_1
x_2	p_2
...	...
x_r	p_r
...	...

15.4 Discrete Probability Distribution

Let X be a discrete random variable, and suppose that the possible values that it can assume are given by $x_1, x_2, x_3, \ldots$, arranged in some order. Suppose also that these values are assumed with probabilities given by

$$P(X = x_k) = f(x_k) \quad k = 1, 2, \ldots \tag{15.4.1}$$

It is convenient to introduce the probability function, also referred to as probability distribution, given by

$$P(X = x) = f(x) \tag{15.4.2}$$

For $x = x_k$, this reduces to (15.4.1) while for other values of x, $f(x) = 0$. In general, $f(x)$ is a probability function if

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$

where the sum in 2 is taken over all possible values of x.

15.5 Continuous Probability Distribution

A continuous distribution describes the probabilities of a continuous random variable's possible values. Since a continuous random variable can take any value within its range, we cannot list all the possible values and their probabilities as in the discrete case.
For continuous random variables we represent probabilities using a probability density function (pdf) (sometimes just called the probability distribution). The pdf of a continuous random variable X, is a function $f(x)$ defined such that:

- Its curve lies on or above the x-axis, i.e. $f(x) \geq 0$ for all x in its range.
- The area under the entire curve is 1.
- The probability $P(a < X < b)$ that X lies between a and b is the area under the curve between a and b.

Below is an example of what a pdf might look like. The region green is $P(2 \leq X \leq 3)$. Tho total shaded area (purple and green) is equal to 1.

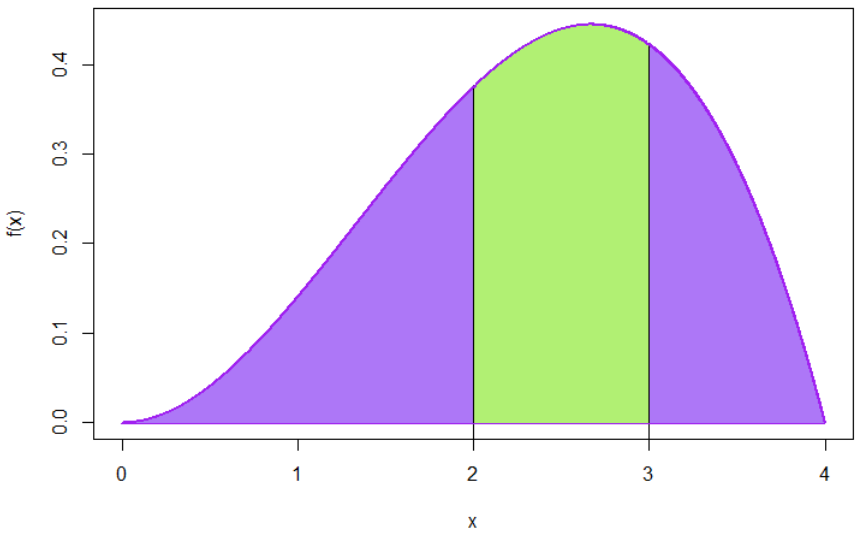


Figure 15.5.1: Example of a continuous pdf.

15.6 Distribution Functions for Random Variables

The cumulative distribution function, or briefly the distribution function, for a random variable X is defined by

$$F(x) = P(X \leq x) \tag{15.6.1}$$

where x is any real number, i.e., $-\infty < x < \infty$.
If X is discrete,

$$F(x) = \sum_{\substack{j \\ X_j \leq x}} p_j \tag{15.6.2}$$

If X is continuous,

$$F(x) = P(-\infty < X \leq x) = \int_{-\infty}^{x} f(x)\, dx \tag{15.6.3}$$

The distribution function F(x) has the following properties:

1. $F(x)$ is non-decreasing [i.e., $F(x) \leq F(y)$ if $x \leq y$].
2. $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$.
3. $F(x)$ is continuous from the right [i.e., $\lim_{h\to0^+} F(x+h) = F(x)$ for all x].
4. Let $\{x_n\}_{n\geq1}$ be a decreasing sequence of real numbers such that $\lim_{n\to\infty} x_n = x$. Then $\lim_{n\to\infty} F(x_n) = F(x)$.

15.7 Continuous Random Variables

A non-discrete random variable X is said to be absolutely continuous, or simply continuous, if its distribution function may be represented as

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\, du \quad (-\infty < x < \infty) \tag{15.7.1}$$

where the function $f(x)$ has the properties

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$

It follows from the above that if X is a continuous random variable, then the probability that X takes on any one particular value is zero, whereas the interval probability that X lies between two different values, say, a and b, is given by

$$P(a < X < b) = \int_{a}^{b} f(x)\, dx \tag{15.7.2}$$

Example 15.7.1. A random variable X has the following probability distribution.

x	-2	-1	0	1	2	3
$p(x)$	0.1	K	0.2	$2K$	0.3	$3K$

(a) Find K, (b) evaluate $P(X < 2)$ and $P(-2 < X < 2)$, (c) find the distribution function of X and (d) evaluate the mean of X.

Solution: (a) Since $\sum P(x) = 1, 6K + 0.6 = 1$
So, $K = \frac{1}{15}$.
Therefore, the probability distribution becomes

x	-2	-1	0	1	2	3
$p(x)$	$1/10$	$1/15$	$1/5$	$2/15$	$3/10$	$1/5$

(b) $P(X < 2) = P(X = -2, -1, 0 \text{ or } 1)$
$= P(X = -2) + P(X = -1) + P(X = 0) + P(X = 1)$
$= \frac{1}{10} + \frac{1}{15} + \frac{1}{5} + \frac{2}{15} = \frac{1}{2}$
(c)

$$\begin{aligned} F(x) &= 0, \quad \text{when} \quad x < -2 \\ &= \frac{1}{10}, \quad \text{when} \quad -2 \le x < -1 \\ &= \frac{1}{6}, \quad \text{when} \quad -1 \le x < 0 \\ &= \frac{11}{30}, \quad \text{when} \quad 0 \le x < 1 \\ &= \frac{1}{2}, \quad \text{when} \quad 1 \le x < 2 \\ &= \frac{4}{5}, \quad \text{when} \quad 2 \le x < 3 \\ &= 1, \quad \text{when} \quad 3 \le x \end{aligned}$$

(d) The mean of X is defined as $E(X) = \sum xp(x)$.
Therefore, mean of $X = \left(-2 \times \frac{1}{10}\right) + \left(-1 \times \frac{1}{15}\right) + \left(0 \times \frac{1}{5}\right) + \left(1 \times \frac{2}{15}\right) + \left(2 \times \frac{3}{10}\right) + \left(3 \times \frac{1}{5}\right)$
$= -\frac{1}{5} - \frac{1}{15} + \frac{2}{15} + \frac{3}{5} + \frac{3}{5} = \frac{16}{15}$

Example 15.7.2. Find the constant c such that the function

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

is a density function and compute $P(1 < X < 2)$.

Solution: Since $f(x)$ satisfies Property 1 if $c \ge 0$, it must satisfy Property 2 in order to be a density function. Now

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^3 cx^2\,dx = \left.\frac{cx^3}{3}\right|_0^3 = 9c$$

and since this must equal 1, we have $c = 1/9$. Now

$$P(1 < X < 2) = \int_1^2 \frac{1}{9}x^2\,dx = \left.\frac{x^3}{27}\right|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

15.8 Joint Distributions

15.8.1 Discrete Case:

If X and Y are two discrete random variables, we define the joint probability function of X and Y by

$$P(X = x, Y = y) = f(x, y) \tag{15.8.1}$$

where

1. $f(x, y) \geq 0$
2. $\sum_x \sum_y f(x, y) = 1$

The joint distribution function of X and Y is defined by

$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{u \leq x} \sum_{v \leq y} f(u, v) \tag{15.8.2}$$

Continuous case:

The case where both variables are continuous is obtained easily by analogy with the discrete case on replacing sums by integrals. Thus the joint probability function for the random variables X and Y (or, as it is more commonly called, the joint density function of X and Y) is defined by

1. $f(x, y) \geq 0$
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$

Properties of $F(x, y)$:

1. $F(-\infty, y) = 0 = F(x, -\infty)$ and $F(\infty, \infty) = 1$
2. $P(a < X < b, Y \leq y) = F(b, y) - F(a, y)$
3. $P(X \leq x, c < Y < d) = F(x, d) - F(x, c)$
4. $P(a < X < b, c < Y < d) = F(b, d) - F(a, d) - F(b, c) + F(a, c)$
5. At points of continuity of $f(x, y)$, $\frac{\partial^2 F}{\partial x \partial y} = f(x, y)$

The joint distribution function of X and Y in this case is defined by

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{y} f(u, v)\, du\, dv \tag{15.8.3}$$

From (15.8.3), we obtain

$$P(X \leq x) = F_1(x) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{\infty} f(u, v)\, du\, dv \tag{15.8.4}$$

$$P(Y \leq y) = F_2(y) = \int_{u=-\infty}^{\infty} \int_{v=-\infty}^{y} f(u, v)\, du\, dv \tag{15.8.5}$$

Eq. (15.8.4) and (15.8.5) are called the marginal distribution functions, or simply the distribution functions, of X and Y, respectively. The derivatives of (15.8.4) and (15.8.5) with respect to x and y are then called the marginal density functions, or simply the density functions, of X and Y and are given by

$$f_1(x) = \int_{v=-\infty}^{\infty} f(x,v)\,dv \qquad f_2(y) = \int_{u=-\infty}^{\infty} f(u,y)\,du \tag{15.8.6}$$

15.9 Change of Variables

Given the probability distributions of one or more random variables, we are often interested in finding distributions of other random variables that depend on them in some specified manner. Procedures for obtaining these distributions are presented in the following theorems for the case of discrete and continuous variables.

15.9.1 Discrete Variables

Theorem 15.9.1. Let X be a discrete random variable whose probability function is $f(x)$. Suppose that a discrete random variable U is defined in terms of X by $U = \phi(X)$, where to each value of X there corresponds one and only one value of U and conversely, so that $X = \psi(U)$. Then the probability function for U is given by

$$g(u) = f[\psi(u)] \tag{15.9.1}$$

Theorem 15.9.2. Let X and Y be discrete random variables having joint probability function $f(x,y)$. Suppose that two discrete random variables U and V are defined in terms of X and Y by $U = \phi_1(X,Y), V = \phi_2(X,Y)$, where to each pair of values of X and Y there corresponds one and only one pair of values of U and V and conversely, so that $X = \psi_1(U,V), Y = \psi_2(U,V)$. Then the joint probability function of U and V is given by

$$g(u,v) = f[\psi_1(u,v), \psi_2(u,v)] \tag{15.9.2}$$

Continuous variables

Theorem 15.9.3. Let X be a continuous random variable with probability density $f(x)$. Let us define $U = \phi(X)$ where $X = \psi(U)$ as in Theorem 15.9.1. Then the probability density of U is given by $g(u)$ where

$$g(u)|du| = f(x)|dx| \tag{15.9.3}$$

$$\text{or } g(u) = f(x)\left|\frac{dx}{du}\right| = f[\psi(u)]\left|\psi'(u)\right| \tag{15.9.4}$$

Theorem 15.9.4. Let X and Y be continuous random variables having joint density function $f(x,y)$. Let us define $U = \phi_1(X,Y)$, $V = \phi_2(X,Y)$ where $X = \psi_1(U,V)$, $Y = \psi_2(U,V)$ as in Theorem 15.9.2. Then the joint density function of U and V is given by $g(u,v)$ where

$$g(u,v)|du\,dv| = f(x,y)|dx\,dy| \tag{15.9.5}$$

$$\text{or } g(u,v) = f(x,y)\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = f[\psi_1(u,v), \psi_2(u,v)]|J| \tag{15.9.6}$$

where Jacobian determinant or briefly Jacobian, is given by

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \tag{15.9.7}$$

Example 15.9.5. The probability function of a random variable X is

$$f(x) = \begin{cases} 2^{-x} & x = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Find the probability function for the random variable $U = x^4 + 1$.

Solution: Since $U = X^4 + 1$, the relationship between the values u and x of the random variables U and X is given by $u = x^4 + 1$ or $x = \sqrt[4]{u-1}$, where $u = 2, 17, 82, \ldots$ and the real positive root is taken. Then the required probability function for U is given by

$$g(u) = \begin{cases} 2^{-\sqrt[4]{u-1}} & u = 2, 17, 82, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Example 15.9.6. If the random variables X and Y have joint density function

$$f(x) = \begin{cases} xy/96 & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

then, find the joint density function $U = XY^2,\ V = X^2Y$.

Solution: Consider $u = xy^2$, $v = x^2y$. Dividing these equations, we obtain $y/x = u/v$ so that $y = ux/v$. This leads to the simultaneous solution $x = v^{2/3}u^{-1/3},\ y = u^{2/3}v^{-1/3}$. The image of $0 < x < 4,\ 1 < y < 5$ in the uv-plane is given by

$$0 < v^{2/3}u^{-1/3} < 4 \qquad 1 < u^{2/3}v^{-1/3} < 5$$

which are equivalent to

$$v^2 < 64u \qquad v < u^2 < 125v$$

The Jacobian is given by

$$J = \begin{vmatrix} -\frac{1}{3}v^{2/3}u^{-4/3} & \frac{2}{3}v^{-1/3}u^{-1/3} \\ \frac{2}{3}u^{-1/3}v^{-1/3} & -\frac{1}{3}u^{2/3}v^{-4/3} \end{vmatrix} = -\frac{1}{3}u^{-2/3}v^{-2/3}$$

Thus the joint density function of U and V is

$$g(u,v) = \begin{cases} \frac{(v^{2/3}u^{-1/3})(u^{2/3}v^{-1/3})}{96}\left(\frac{1}{3}u^{-2/3}v^{-2/3}\right) & v^2 < 64u,\ v < u^2 < 125v \\ 0 & \text{otherwise} \end{cases}$$

$$\Rightarrow g(u,v) = \begin{cases} u^{-1/3}v^{-1/3}/288 & v^2 < 64u,\ v < u^2 < 125v \\ 0 & \text{otherwise} \end{cases}$$

15.10 Convolutions

As a particular consequence of the above theorems, we can show the density function of the sum of two continuous random variables X and Y, i.e., of $U = X + Y$, having joint density function $f(x, y)$ is given by

$$g(u) = \int_{-\infty}^{\infty} f(x, u - x)\, dx \tag{15.10.1}$$

In special case where the X and Y are independent, $f(x,y) = f_1(x)f_2(y)$, and (15.10.1) reduces to

$$g(u) = \int_{-\infty}^{\infty} f_1(x)f_2(u-x)\,dx \tag{15.10.2}$$

which is called the convolution of f_1 and f_2, abbreviated, $f_1 * f_2$. The following are some important properties of the convolution:

1. $f_1 * f_2 = f_2 * f_1$
2. $f_1 * (f_2 * f_3) = (f_1 * f_2) * f_3$
3. $f_1 * (f_2 + f_3) = (f_1 * f_2) + (f_1 * f_3)$

These results show that f_1, f_2, f_3 obey the commutative, associative and distributive laws of algebra with respect to the operation of convolution.

Theorem 15.10.1. Let X and Y be random variables having joint density function $f(x,y)$. Prove that the density function of $U = X + Y$ is

$$g(u) = \int_{-\infty}^{\infty} f(v, u-v)\,dv$$

Proof. Let $U = X+Y, V = X$, where we have arbitrary added the second equation. Corresponding to these we have $u = x + y,\ v = x$ or $x = v,\ y = u - v$. The Jacobian of the transformation is given by

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$

Thus, the joint density function of U and V is

$$g(u,v) = f(v, u-v)$$

Therefore, the marginal density function of U is

$$g(u) = \int_{-\infty}^{\infty} f(v, u-v)\,dv$$

□

Example 15.10.2. If X and Y are independent random variables having density functions

$$f_1(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad f_2(x) = \begin{cases} 3e^{-3y} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

find the density function of their sum, $U = X + Y$.

Solution: The required density function is the the convolution of f_1 and f_2 is given by

$$g(u) = f_1 * f_2 = \int_{-\infty}^{\infty} f_1(v)f_2(u-v)\,dv$$

In the integrand f_1 vanish when $v < 0$ and f_2 vanishes when $v > u$. Hence

$$\begin{aligned} g(u) &= \int_0^u (2e^{-2v})(3e^{-3(u-v)}dv \\ &= 6e^{-3u}\int_0^u e^v dv = 6e^{-3u}(e^u - 1) = 6(e^{-2u} - e^{-3u}) \end{aligned}$$

if $u \geq 0$ and $g(u) = 0$ if $u < 0$. Check

$$\int_{-\infty}^{\infty} g(u)\, du = 6\int_0^{\infty}(e^{-2u} - e^{-3u})\, du = 6\left(\frac{1}{2} - \frac{1}{3}\right) = 1$$

Unit 16

Course Structure

- Expectation. Moments. Moment inequalities, Characteristic function. Convergence of sequence of random variables-weak convergence, strong convergence and convergence in distribution, continuity theorem for characteristic functions. Weak and strong law of large numbers. Central Limit Theorem.

16.1 Mathematical Expectation

A very important concept in probability and statistics is that of the mathematical expectation, expected value, or briefly the expectation, of a random variable. For a discrete random variable X having the possible values $x_1, \ldots, x_n$, the expectation of X is defined as

$$E(X) = x_1 P(X = x_1) + \ldots + x_n P(X = x_n) = \sum_{j=1}^{n} x_j P(X = x_j) \tag{16.1.1}$$

For a continuous random variable X having density function $f(x)$, the expectation of X is defined as

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx \tag{16.1.2}$$

Another quantity of great importance in probability and statistics is called the variance and is defined by

$$\text{Var}(X) = E[(X - \mu)^2] \tag{16.1.3}$$

The variance is a non-negative number. The positive square root of the variance is called the standard deviation and is given by

$$\sigma_X = \sqrt{\text{Var}(X)} = \sqrt{E[(X - \mu)^2]} \tag{16.1.4}$$

If X is a continuous random variable having density function $f(x)$, then the variance is given by

$$\sigma_X^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx \tag{16.1.5}$$

provided that the integral converges.

Example 16.1.1. Two cards are drawn at random with replacement from a box which contains 4 cards numbered 1, 1, 2 and 2. If X denotes the sum of the numbers shown on the two cards, find the mean and variance of X

Solution: X can take the values 2, 3 and 4.

$$\begin{aligned} P(X=2) &= P \text{ (two 1's are drawn)} \\ &= \frac{{}^2C_2}{{}^4C_2} = \frac{1}{6} \end{aligned}$$

$$\begin{aligned} P(X=3) &= P \text{ (one 1 and one 2 are drawn)} \\ &= \frac{{}^2C_1.{}^2C_1}{{}^4C_2} = \frac{4}{6} \end{aligned}$$

$$\begin{aligned} P(X=4) &= P \text{ (two 2's are drawn)} \\ &= \frac{{}^2C_1}{{}^4C_2} = \frac{1}{6} \end{aligned}$$

The probability distribution of X is as shown.

$X = x_i$	2	3	4
p_i	$\frac{1}{6}$	$\frac{4}{6}$	$\frac{1}{6}$

$E(X) = \sum p_i x_i = \frac{2}{6} + \frac{12}{6} + \frac{4}{6} = 3$
$E(X^2) = \sum p_i {x_i}^2 = \frac{4}{6} + \frac{36}{6} + \frac{16}{6} = \frac{28}{3}$
$Var(X) = E(X^2) - E(X)^2 = \frac{28}{3} - 9 = \frac{1}{3}$

16.2 Moments

The r-th moment of a random variable X about the mean μ , also called the r-th central moment, is defined as

$$\mu_r = E[(X-\mu)^r] \tag{16.2.1}$$

where $r = 0, 1, 2, \ldots$. It follows that $\mu_0 = 1, \mu_1 = 0$ and $\mu_2 = \sigma^2$, i.e., the second central moment or second moment about the mean is the variance. We have, assuming absolute convergence,

$$\mu_r = \sum (x-\mu)^r f(x) \quad \text{(discrete variable)} \tag{16.2.2}$$

$$\mu_r = \int_{-\infty}^{\infty} (x-\mu)^r f(x)\, dx \quad \text{(continuous variable)} \tag{16.2.3}$$

The r-th moment of X about the origin, also called the r-th raw moment, is defined as

$$\mu'_r = E(X^r) \tag{16.2.4}$$

where $r = 0, 1, 2, \ldots$, and in this case there are formulas analogous to (16.2.2) and (16.2.3) in which $\mu = 0$. The relation between these moments is given by

$$\mu_r = \mu'_r - \binom{r}{1}\mu'_{r-1}\mu + \ldots + (-1)^j \binom{r}{j}\mu'_{r-j}\mu^j + \ldots + (-1)^r \mu'_0 \mu^r \tag{16.2.5}$$

Proof.

$$\begin{aligned}
\mu_r &= E(X-\mu)^r] \\
&= E\left[X^r - \binom{r}{1}X^{r-1}\mu + \cdots + (-1)^j\binom{r}{j}X^{r-j}\mu^j + \ldots + (-1)^{r-1}\binom{r}{r-1}X\mu^{r-1} + (-1)^r\mu^r\right] \\
&= E(X^r) - \binom{r}{1}E(X^{r-1})\mu + \ldots + (-1)^j\binom{r}{j}E(X^{r-j})\mu^j + \ldots + (-1)^{r-1}\binom{r}{r-1}E(X)\mu^{r-1} \\
&\qquad\qquad +(-1)^r\mu^r \\
&= \mu'_r - \binom{r}{1}\mu'_{r-1}\mu + \ldots + (-1)^j\binom{r}{j}\mu'_{r-j}\mu^j + \ldots + (-1)^{r-1}r\mu^r + (-1)^{-r}\mu^r
\end{aligned}$$

where the last two terms can be combined to give $(-1)^{r-1}(r-1)\mu^r$. □

16.3 Moment Generating Functions

The moment generating function of X is defined by

$$M_X(t) = E(e^{tX}) \tag{16.3.1}$$

Using the power series expansion, we have

$$\begin{aligned}
M_X(t) &= E(e^tX) = E\left(1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \ldots\right) \\
&= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \ldots \\
&= 1 + \mu t + \mu'_2\frac{t^2}{2!} + \mu'_3\frac{t^3}{3!} + \ldots
\end{aligned} \tag{16.3.2}$$

Since the coefficients in this expansion enable us to find the moments, the reason for the name moment generating function is apparent. From the expansion we can show that

$$\mu'_r = \frac{d^r}{dt^r}M_X(t)\bigg|_{t=0} \tag{16.3.3}$$

i.e., μ'_r is the r-th derivative of $M_X(t)$ evaluated at $t = 0$.

Example 16.3.1. A random variable X has density function given by

$$f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{16.3.4}$$

Find the moment generating function and the first four moments about origin.

Solution: We have

$$\begin{aligned}
M(t) &= E(e^{tX}) = \int_{-\infty}^{\infty} e^{tX}f(x)\,dx \\
&= \int_0^{\infty} e^{tX}(2e^{-2x})\,dx = 2\int_0^{\infty} e^{(t-2x)}dx \\
&= \frac{2e^{(t-2)x}}{t-2}\bigg|_0^{\infty} = \frac{2}{2-t}, \quad \text{assuming } t < 2
\end{aligned}$$

If $|t| < 2$, we have

$$\frac{2}{2-t} = \frac{1}{1-\frac{t}{2}} = 1 + \frac{t}{2} + \frac{t^2}{4} + \frac{t^3}{8} + \frac{t^4}{16} + \cdots$$

But

$$M(t) = 1 + \mu t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \mu_4' \frac{t^4}{4!} + \cdots$$

Therefore, on comparing terms, we have $\mu = \frac{1}{2},\ \mu_2' = \frac{1}{2},\ \mu_3' = \frac{3}{4},\ \mu_4' = \frac{3}{2}$.

16.4 Characteristic Function

If we let $t = i\omega$, where i is the imaginary unit, in the moment generating function we obtain an important function called the characteristic function. We denote this by

$$\phi_X(\omega) = M_X(i\omega) = E(e^{i\omega X}) \tag{16.4.1}$$

It follows that

$$\phi_X(\omega) = \sum e^{i\omega x} f(x) \quad \text{(discrete variable)} \tag{16.4.2}$$

$$\phi_X(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} f(x) \quad \text{(continuous variable)} \tag{16.4.3}$$

Since $|e^{i\omega t}| = 1$, the series and the integral always converge absolutely.
The corresponding results (16.3.2) and (16.3.3) becomes

$$\phi_X(\omega) = 1 + i\mu\omega - \mu_2' \frac{\omega^2}{2!} + \cdots + i^r \mu_r' \frac{\omega^r}{r!} + \cdots \tag{16.4.4}$$

where

$$\mu_r' = (-1)^r i^r \frac{d^r}{d\omega^r} \phi_X(\omega)\Big|_{\omega=0} \tag{16.4.5}$$

Theorem 16.4.1. If $\phi_X(\omega)$ is the characteristic function of the random variable X and a and b $(b \neq 0)$ are constants, then the characteristic function of $(X + a)/b$ is

$$\phi_{(X+a)/b}(\omega) = e^{ai\omega/b} \phi_X\left(\frac{\omega}{b}\right) \tag{16.4.6}$$

Theorem 16.4.2. If X and Y are independent random variables having characteristic functions $\phi_X(\omega)$ and $\phi_Y(\omega)$, respectively, then

$$\phi_{X+Y}(\omega) = \phi_X(\omega)\phi_Y(\omega) \tag{16.4.7}$$

Example 16.4.3. Find the characteristic function of the random variable X having density function given by

$$f(x) = \begin{cases} 1/2a & |x| < a \\ 0 & \text{otherwise} \end{cases}$$

Solution: The characteristic function is given by

$$E(e^{i\omega X}) = \int_{-\infty}^{\infty} e^{i\omega x} f(x)\, dx = \frac{1}{2a} \int_{-a}^{a} e^{i\omega x} dx$$

$$= \frac{1}{2a} \frac{e^{i\omega x}}{i\omega}\Big|_{-a}^{a} = \frac{e^{ia\omega} - e^{-ia\omega}}{2ia\omega} = \frac{\sin a\omega}{a\omega}$$

Example 16.4.4. Find the characteristic function of the random variable X having density function $f(x) = ce^{-a|x|}$, $-\infty < x < \infty$, where $a > 0$, and c is a suitable constant.

Solution: Since $f(x)$ is a density function, we must have

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

so that

$$\begin{aligned} c\int_{-\infty}^{\infty} e^{-a|x|}dx &= c\left[\int_{-\infty}^{0} e^{-a(-x)}dx + \int_{0}^{\infty} e^{-ax}dx\right] \\ &= c\frac{e^{ax}}{a}\bigg|_{-\infty}^{0} + c\frac{e^{-ax}}{-a}\bigg|_{0}^{\infty} = \frac{2c}{a} = 1 \end{aligned}$$

Then $c = a/2$. The characteristic function is therefore given by

$$\begin{aligned} E(e^{i\omega X}) &= \int_{-\infty}^{\infty} e^{i\omega x} f(x)\,dx \\ &= \frac{a}{2}\left[\int_{-\infty}^{0} e^{i\omega x}e^{-a(-x)}dx + \int_{0}^{\infty} e^{i\omega x}e^{-ax}dx\right] \\ &= \frac{a}{2}\left[\int_{-\infty}^{0} e^{(a+i\omega)x}dx + \int_{0}^{\infty} e^{-(a-i\omega)x}e^{-ax}dx\right] \\ &= \frac{a}{2}\frac{e^{(a+i\omega)x}}{a+i\omega}\bigg|_{-\infty}^{0} + a\frac{e^{-(a-i\omega)x}}{-(a-i\omega)}\bigg|_{0}^{\infty} \\ &= \frac{a}{2(a+i\omega)} + \frac{a}{2(a-i\omega)} \\ &= \frac{a^2}{a^2+\omega^2} \end{aligned}$$

16.5 Chebyshev's Inequality

Suppose that X is a random variable (discrete or continuous) having mean μ and variance σ^2, which are finite. Then if ϵ is any positive number,

$$P(|X-\mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2} \tag{16.5.1}$$

or, with $\epsilon = k\sigma$

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2} \tag{16.5.2}$$

Proof. We shall present the proof fro continuous random variables. A proof for discrete variables is similar if integrals are replaced by sums. If $f(x)$ is the density function of X, then

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

Since the integrand is non-negative, the value of the integral can only decrease when the range of integration is diminished. Therefore,

$$\sigma^2 \geq \int\limits_{|x-\mu|\geq\epsilon} (x-\mu)^2 f(x)\,dx \geq \int\limits_{|x-\mu|\geq\epsilon} \epsilon^2 f(x)\,dx = \epsilon^2 \int\limits_{|x-\mu|\geq\epsilon} f(x)\,dx$$

But the last integral is equal to $P(|X-\mu|\geq\epsilon)$. Hence,

$$P(|X-\mu|\geq\epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

□

16.6 Law of Large Numbers

Theorem 16.6.1. Let $X_1, X_2, \ldots, X_n$ be mutually independent random variables (discrete or continuous), each having finite mean μ and variance σ^2. Then if $S_n = X_1 + X_2 + \cdots + X_n (n = 1, 2, \ldots)$,

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0 \tag{16.6.1}$$

Proof. We have

$$E(X_1) = E(X_2) = \cdots = E(X_n) = \mu$$
$$\text{Var}(X_1) = \text{Var}(X_2) = \cdots = \text{Var}(X_n) = \sigma^2$$

Then

$$E\left(\frac{S_n}{n}\right) = E\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{1}{n}[E(X_1) + \cdots + E(X_n)] = \frac{1}{n}(n\mu) = \mu$$
$$\text{Var}(S_n) = \text{Var}(X_1 + \cdots + X_n) = \text{Var}(X_1) + \cdots + \text{Var}(X_n) = n\sigma^2$$

so that

$$\text{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\text{Var}(S_n) = \frac{\sigma^2}{n}$$

Therefore, by Chebyshev's inequality with $X = S_n/n$, we have

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}$$

Taking the limit as $n \to \infty$, this becomes, as required,

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0$$

□

Note: Since S_n/n is the arithmetic mean of $X_1, \ldots, X_n$, this theorem states that the probability of the arithmetic mean S_n/n differing from its expected value μ by more than ϵ approaches zero as $n \to \infty$. A stronger result which we might expect to be true, is that

$$\lim_{n\to\infty} \frac{S_n}{n} = \mu,$$

but this is actually false. However, we can prove that $\lim\limits_{n\to\infty} S_n/n = \mu$ with probability one. This result is often called the strong law of large numbers, and, by contrast Theorem 16.6.1 is called the weak law of large numbers.

16.7 Special Probability Distributions

16.7.1 The Binomial Distribution

Let p be the probability that an event will happen in any single Bernoulli trial (called the probability of success). Then $q = 1 - p$ is the probability that the event will fail to happen in any single trial (called the probability of failure). The probability that the event will happen exactly x times in n trials (i.e., x successes and $n - x$ failures will occur) is given by the probability function

$$f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x} \tag{16.7.1}$$

where the random variable X denotes the number of successes in n trials and $x = 0, 1, \ldots, n$.

1. Mean of Binomial Distribution, $\mu = np$
2. Variance of Binomial Distribution, $\sigma^2 = npq$
3. Moment generating function $M(t) = (q + pe^t)^n$
4. Characteristic function $\phi(\omega) = (q + pe^{i\omega})^n$

16.7.2 The Normal Distribution

One of the most important examples of a continuous probability distribution is the normal distribution, some times called the Gaussian distribution. The density function for this distribution is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty \tag{16.7.2}$$

where μ and σ are the mean and standard deviation, respectively. The corresponding distribution function is given by

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(v-\mu)^2}{2\sigma^2}} dv \tag{16.7.3}$$

If X has the distribution function given by (16.7.3), we say that the random variable X is normally distributed with mean μ and variance σ^2. If we let Z be the standardized variable corresponding to X, i.e., if we let

$$Z = \frac{X - \mu}{\sigma} \tag{16.7.4}$$

then the mean or expected value of Z is 0 and the variance is 1. In such case the density function for Z can be obtained from (16.7.2) by formally placing $\mu = 0$ and $\sigma = 1$, yielding

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \tag{16.7.5}$$

This is often referred to as the standard normal density function.

1. Mean of Normal Distribution, μ
2. Variance of Binomial Distribution, σ^2
3. Moment generating function $M(t) = e^{ut + (\sigma^2 t^2/2)}$
4. Characteristic function $\phi(\omega) = e^{i\mu\omega - (\sigma^2\omega^2/2)}$

16.7.3 Relation Between Binomial and Normal Distributions

If n is large and if neither p nor q is too close to zero, the binomial distribution can be closely approximated by a normal distribution with standardized random variable given by

$$Z = \frac{X - np}{\sqrt{npq}} \tag{16.7.6}$$

Here X is the random variable giving the number of successes in n Bernoulli trials and p is the probability of success. The approximation becomes better with increasing n and is exact in the limiting case. The fact that the binomial distribution approaches the normal distribution can be described by writing

$$\lim_{n\to\infty} P\left(a \leq \frac{X - np}{\sqrt{npq}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2} du \tag{16.7.7}$$

In words, we say that the standardized random variable $(X - np)/\sqrt{npq}$ is asymptotically normal.

16.7.4 The Poisson Distribution

Let X be a discrete random variable that can take on the values $0, 1, 2, \ldots$ such that the probability function of X is given by

$$f(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots \tag{16.7.8}$$

where λ is a given positive constant. This distribution is called the Poisson distribution and random variable having this distribution is said to be Poisson distributed.

1. Mean of Poisson Distribution, $\mu = \lambda$
2. Variance of Poisson Distribution, $\sigma^2 = \lambda$
3. Moment generating function $M(t) = e^{\lambda(e^t - 1)}$
4. Characteristic function $\phi(\omega) = e^{\lambda(e^{i\omega} - 1}$

16.7.5 Relation Between the Poisson and Normal Distribution

We can show that if X is the Poisson random variable of (16.7.8) and $(X - \lambda)/\sqrt{\lambda}$ is the corresponding standardized random variable, then

$$\lim_{\lambda\to\infty} P\left(a \leq \frac{X - \lambda}{\sqrt{\lambda}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2} du \tag{16.7.9}$$

i.e., the Poisson distribution approaches the normal distribution as $\lambda \to \infty$ or $(X - \lambda)/\sqrt{\lambda}$ is asymptotically normal.

16.8 The Central Limit Theorem

The similarity between (16.7.7) and (16.7.9) naturally leads us to ask whether there are any other distribution besides the binomial and Poisson that have the normal distribution as the limiting case. The following remarkable theorem reveal that actually a large class of distribution have this property.

Theorem 16.8.1. Let $X_1, X_2, \ldots, X_n$ be independent random variables that are identically distributed (i.e., all have the same probability function in the discrete case or density function in the continuous case) and have finite mean μ and variance σ^2. Then is $S_n = X_1 + X_2 + \ldots + X_n$ $(n = 1, 2, \ldots)$,

$$\lim_{n\to\infty} P\left(a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2} du \tag{16.8.1}$$

that is, the random variable $(S_n - n\mu)/\sigma\sqrt{n}$, which is the standardized variable corresponding to S_n, is asymptotically normal

Proof. For $n = 1, 2, \ldots$, we have $S_n = X_1 + X_2 + \ldots + X_n$. Now $X_1, X_2, \ldots, X_n$ each have mean μ and variance σ^2. Thus,

$$E(S_n) = E(X_1) + E(X_2) + \cdots + E(X_n) = n\mu$$

and, because the X_k are independent,

$$\text{Var}(S_n) = \text{Var}(X_1) + \text{Var}(X_2) + \cdots + \text{Var}(X_n) = n\sigma^2$$

It follows that the standardized random variable corresponding to S_n is

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

The moment generating function for S_n^* is

$$\begin{aligned} E(e^{tS_n^*}) &= E[e^{t(S_n - n\mu)/\sigma\sqrt{n}}] \\ &= E[e^{t(X_1-\mu)/\sigma\sqrt{n}} e^{t(X_2-\mu)/\sigma\sqrt{n}} \cdots e^{t(X_n-\mu)/\sigma\sqrt{n}}] \\ &= E[e^{t(X_1-\mu)/\sigma\sqrt{n}}] \cdot E[e^{t(X_2-\mu)/\sigma\sqrt{n}}] \cdots E[e^{t(X_n-\mu)/\sigma\sqrt{n}}] \\ &= \{E[e^{t(X_1-\mu)/\sigma\sqrt{n}}]\}^n \end{aligned}$$

where, in the last two steps, we have respectively used the facts that the X_k are independent and are identically distributed. Now, by a Taylor series expansion,

$$\begin{aligned} E[e^{t(X_1-\mu)/\sigma\sqrt{n}}] &= E\left[1 + \frac{t(X_1-\mu)}{\sigma\sqrt{n}} + \frac{t^2(X_1-\mu)^2}{2\sigma^2 n} + \cdots\right] \\ &= E(1) + \frac{t}{\sigma\sqrt{n}} E(X_1 - \mu) + \frac{t^2}{2\sigma^2 n} E[(X_1-\mu)^2] + \cdots \\ &= 1 + \frac{t}{\sigma\sqrt{n}}(0) + \frac{t^2}{2\sigma^2 n}(\sigma^2) + \cdots \\ &= 1 + \frac{t^2}{2n} + \cdots \end{aligned}$$

so that

$$E(e^{tS_n^*}) = \left(1 + \frac{t^2}{2n} + \cdots\right)^n$$

But the limit of this as $n \to \infty$ is $e^{t^2/2}$, which is the moment generating function of the standardized normal distribution. Hence, the required result follows. □

Unit 17

Course Structure

- Definition and classification of stochastic processes
- Markov chains with finite and countable state space
- Classification of states.

17.1 Introduction

Since the last century there have been marked changes in the approach to scientific enquiries. There has been greater realisation that probability (or non-deterministic) models are more realistic than deterministic models in many situations. Observations taken at different time points rather than those taken at a fixed period of time began to engage the attention of probabilist. This lead to a new concept of indeterminism: indeterminism in dynamic studies. This has been called "dynamic indeterminism". Many phenomenon occurring in physical and life sciences are studied now not only as a random phenomenon but also as one changing with time or space. Similar considerations are also made in other areas, such as, social sciences, engineering and management and so on. The scope of applications of random variables which are functions of time or space or both has been on the increase.

Families of random variables which are functions of say, time, are known as stochastic processes (or random processes, or random functions). A few simple examples are given as illustrations.

Example 17.1.1. Consider a simple experiment like throwing a true die.

(i) Suppose that X_n is the outcome of the n-th throw, $n \geq 1$. Then $\{X_n,\ n \geq 1\}$ is a family of random variables such that for a distinct value of $n\ (= 1, 2, \ldots)$, one get a distinct random variable X_n; $\{X_n,\ n \geq 1\}$ constitutes a stochastic process, known as Bernoulli process.

(ii) Suppose that X_n is the number of sixes in the first n throws. For a distinct value of $n = 1, 2, \ldots,$ we get a distinct binomial variable X_n; $\{X_n,\ n \geq 1\}$ which gives a family of random variables is a stochastic process.

(iii) Suppose that X_n is the maximum number shown in the first n throws. Here $\{X_n,\ n \geq 1\}$ constitutes a stochastic process.

Example 17.1.2. Consider that there are r cells and an infinitely large number of identical balls and that balls are thrown at random, one by one, into the cells, the ball thrown being equally likely to go into any one of the cells. Suppose that X_n is the number of occupied cells after n throws. Then $\{X_n,\ n \geq 1\}$ constitutes a stochastic process.

17.2 Specification of Stochastic Processes

The set of possible values of a single random variable X_n of a stochastic process $\{X_n,\ n \geq 1\}$ is known as its *state space.* The state space is discrete if it contains a finite or a denumerable infinity of points; otherwise, it is continuous.

For example, if X_n is the toal number of sixes appearing in the first n throws of a die, the set of possible values of X_n is finite set of non-negative integers $0, 1, \ldots, n$. Here, the state space of X_n is discrete. We can write

$$X_n = Y_1 + Y_2 + \ldots + Y_n,$$

where Y_i is a discrete random variable denoting the outcome of the i-th throw and $Y_i = 1$ or 0 according as the i-th throw shows six or not. Secondly, consider

$$X_n = Z_1 + Z_2 + \ldots + Z_n,$$

where Z_i is a continuous random variable assuming values $[0, \infty)$. Here, the set of possible values of X_n is the interval $[0, \infty)$, and so the state space of X_n is continuous.

In the above two examples, we assume that the parameter n of X_n is restricted to the non-negative integers $n = 0, 1, 2, \ldots$ We consider the state of the system at distinct time points $n = 0, 1, 2, \ldots$, only. Here the word *time* is used in a wide sense. We note that in the first case considered above the "time n" implies throw number n.

On the other hand, one can visualise a family of random variables $\{X_t,\ t \in T\}$ (or $\{X(t),\ t \in T\}$) such that the state of the system is characterized at every instant over a finite or infinite interval. The system is then defined for a continuous range of time and we say that we have a family of random variable in *continuous* time. A stochastic process in continuous time may have either a discrete or a continuous state space. For example, suppose that $X(t)$ gives the number of incoming calls at a switchboard in an interval $(0, t)$. Here the state space of $X(t)$ is discrete through $X(t)$ is defined for a continuous range of time. We have a process in continuous time having a discrete state space. Suppose that $X(t)$ represents the maximum temperature at a particular place in $(0, t)$, then the set of possible values of $X(t)$ is continuous. Here we have a system in continuous time having a continuous state space.

So far we have assumed that the values assumed by the random variable X_n (or $X(t)$) are one-dimensional, but he process $\{X_n\}$ (or $\{X(t)\}$) may be multi-dimensional. Consider $X(t) = (X_1(t), X_2(t))$, where X_1 represents the maximum and X_2 the minimum temperature at a place in an interval of time $(0, t)$, We have here a two-dimensional stochastic process in continuous time having continuous state space. One can similarly have multi-dimensional processes. One-dimensional processes can be classified, in general, into the following four types of processes:

- Discrete time, discrete state space
- Discrete time, continuous state space

- Continuous time, discrete state space
- Continuous time, continuous state space.

All the four types may be represented by $\{X(t),\ t \in T\}$. In case of discrete time, the parameter generally used is n, i.e., the family is represented by $\{X_n,\ n = 0, 1, 2, \ldots\}$. In case of continuous time both the symbols $\{X_t,\ t \in T\}$ and $\{X(t),\ t \in T\}$ (where T is a finite or infinite interval) are used . The parameter t is usually interpreted as time, though it may represent such characters as distance, length, thickness and so on. We shall use the notation $\{X(t),\ t \in T\}$ both in the cases of discrete and continuous parameters and shall specify, whenever necessary.

Relationship

In some of the cases, the random variable X_n, i.e., members of the family $\{X_n,\ n \geq 1\}$ are mutually independent, but more often they are not. We generally come across processes whose members are mutually dependent. The relationship among them is often of great importance.

The nature of dependence could be infinitely varied. Here dependence of some special types, which occurs quite often and is of great importance, will be considered. We may broadly describe some stochastic processes according to the nature of dependence relationship existing among the members of the family.

Processes with independent increments

If for all $t_1, \ldots, t_n$, $t_1 < t_2 < \ldots < t_n$, the random variables

$$X(t_2) - X(t_1),\ \ X(t_3) - X(t_2),\ \ \ldots,\ \ X(t_n) - X(t_{n-1})$$

are independent, then $\{X(t),\ t \in T\}$ is said to be a process with independent increments.

Suppose that we wish to consider the discrete parameter case. Consider a process in discrete time with independent increments. Writing

$$T = \{0, 1, 2, \ldots\},\ \ t_i = i - 1,\ X(t_i) = X_{i-1},$$
$$Z_i = X_i - X_{i-1},\ \ i = 1, 2, \ldots \ \text{ and } \ Z_0 = X_0,$$

We have a sequence of independent random variables $\{Z_n,\ n \geq 0\}$.

Markov Process If $\{X(t), t \in T\}$ is a stochastic process such that, given the value $X(s)$, the values of $X(t), t > s$, do not depend on the values of $X(u),\ u < s$, then the process is said to be a Markov process. A definition of such a process is given below. If, for, $t_1 < t_2 < \ldots < t_n < t$,

$$P\Big\{a \leq X(t) \leq b | X(t_1) = x_1, \ldots, X(t_n) = x_n\Big\} = P\Big\{a \leq X(t) \leq b | X(t_n) = x_n\}$$

the process $\{X(t),\ t \in T\}$ is a *Makrov process*. A discrete parameter Markov process is known as a *Markov chain.*

17.3 Markov Chains

Consider a simple coin tossing experiment repeated for a number of times. The possible outcomes at each trial are two: head with probability, say, p and tail with probability q, $p + q = 1$. Let us denote head by 1 and

tail by 0 and the random variable denoting the result of the n-th toss by X_n. Then for $n = 1, 2, 3, \ldots$

$$P\{X_n = 1\} = p, \quad P\{X_n = 0\} = q.$$

Thus we have a sequence of random variables $X_1,\ X_2, \ldots$. The trials are independent and the result of the n-th trial does not depend in any way on the previous trials numbered $1,\ 2,\ \ldots,\ (n-1)$. The random variables are independent.

Consider now the random variable given by the partial sum $S_n = X_1 + \ldots + X_n$. The sum S_n gives the accumulated number of heads in the first n trials and its possible values are $0,\ 1,\ \ldots, n$.

We have $S_{n+1} = S_n + X_{n+1}$. Given that $S_n = j$ $(j = 0, 1, \ldots, n)$, the random variable S_{n+1} can assume only two possible values: $S_{n+1} = j$ with probability q and $S_{n+1} = j+1$ with probability p; these probabilities are not at all affected by the values of the variables $S_1, S_2, \ldots, S_{n-1}$. Thus

$$P\{S_{n+1} = j + 1 | S_n = j\} = p$$
$$P\{S_{n+1} = j | S_n = j\} = q.$$

We have an example of a Markov chain, a case of simple dependence that the outcome of $(n + 1)$-st trial depends directly on that of n-th trial and *only* on it. The conditional probability of S_{n+1} given S_n depends on the value of S_n and the manner in which the value of S_n was reached is of no consequence.

Definition 17.3.1. The stochastic process $\{X_n,\ n = 0, 1, 2, \ldots\}$ is called a Markov chain, if, for $j,\ k,\ j_1, \ldots, j_{n-1} \in N$ (or any subset of I),

$$P\{X_n = k | X_{n-1} = j,\ X_{n-1} = j_1, \ldots, X_0 = j_{n-1}\}$$
$$P\{X_n = k | X_{n-1} = j\} = p_{jk} \ \text{(say)}$$

whenever the first member is defined.

The outcomes are called the states of the Markov chain; if X_n has the outcome j (i.e., $X_n = j$), the process is said to be at state j at n-th trial. To a pair of states (j,k) at the two successive trials (say, n-th and (n+1)-th trials) there is an associated conditional probability p_{jk}. It is the probability of transition from the state j at n-th trial to the state k at (n+1)-th trial. The transition probabilities p_{jk} are basic to the study of the structure of the Markov chain.

The transition probability may or may not be independent of n. If the transition probability p_{jk} is independent of n, the Markov chain is said to be *homogeneous* (or to have *stationary transition probabilities*). If it is dependent on n, the chain is said to be non-homogeneous. Here we shall confine to *homogeneous chains*.

17.4 Transition Probabilities and Transition Matrix

For a finite Markov chain with m states $E_1, E_2, \ldots, E_m$, introduce the notation

$$p_{ij} = P\{X_n = j | X_{n-1} = i\} \tag{17.4.1}$$

where $i, j = 1, 2, \ldots, m$. The numbers p_{ij} are known as the transition probabilities of the chain, and must satisfy

$$p_{ij} \geq 0, \quad \sum_{j=1}^{m} p_{ij} = 1$$

for each $i = 1, 2, \ldots, m$.

Transition probabilities form an $m \times m$ array which can be assembled into a transition matrix T, where

$$T = [p_{ij}] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix} \tag{17.4.2}$$

Note that each row of T is a probability distribution. Any square matrix for which $p_{ij} \geq 0$ and $\sum_{j=1}^{m} p_{ij} = 1$ is said to be row-stochastic.

Example 17.4.1. The matrix $A = [a_{ij}]$ and $B = [b_{ij}]$ are $m \times n$ row-stochastic matrices. Show that $C = AB$ is also row-stochastic.

Solution: By the multiplication rule for the matrices

$$C = AB = [a_{ij}][b_{ij}] = \sum_{k=1}^{m} a_{ik} b_{kj}.$$

Since $a_{ij} \geq 0$ and $b_{ij} \geq 0$ for all $i, j = 1, 2, \ldots, m$, it follows that $c_{ij} \geq 0$. Also

$$\sum_{j=1}^{m} c_{ij} = \sum_{j=1}^{m} \sum_{k=1}^{m} a_{ik} b_{kj} = \sum_{k=1}^{m} a_{ik} \sum_{j=1}^{m} b_{kj} = \sum_{k=1}^{m} a_{ik} \cdot 1 = 1,$$

since, $\sum_{j=1}^{m} b_{kj} = 1$ and $\sum_{k=1}^{m} a_{ik} = 1$.

It follows from this example that any power T^n of the transition matrix T must also be row-stochastic.

17.5 Classification of States

Let us consider the general m-state chain with states $E_1, E_2, \ldots, E_m$ and transition matrix

$$T = [p_{ij}], \quad (1 \leq i, j \leq m)$$

For a homogeneous chain, recollect that p_{ij} is the probability that a transition occurs between E_i and E_j at any step or change of state in the chain. We intend to investigate and classify some of the more common types of states which can occur in Markov chains.

(a) **Absorbing state:** An absorbing state E_i is characterised by the probabilities

$$p_{ii} = 1, \quad p_{ij} = 0, \ \ (i \neq j, \ \ j = 1, 2, \ldots, m)$$

in the i-th row of T.

(a) **Periodic state:** The probability of a return to E_i at step n is $p_{ii}^{(n)}$. Let t be an integer greater than 1. Suppose that

$$\begin{aligned} p_{ii}^{(n)} &= 0 \quad \text{for} \quad n \neq t, \ 2t, \ 3t, \ldots \\ p_{ii}^{(n)} &\neq 0 \quad \text{for} \quad n = t, \ 2t, \ 3t, \ldots \end{aligned}$$

In this case, the state E_i is said to be periodic with period t. If, for a state, no such t exists with this property, then the state is described as aperiodic. Let

$$d(i) = \gcd\{n | p_{ii}^{(n)} > 0\}, \tag{17.5.1}$$

that is, the greatest common divisor of the set of integers n for which $p_{ii}^{(n)} > 0$. Then the state E_i is said to be periodic if $d(i) > 1$ and aperiodic if $d(i) = 1$.

Example 17.5.1. A four-state Markov chain has the transition matrix

$$T = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Show that all states have period 3.

Solution: The transition diagram is shown in Fig. 17.5.1, from which it is clear that all states are period 3. For example, if the chain start in E_1, then returns to E_1 are only possible at steps 3, 6, 9, ... either through E_2 or E_3.

The analysis of chains with periodic states can be complicated. However, one can check for a suspected

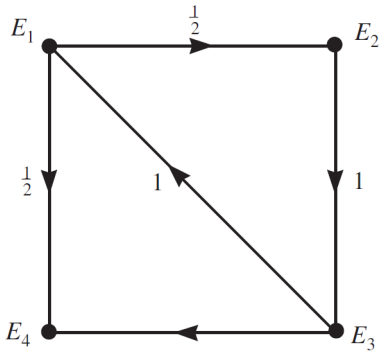


Figure 17.5.1: The transition diagram for Example 17.5.1

periodicity as follows. By direct computation

$$S = T^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

In this example,

$$S^2 = T^6 = S \cdot S = S,$$

so that

$$S^r = T^{3r} = S, \quad (r = 1, 2, \ldots),$$

which always has non-zero elements on its diagonal. On the other hand,

$$S^{r+1} = S^r S = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad S^{r+2} = S^r S^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

and both these matrices have zero diagonal elements for $r = 1, 2, 3 \ldots$. Hence, for $i = 1, 2, 3, 4$,

$$\begin{aligned} p_{ii}^{(n)} &= 0 \quad \text{for} \quad n \neq 3, 6, 9, \ldots, \\ p_{ii}^{(n)} &\neq 0 \quad \text{for} \quad n = 3, 6, 9, \ldots, \end{aligned}$$

which means that all the states are period 3.

(c) **Persistent State:** Let $f_j^{(n)}$ be the probability that the first return or visit to E_j occurs at the n-th step. This probability is not the same as $p_{jj}^{(n)}$ which is the probability that a return occurs at the n-th step, and includes possible returns at steps 1, 2, 3, ..., $n-1$ also. It follows that

$$p_{jj}^{(1)} (= p_{jj}) = f_j^{(1)}, \tag{17.5.2}$$

$$p_{jj}^{(2)} = f_j^{(2)} + f_j^{(1)} p_{jj}^{(1)}, \tag{17.5.3}$$

$$p_{jj}^{(3)} = f_j^{(3)} + f_j^{(1)} p_{jj}^{(2)} + f_j^{(2)} p_{jj}^{(1)}, \tag{17.5.4}$$

and, in general,

$$p_{jj}^{(n)} = f_j^{(n)} + \sum_{r=1}^{n-1} f_j^{(r)} p_{jj}^{(n-r)} \quad (n \geq 2). \tag{17.5.5}$$

The terms in Eqn.(17.5.4) imply that the probability of a return at the third step is the probability of a first return at the third step, or the probability of a first return at the first step and a return two steps later, or the probability of a first return at the second step and a return one step later.

Equations (17.5.2) and (17.5.5) become iterative formulas for the sequence of first returns $f_j^{(n)}$ which can be expressed as:

$$f_j^{(1)} = p_{jj}, \tag{17.5.6}$$

$$f_j^{(n)} = p_{jj}^{(n)} - \sum_{r=1}^{n-1} f_j^{(r)} p_{jj}^{(n-r)} \quad (n \geq 2). \tag{17.5.7}$$

The probability that a chain returns at some step to the state E_j is

$$f_j = \sum_{n=1}^{j} f_j^{(n)}.$$

If $f_j = 1$, then a return to E_j is certain, and E_j is called a persistent state.

Example 17.5.2. A three-state Markov chain has the transition matrix

$$T = \begin{bmatrix} p & 1-p & 0 \\ 0 & 0 & 1 \\ 1-q & 0 & q \end{bmatrix}$$

where $0 < p < 1,\ 0 < q < 1$. Show that the state E_1 is persistent.

Solution: For simple chains a direct approach using the transition diagram is often easier than the formula (17.5.7) for $f_j^{(n)}$. For this example the transition diagram is shown in Fig. 17.5.2.

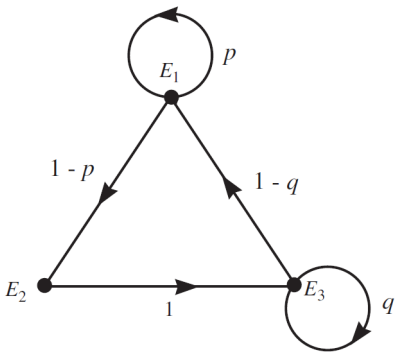


Figure 17.5.2: The transition diagram for Example 17.5.2

If a sequence starts in E_1, then it can be seen that first returns to E_1 can be made to E_1 at every step except for $n = 2$, since after two steps the chain must be in state E_3. From the figure it can be argued that

$$f_1^{(1)} = p, \quad f_1^{(2)} = 0, \quad f_1^{(3)} = (1-p)\cdot 1\cdot(1-q),$$
$$f_1^{(n)} = (1-p)\cdot 1\cdot q^{n-3}\cdot(1-q), \quad (n \geq 4).$$

The last result for $f_1^{(n)}$ for $n \geq 4$ follows from the following sequence of transitions:

$$E_1 E_2 \underbrace{E_3 E_3 \cdots E_3}_{(n-3) \text{ times}} E_1.$$

The probability f_1 that the system returns at least once to E_1 is

$$\begin{aligned} f_1 = \sum_{n=1}^{\infty} = f_1^{(n)} &= p + \sum_{n=3}^{\infty}(1-p)(1-q)q^{n-3} \\ &= p + (1-p)(1-q)\sum_{s=0}^{\infty} q^s \quad (s = n-3) \\ &= p + (1-p)\frac{(1-q)}{(1-q)} \\ &= 1 \end{aligned}$$

Hence, $f_1 = 1$, and consequently the state E_1 is persistent.

The mean recurrence time μ_j of a persistent state E_j, for which $\sum_{n=1}^{\infty} f_j^{(n)} = 1$, is given by

$$\mu_j = \sum_{n=1}^{\infty} n f_j^{(n)}. \tag{17.5.8}$$

In Example 17.5.2, the state E_1 is persistent and its mean recurrence time is given by

$$\begin{aligned}\mu_1 = \sum_{n=1}^{\infty} n f_1^{(n)} &= p + (1-p)(1-q)\sum_{n=3}^{\infty} nq^{n-3} \\ &= p + (1-p)(1-q)\left[\frac{3-2q}{(1-q)^2}\right] \\ &= \frac{3-2p-2q+pq}{1-q}\end{aligned}$$

which is finite. For some chains, however, the mean recurrence time can be infinite; in other words, the mean number of steps to a first return is unbounded.

A persistent state E_j is said to be null if $\mu_j = \infty$ and nonnull if $\mu_j < \infty$.

Example 17.5.3. A three-state inhomogeneous Markov chain has the transition matrix

$$T_n = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 1/(n+1) & 0 & n/(n+1) \end{bmatrix}$$

where T_n is transition matrix at step n. Show that E_1 is a persistent null state.

Solution: The transition diagram at a general step n is shown in Fig. 17.5.3 From the figure, we have

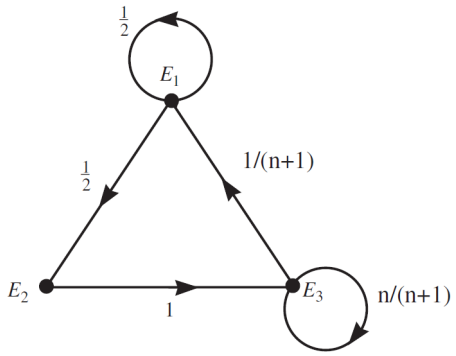


Figure 17.5.3: The transition diagram for Example 17.5.3

.

$$\begin{aligned} f_1^{(1)} &= \frac{1}{2}, \quad f_1^{(2)} = 0, \quad f_1^{(3)} = \frac{1}{2}\cdot 1 \cdot \frac{1}{4}, \\ f_1^{(n)} &= \frac{1}{2}\cdot 1 \cdot \frac{3}{4}\cdot\frac{4}{5}\cdots\frac{n-1}{n}\cdot\frac{1}{n+1} = \frac{3}{2n(n+1),} \quad (n \geq 4). \end{aligned}$$

Hence,

$$f_1 = \frac{1}{2} + \frac{1}{8} + \frac{3}{2}\sum_{n=4}^{\infty}\frac{1}{n(n+1)}.$$

Since

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1},$$

it follows that

$$\sum_{n=4}^{\infty} \frac{1}{n(n+1)} = \lim_{N\to\infty} \sum_{n=4}^{N} \left(\frac{1}{n} - \frac{1}{n+1} \right) = \lim_{N\to\infty} \left(\frac{1}{4} - \frac{1}{N+1} \right) = \frac{1}{4}.$$

Hence

$$f_1 = \frac{5}{8} + \frac{3}{8} = 1,$$

which means E_1 is persistent. On the other hand, the mean recurrence time

$$\begin{aligned} \mu_j = \sum_{n=1}^{\infty} n f_1^{(n)} &= \frac{7}{8} + \frac{3}{2} \sum_{n=4}^{\infty} \frac{n}{n(n+1)} \\ &= \frac{7}{8} + \frac{3}{2} \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \cdots \right) \\ &= \frac{7}{8} + \frac{3}{2} \sum_{n=5}^{\infty} \frac{1}{n}. \end{aligned}$$

The series in the previous equation is the harmonic series

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots, \tag{17.5.9}$$

minus the first four terms. The harmonic series is a well-known divergent series, which means that $\mu_1 = \infty$. Hence E_1 is persistent and null.

(d) **Transient state:** For a persistent state the probability of a first return at some step in the future is certain. For some states,

$$f_j = \sum_{n=1}^{\infty} f_j^{(n)} < 1, \tag{17.5.10}$$

which means that the probability of a first return is not certain. Such states are described as transient.

Example 17.5.4. A four state Markov chain has the transition matrix

$$T = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Show that E_1 is a transient state.

Solution: The transition diagram is shown in Fig. 17.5.4. From the figure

$$f_1^{(1)} = 0, \quad f_1^{(2)} = \frac{1}{2} \cdot \frac{1}{2} = \left(\frac{1}{2} \right)^2, \quad f_1^{(3)} = \left(\frac{1}{2} \right)^3, \quad f_1^{(n)} = \left(\frac{1}{2} \right)^n.$$

Hence

$$f_1 = \sum_{n=1}^{\infty} f_1^{(n)} = \sum_{n=2}^{\infty} \left(\frac{1}{2} \right)^n = \frac{1}{2} < 1$$

implying that E_1 is a transient state. The reason for the transience of E_1 can be seen from Fig. 17.5.4, where transitions from E_3 or E_4 to E_1 or E_2 are not possible.

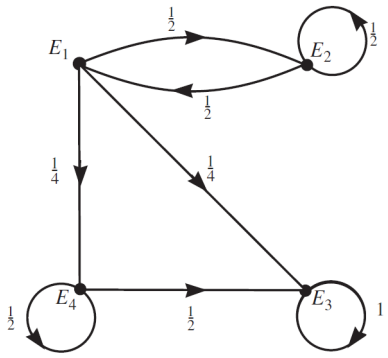


Figure 17.5.4: The transition diagram for Example 17.5.4

(d) **Ergodic states:** A state which is persistent, nonnull and aperiodic is called ergodic state.

Example 17.5.5. A three-state Markov chain has the transition matrix

$$T = \begin{bmatrix} p & 1-p & 0 \\ 0 & 0 & 1 \\ 1-q & 0 & q \end{bmatrix}$$

where $0 < p < 1,\ 0 < q < 1$. Show that the state E_1 is ergodic.

Solution: It was already shown in Example 17.5.2, that E_1 is persistent with

$$f_1^{(1)} = p, \quad f_1^{(2)} = 0, \quad f_1^{(n)} = (1-p)(1-q)q^{n-3}, \ \ (n \geq 3).$$

It follows that its mean recurrence time is

$$\mu_1 = \sum_{n=1}^{\infty} n f_1^{(1)} = p + (1-p)(1-q)\sum_{n=3}^{\infty} nq^{n-3} = \frac{3-2q}{(1-q)^2} < \infty.$$

The convergence of μ_1 implies that E_1 is nonnull. Also the diagonal elements $p_{ii}^{(n)} > 0$ for $n \geq 3$ and $i = 1, 2, 3$, which means that E_1 is aperiodic. Hence from the definition above E_1 (and E_2 and E_3 also) is ergodic.

Unit 18

Course Structure

- Statistical Inference
- Estimation of Parameters
- Minimum Variance Unbiased Estimator
- Method of Maximum Likelihood for Estimation of a parameter

18.1 Introduction

To study the features of any population we first select a sample from the population. A carefully selected sample may be expected to possess the characteristics of the population. A scientific theory developed to get an idea regarding the properties of a population on the basic of the knowledge of the properties of a sample drawn from it is known as *Statistical Inference*.

Statistical Inference may be classified into two main categories :

(i) Problems of Estimation.

(ii) Problems of Testing of Hypothesis or Testing of Significance.

18.2 Estimation of Parameters

Let the distribution function of a population contains one or more unknown parameters and let our task is to make a guess about them on the basis of a sample. The theory regarding this is called *theory of estimation.* In particular, let $x_1, x_2, \ldots, x_n$ be n samples drawn from a population whose distribution has an unknown parameter θ. The problem to replace this unknown θ by a suitable statistic (i.e., a function of the sample values) $\hat{\theta}(x_1, x_2, \ldots, x_n)$ is the problem of estimation.

There are two types of estimation:

(i) Point Estimation, and (ii) Interval Estimation.

In the case of point estimation, the value of θ may vary from sample to sample and this function is known as *'estimator'* of the parameter and its value for a particular sample is called and *'estimate'*.

In the case of interval estimation, two statistics $\hat{\theta}_1(x_1, x_2, \ldots, x_n)$ and $\hat{\theta}_2(x_1, x_2, \ldots, x_n)$ are selected within which the value of the parameter θ is expected to lie. This interval is known as Confidence Interval and the two quantities used to specify the interval are known as Confidence Limits.

According to R. A. Fisher, a good estimator must have the following characteristics:

(i) Unbiasedness,

(ii) Consistency,

(iii) Efficiency,

(iv) Sufficiency.

18.3 Unbiasedness

A statistic T is said to be an unbiased estimator of a parameter θ if the expected value of the statistic coincides with the actual value of the parameter, i.e., if

$$E(T) = \theta$$

Otherwise, the estimation will be called biased. $E(T) - \theta$ is called the bias of the statistic T in estimating θ. It will be called positively or negatively biased according as $E(T) - \theta$ is greater or less than zero.

Theorem 18.3.1. The sample mean is an unbiased estimate of the population mean.

Proof. Let $x_1, x_2, \ldots, x_n$ be n simple samples drawn from a finite population $X_1, X_2, \ldots, X_N$ with replacement. In this case, each x_i have equal chance to be selected from any of the N population values. Therefore,

$$\begin{aligned} E(x_i) &= \frac{1}{N}X_1 + \frac{1}{N}X_2 + \ldots + \frac{1}{N}X_N \\ &= \frac{1}{N}(X_1 + X_2 + \ldots + X_N) \\ &= m, \quad \text{the population mean,} \quad i = 1, 2, \ldots, n \end{aligned} \tag{18.3.1}$$

Again,

$$\begin{aligned} \overline{x} &= \text{sample mean} \\ &= \frac{x_1 + x_2 + \ldots + x_n}{n} \end{aligned}$$

Now,

$$\begin{aligned} E(\overline{x}) &= E\left(\frac{x_1+x_2+\ldots+x_n}{n}\right) \\ &= \frac{1}{n}\Big[E(x_1)+E(x_2)+\ldots+E(x_n)\Big] \\ &= \frac{1}{n}\Big[m+m+\ldots+m\Big] \\ &= \frac{nm}{n} \\ &= m = \text{ population mean} \end{aligned} \tag{18.3.2}$$

Therefore, the sample mean $\overline{x}$ is an unbiased estimate of the population mean m. □

Theorem 18.3.2. The sample variance is a biased estimator of the population variance.

Proof. Let m and σ^2 be the population mean and variance respectively and let $\overline{x}$ and S^2 be the corresponding sample mean and variance.

It is easy to note that $E(x_i) = m$ and $Var(x_i) = E\{(x_i - m)^2\} = \sigma^2$ for $i = 1, 2, \ldots, n$.

Again, Sample mean $\overline{x} = \dfrac{x_1+x_2+\ldots+x_n}{n}$ and sample variance

$$\begin{aligned} S^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-m+m-\overline{x})^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\{(x_i-m)^2 - 2(x_i-m)(\overline{x}-m)+(\overline{x}-m)^2\} \\ &= \frac{1}{n}\left[\sum_{i=1}^{n}(x_i-m)^2 - 2(\overline{x}-m)\sum_{i=1}^{n}(x_i-m)+n(\overline{x}-m)^2\right] \\ &= \frac{1}{n}\left[\sum_{i=1}^{n}(x_i-m)^2 - 2(\overline{x}-m)\left(\sum_{i=1}^{n}x_i-nm\right)+n(\overline{x}-m)^2\right] \\ &= \frac{1}{n}\left[\sum_{i=1}^{n}(x_i-m)^2 - 2(\overline{x}-m)(n\overline{x}-nm)+n(\overline{x}-m)^2\right] \\ &= \frac{1}{n}\left[\sum_{i=1}^{n}(x_i-m)^2\right] - (\overline{x}-m)^2. \end{aligned}$$

Therefore,

$$\begin{aligned} E(S^2) &= E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2\right\} - E\{(\overline{x} - m)^2\} \\ &= \frac{1}{n}\sum_{i=1}^{n} E\{(x_i - m)^2\} - Var(\overline{x}) \\ &= \frac{\sum_{i=1}^{n}\sigma^2}{n} - \frac{\sigma^2}{n} \\ &= \sigma^2 - \frac{\sigma^2}{n} \\ &= \frac{n-1}{n}\sigma^2 \end{aligned}$$

Since, $E(S^2) \neq \sigma^2$, so S^2 is not an unbiased estimate of σ^2. Again,

$$\begin{aligned} \text{bias} &= E(S^2) - \sigma^2 \\ &= \frac{n-1}{n}\sigma^2 - \sigma^2 \\ &= -\frac{\sigma^2}{n}. \end{aligned}$$

Again, if we write

$$s^2 = \frac{n}{n-1}S^2, \tag{18.3.3}$$

$$\begin{aligned} \text{then} \quad E(s^2) &= \frac{n}{n-1}E(S^2) \\ &= \frac{n}{n-1}\cdot\frac{n-1}{n}\sigma^2 \\ &= \sigma^2. \end{aligned}$$

Thus s^2 as defined by (18.3.3) is an unbiased estimate of σ^2. □

18.4 Minimum-Variance Unbiased (M.V.U.) Estimator

Among all the unbiased estimators the minimum-variance unbiased estimator will be that one which has the minimum variance. Thus, if T_m be the minimum-variance unbiased estimator of any parameter θ, then $E(T_m) = \theta$ and Var$(T_m) <$ Var(T), where T is any other unbiased estimator of θ, i.e., $E(T) = \theta$.

18.4.1 Consistent Estimator:

A statistic T_n computed from a sample of n observations is said to be a consistent estimator of a population parameter θ if

$$T_n \xrightarrow[\text{in P}]{} \theta \quad \text{as} \quad n \to \infty. \tag{18.4.1}$$

In other notation,

$$\lim_{n\to\infty} P(|T_n - \theta| < \epsilon) = 1 \tag{18.4.2}$$

or its equivalent,

$$\lim_{n\to\infty} P(|T_n - \theta| \geq \epsilon) = 0 \tag{18.4.3}$$

Thus, a consistent estimator is expected to come more closer to the parameter as the size of the sample becomes larger.

It may be shown that two sufficient conditions for an estimator T_n to be consistent estimator of θ are

$$\text{(i) } E(T_n) \to \theta \text{ and (ii) Var } (T_n) \to 0 \text{ as } n \to \infty.$$

18.5 Efficient Estimator

Among all consistent estimators that one which has minimum asymptotic variance is called the most efficient estimator. Thus a consistent estimator T_n' is said to be most efficient estimator if its sampling variance is less than that of any other consistent estimator T_n, i.e., in this case

$$\text{Var}(T_n') < \text{Var}(T_n).$$

If V_m be the variance of the most efficient estimator and V be the variance of another estimator for a parameter θ, then the efficiency of the estimator is defined as

$$\text{Efficiency} = \frac{V_m}{V}.$$

Since, $V_m \leq V$, so efficiency cannot exceed 1.

18.6 Sufficient Estimator

A statistic T is said to be a sufficient estimator for a parameter θ if it contains all information in the sample about θ. In this case, the joint distribution of the sample can be expressed as the product of two factors, one of which is the sampling distribution of T and contains θ, but the other factor is independent of θ.

Thus for a random sample $x_1, x_2, \ldots, x_n$ from a population whose probability density function (p.m.f) is $f(x, \theta)$ if T be a sufficient estimator of θ, then

$$f(x_1, \theta) \cdot f(x_2, \theta) \ldots f(x_{n_1}, \theta) = f_1(T, \theta) \cdot f_2(x_1, x_2, \ldots, x_n)$$

where $f_1(T, \theta)$ is the sampling of T and contains θ, but $f_2(x_1, x_2, \ldots, x_n)$ is independent of θ.

18.7 Method of Maximum Likelihood for Estimation of a parameters

There are many methods generally used for estimation of parameters of a distribution. Among these, the Method of Maximum Liklihood is one of the most familiar methods.

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n drawn from a population and let $\theta_1, \theta_2, \ldots, \theta_k$ be k parameters of the distribution. This event can be denoted by $(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$ and the probability of this is clearly a function of sample values $x_1, x_2, \ldots, x_n$ and the parameters $\theta_1, \theta_2, \ldots, \theta_k$. This function is known as likelihood function of the sample and it is generally denoted by $L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_k)$. Thus,

$$L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_k) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

Since $X_1, X_2, \ldots, X_n$ are mutually independent random variables each having the distribution of the population, then in the discrete case

$$P(X = x_i) = f_{x_i}(\theta_1, \theta_2, \ldots, \theta_k)$$

and in the continuous case

$$P(X = x_i) = f(x_i, \theta_1, \theta_2, \ldots, \theta_k).$$

Then the liklihood function L in the two cases are given as

$$\begin{aligned} L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_n) &= P(X_1 = x_1)P(X_2 = x_2)\ldots P(X_n = x_n) \\ \text{(In discrete case)} &= f_{x_1}(\theta_1, \theta_2, \ldots, \theta_k) f_{x_2}(\theta_1, \theta_2, \ldots, \theta_k)\ldots f_{x_n}(\theta_1, \theta_2, \ldots, \theta_k) \\ \text{(In continuous case)} &= f(x_1, \theta_1, \theta_2, \ldots, \theta_k) f(x_2, \theta_1, \theta_2, \ldots, \theta_k)\ldots f(x_n, \theta_1, \theta_2, \ldots, \theta_k) \end{aligned}$$

Now, this method states that regarding the sample values as fixed, we shall try to find the values of $\theta_1, \theta_2, \ldots, \theta_k$ such that for these values the likelihood function L will be maximised. Since $L > 0$, so when L is maximum, then $\log L$ is also maximum. The corresponding equations for determining $\theta_1, \theta_2, \ldots, \theta_k$ are

$$\frac{\partial \log L}{\partial \theta_1} = 0, \quad \frac{\partial \log L}{\partial \theta_2} = 0, \quad \ldots \quad , \frac{\partial \log L}{\partial \theta_k} = 0,$$

which are called *likelihood equations.* Solving these k equations we get likelihood estimates of $\theta_1, \theta_2, \ldots, \theta_k$ and they are generally denoted by

$$\theta_1 = \hat{\theta_1}(x_1, x_2, \ldots, x_n), \theta_2 = \hat{\theta_2}(x_1, x_2, \ldots, x_n), \ldots, \theta_k = \hat{\theta_k}(x_1, x_2, \ldots, x_n).$$

Also it may tested that for these values of $\hat{\theta_1}, \hat{\theta_2}, \ldots, \hat{\theta_k}$, L is maximum.

Example 18.7.1. Let T_1 and T_2 be two estimators of the parameter θ. Under what condition $aT_1 + bT_2$ will be an unbiased estimator of θ?

Solution: Since T_1 and T_2 are two unbiased estimator of θ, so $E(T_1) = E(T_2) = \theta$. Again, if $(aT_1 + bT_2)$ be an unbiased estimator of θ, then

$$\begin{aligned} & E(aT_1 + bT_2) = \theta \\ \Rightarrow\ & aE(T_1) + bE(T_2) = \theta \\ \Rightarrow\ & a\theta + b\theta = \theta \\ \Rightarrow\ & a + b = 1, \quad \text{which is the required condition.} \end{aligned}$$

Example 18.7.2. If $X_1, X_2, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$ population, show that the estimator

$$T = \frac{1}{n+1}\sum_{i=1}^{n} X_i$$

is a biased but consistent for μ. Hence obtain the unbiased estimator for μ.

Solution: We have

$$\begin{aligned} T &= \frac{1}{n+1}\sum_{i=1}^{n} X_i \\ &= \frac{n}{n+1} \cdot \frac{1}{n}\sum_{i=1}^{n} X_i \\ &= \frac{n}{n+1}\overline{X} \end{aligned}$$

We know, $X \xrightarrow{\text{in p}}$ as $n \to \infty$ and $\dfrac{n}{n+1} \to 1$ as $n \to \infty$. So

$$T \xrightarrow{\text{in p}} \mu \quad \text{as} \quad n \to \infty.$$

Thus, T is a consistent estimator of μ. Again

$$E(T) = \frac{1}{n+1}\sum_{i=1}^{n} E(X_i) = \frac{1}{n+1} n\mu = \frac{n}{n+1}\mu \quad (\neq \mu)$$

So T is a biased estimator of μ. If we put $T_1 = \dfrac{n+1}{n}$, then, $E(T_1) = \dfrac{n+1}{n}E(T) = \mu$.
Thus, $T_1 = \dfrac{n+1}{n}T$ is the unbiased estimator for μ.

Example 18.7.3. Maximum likelihood estimate of the parameter p of the Binomial (N, p) population for n sample values.

Solution: For Binomial (N, P) population, the density function is given by

$$f_{x_i} = {}^{N}C_{x_i} p^{x_i}(1-p)^{N-x_i}, \quad i = 0, 1, 2, \ldots, n.$$

Now, the likelihood function L is given by

$$\begin{aligned} L &= f_{x_1} \cdot f_{x_2} \cdots f_{x_n} \\ &= {}^{N}C_{x_1} p^{x_1}(1-p)^{N-x_1} \cdot {}^{N}C_{x_2} p^{x_2}(1-p)^{N-x_2} \cdots {}^{N}C_{x_n} p^{x_n}(1-p)^{N-x_n} \\ &= {}^{N}C_{x_1}\, {}^{N}C_{x_2} \cdots {}^{N}C_{x_n} p^{x_1+x_2+\cdots+x_n}(1-p)^{nN-(x_1+x_2+\ldots+x_n)} \end{aligned}$$

So,

$$\log L = (x_1 + x_2 + \ldots + x_n)\log p + [nN - (x_1 + x_2 + \ldots + x_n)]\log(1-p) + \text{terms independent of } p.$$

Now,

$$\frac{\partial \log L}{\partial p} = \frac{n\overline{x}}{p} = \frac{nN - n\overline{x}}{1-p} \quad \left[\because \overline{x} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)\right]$$

Thus, $\dfrac{\partial \log L}{\partial p} = 0$ gives

$$\begin{aligned} \frac{n\overline{x}}{p} &= \frac{nN - n\overline{x}}{1-p} \\ \Rightarrow p &= \frac{\overline{x}}{N}. \end{aligned}$$

Thus, $\hat{p} = \dfrac{\overline{x}}{N}$ is the likelihood estimate of p.

It can be verified that

$$\left[\frac{\partial^2 L}{\partial p}\right]_{p=\hat{p}} < 0.$$

Example 18.7.4. Maximum likelihood estimator of the parameter of a Poisson distribution.

Solution: Let $x_1, x_2, \ldots, x_n$ be n sample values drawn from a Poisson distribution having parameter μ. Then

$$f(x, \mu) = \frac{e^{-\mu}\mu^x}{x!} \quad (x = 0, 1, 2, \ldots, \infty)$$

The likelihood function L of the sample observations is given by

$$\begin{aligned} L &= f(x_1, \mu) \cdot f(x_2, \mu) \cdots f(x_n, \mu) \\ &= \frac{e^{-\mu}\mu^{x_1}}{x_1!} \cdot \frac{e^{-\mu}\mu^{x_2}}{x_2!} \cdots \frac{e^{-\mu}\mu^{x_n}}{x_n!} \\ &= \frac{e^{-n\mu}\mu^{x_1+x_2+\ldots+x_n}}{(x_1!)(x_2!)\ldots(x_n!)} \end{aligned} \tag{18.7.1}$$

So

$$\begin{aligned} \log L &= \log(e^{-n\mu}) + \log(\mu^{x_1+x_2+\ldots+x_n}) - \log(x_1!\, x_2!\, \cdots x_n!) \\ &= -n\mu + \left(\sum_{i=1}^{n} x_i\right) \log \mu - \sum_{i=1}^{n} \log(x_i!). \end{aligned}$$

If $\hat{\mu}$ be the likelihood estimator of μ, then it will be given by

$$\left[\frac{\partial \log L}{\partial \mu}\right]_{\mu=\hat{\mu}} = 0 \quad \text{and} \quad \left[\frac{\partial^2 \log L}{\partial \mu^2}\right]_{\mu=\hat{\mu}} < 0.$$

From above, we have

$$\frac{\partial \log L}{\partial \mu} = -n + \frac{1}{\mu} + \sum_{i=1}^{n} x_i \quad \text{and} \quad \frac{\partial^2 \log L}{\partial \mu^2} = -\frac{1}{\mu^2} + \sum_{i=1}^{n} x_i$$

Now, $\dfrac{\partial \log L}{\partial \mu} = 0$ gives,

$$\begin{aligned} -n + \frac{1}{\mu}\sum_{i=1}^{n} x_i &= 0 \\ \Rightarrow \mu = \frac{1}{n}\sum_{i=1}^{n} x_i &= \overline{x}. \end{aligned}$$

Therefore, $\overline{\mu} = \overline{x}$. Again

$$\left[\frac{\partial^2 \log L}{\partial \mu^2}\right]_{\mu=\hat{\mu}} = -\frac{1}{\hat{\mu}^2}\sum_{i=1}^{n} x_i = -\frac{n\overline{x}}{(\overline{x})^2} = -\frac{n}{\overline{x}} < 0.$$

Thus, $\hat{\mu} = \overline{x}$, the sample means is the likelihood estimate of the parameter μ of a Poisson distribution.

Example 18.7.5. Maximum likelihood estimates of the parameter m and σ in Normal (m, σ) population for a sample of size n.

Solution: We know for a Normal (m, σ) distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2}\sigma^2}, \quad -\infty < x < \infty.$$

So, the likelihood function L is given by

$$\begin{aligned} L &= f(x_1)f(x_2)\cdots f(x_n) \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-m)^2}{2\sigma^2}}. \end{aligned}$$

Then,

$$\log L = -n\log(\sqrt{2\pi}) - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-m)^2.$$

Then, $\dfrac{\partial L}{\partial m} = 0$ gives $\displaystyle\sum_{i=1}^{n}(x_i-m) = 0 \Rightarrow m = \frac{\sum_{i=1}^{n}x_i}{n} \Rightarrow \hat{m} = \overline{x}.$

Also, $\dfrac{\partial \log L}{\partial \sigma} = 0$ gives

$$\begin{aligned} -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-m)^2 &= 0 \\ \Rightarrow \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-m)^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2 = S^2. \end{aligned}$$

Thus, $\hat{\sigma}^2 = S^2$.

Example 18.7.6. Find the maximum likelihood estimate of the parameter λ for the Weibuzl distribution

$$f(x) = \lambda\alpha x^{\alpha-1}e^{-\lambda x^\alpha}, \quad (x > 0)$$

using a sample of size n assuming that α is known.

Solution: If $x_1, x_2, \ldots, x_n$ be n sample values, then the maximum likelihood function L is given by

$$L = \lambda^n\alpha^n(x_1, x_2, \ldots, x_n)^{n-1}e^{-\lambda(x_1^\alpha + x_2^\alpha + \ldots + x_n^\alpha)}$$

Then,

$$\log L = n\log\lambda - \lambda(x_1^\alpha + x_2^\alpha + \ldots + x_n^\alpha) \ + \text{ terms independent of } \lambda.$$

So, the likelihood equation $\dfrac{\partial \log L}{\partial \lambda} = 0$ gives

$$\begin{aligned} \frac{n}{\lambda} - (x_1^\alpha + x_2^\alpha + \ldots + x_n^\alpha) &= 0 \\ \Rightarrow \hat{\lambda} = \frac{n}{\sum_{i=1}^{n}x_i^\alpha} & \end{aligned}$$

Again $\left[\dfrac{\partial^2 \log L}{\partial \lambda^2}\right]_{\lambda=\hat{\lambda}} = -\dfrac{n}{\lambda^2}$. So for this $\hat{\lambda}$, L is maximum.

Example 18.7.7. Prove that the maximum likelihood estimate of the parameter α of a population having density function

$$\frac{2}{\alpha^2}(\alpha - x), \quad 0 < x < \alpha,$$

for a sample of unit size is $2x$, x being the sample value. Show also that the estimate is biased.

Solution: Since the sample is of unit size, so the likelihood function L is given by

$$\begin{aligned} L &= \frac{2}{\alpha^2}(\alpha - x) \\ \Rightarrow \log L &= \log 2 - 2\log\alpha + \log(\alpha - x) \end{aligned}$$

Now, the likelihood equation $\dfrac{\partial \log L}{\partial \alpha} = 0$ gives

$$\begin{aligned} -\frac{2}{\alpha} + \frac{1}{\alpha - x} &= 0 \\ \Rightarrow \alpha &= 2x \end{aligned}$$

So, $\hat{\alpha} = 2x$. Also it can be shown that

$$\frac{\partial^2 \log L}{\partial \alpha^2} = \frac{2}{\alpha^2} - \frac{1}{(\alpha - x)^2} < 0 \quad \text{for} \quad \hat{\alpha} = 2x.$$

Thus, the maximum likelihood estimate of α is $\hat{\alpha} = 2x$. Again,

$$\begin{aligned} E(2x) &= \int_0^\alpha 2x \cdot \frac{2}{\alpha^2}(\alpha - x)\, dx = \frac{4}{\alpha^2}\int_0^\alpha (\alpha x - x^2)\, dx \\ &= \frac{4}{\alpha^2}\left[\alpha\frac{x^2}{2} - \frac{x^3}{3}\right]_0^\alpha = \frac{4}{\alpha^2}\left[\frac{\alpha^3}{2} - \frac{\alpha^3}{3}\right] \\ &= \frac{2\alpha}{3} \neq \alpha \end{aligned}$$

Thus, $\hat{\alpha}$ is a biased estimate of α.

Unit 19

Course Structure

- Interval estimation
- Method for finding confidence intervals
- Statistical hypothesis
- Level of significance; Power of the test

19.1 Introduction

We have studied the problem of estimation of a parameter occurring in a distribution such an estimate is called parameter estimate and the corresponding problem is known as the problem of estimation. Such and estimate always associated with random error. For this reason, it is sometime desirable to find a $\delta > 0$ for a given small ϵ where $0 < \epsilon < 1$ such that an estimate $\hat{\theta}$ for the parameter θ satisfies

$$P(\hat{\theta} - \delta < \theta < \hat{\theta} + \delta) = 1 - \epsilon$$

19.2 Interval Estimation

Let θ be a population parameter and let T_1 and T_2 be two functions based on sample observations such that

$$P(T_1 \leq \theta \leq T_2) = 1 - \epsilon \tag{19.2.1}$$

where $\epsilon(0 < \epsilon < 1)$ is a parameter. Then the interval (T_1, T_2) is called an interval estimate or a confidence interval for the parameter θ with confidence coefficient $1 - \epsilon$; the statistics T_1 and T_2 are respectively called the lower and upper confidence limits for θ.

A practical interpretation of this result is that if a long sequence of random samples, are drawn from a population under uniform conditions and the statistics T_1 and T_2 are computed in each time, then

$$\frac{\text{The number of times the interval } (T_1, T_2) \text{ includes the true parameter } \theta}{\text{The total number of samples drawn}} = 1 - \epsilon$$

The number ϵ is usually chosen to be very small, like 0.05, 0.01, 0.001 etc. and the corresponding confidence coefficients are 0.95, 0.99, 0.999 etc. and then the corresponding confidence intervals will be called 95%, 99%, 99.9% etc. confidence intervals.

The length of the interval $(T_2 - T_1)$ is used as an inverse measure of precision of the interval estimate.

19.3 Method for finding confidence intervals

To find the confidence interval for a parameter θ, the following steps to be followed.

1. We choose, if possible, a suitable statistic $z = z(x_1, x_2, \ldots, x_n, \theta)$ whose sampling distribution is independent of the parameter θ but which itself depends on θ.

2. Now we choose two numbers $\alpha_\epsilon, \beta_\epsilon (> \alpha_\epsilon)$ such that

$$P(\alpha_\epsilon < z < \beta_\epsilon) = 1 - \epsilon \tag{19.3.1}$$

3. We rewrite the above equation (19.3.1) as

$$P(T_1 < \theta < T_2) = 1 - \epsilon \tag{19.3.2}$$

Then (T_1, T_2) is the desired confidence interval for the population parameter θ.

19.4 Confidence interval for some special cases

(a) The confidence interval for m for a Normal (m, σ) population.
Case 1. σ known: The suitable statistic for this case will be chosen as

$$z = \frac{\overline{x} - m}{\sigma/\sqrt{n}}$$

whose sampling distribution is normal $(0, 1)$ and which depends on the parameter m.

Since normal curve is symmetrical curve, so we take two points $\pm u_\epsilon$ symmetrically about the origin, Fig. 19.4.1, such that

$$\begin{aligned} &P(-u_\epsilon < z < u_\epsilon) = 1 - \epsilon \\ \Rightarrow\ &P\left(-u_\epsilon < \frac{\overline{x} - m}{\sigma/\sqrt{n}} < u_\epsilon\right) = 1 - \epsilon \end{aligned}$$

which can be rewritten as

$$P\left(\overline{x} - \frac{\sigma u_\epsilon}{\sqrt{n}} < m < \overline{x} + \frac{\sigma u_\epsilon}{\sqrt{n}}\right) = 1 - \epsilon.$$

Hence a confidence interval for m having confidence coefficient $1 - \epsilon$ is

$$\left(\overline{x} - \frac{\sigma u_\epsilon}{\sqrt{n}}, \overline{x} + \frac{\sigma u_\epsilon}{\sqrt{n}}\right) \tag{19.4.1}$$

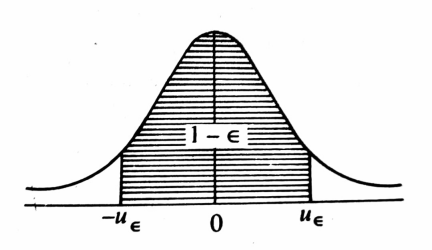


Figure 19.4.1

where u_ϵ is given by $P(-u_\epsilon < z < u_\epsilon) = 1 - \epsilon$ or from symmetry $P(z > u_\epsilon) = \frac{1}{2}\epsilon$. For 95% confidence interval, $1 - \epsilon = 0.95$ and $u_\epsilon = 1.96$, then the corresponding confidence interval for the population mean m will be

$$\left(\overline{x} - 1.96\frac{\sigma}{\sqrt{m}}, \overline{x} + 1.96\frac{\sigma}{\sqrt{m}}\right) \tag{19.4.2}$$

Case II: σ unknown: In this case, the suitable statistic will be

$$t = \frac{\overline{x} - m}{s/\sqrt{n}}, \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

whose sampling distribution has t-distribution with $n - 1$ degrees of freedom.

Now procedding exactly as in the Case I, we can calculate two numbers $\pm t_\epsilon$, Fig. 19.4.2, by

$$P(-t_\epsilon < t < t_\epsilon) = 1 - \epsilon$$

which gives the confidence interval of m as

$$\left(\overline{x} - \frac{st_\epsilon}{\sqrt{n}}, \overline{x} + \frac{st_\epsilon}{\sqrt{n}}\right) \tag{19.4.3}$$

Here t_ϵ is given by $P(-t_\epsilon < t < t_\epsilon) = 1 - \epsilon$ or by $P(t > t_\epsilon) = \frac{1}{2}\epsilon$. In case of large samples, if σ is unknown, then the approximate call interval for m may be obtained by replacing σ by s or S in (19.4.2).

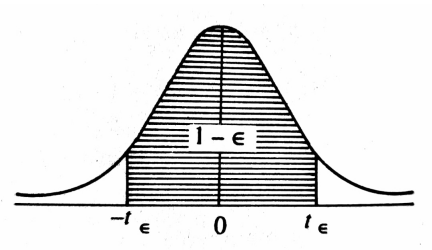


Figure 19.4.2

(b) Confidence interval for σ

It is known that the statistic

$$\chi^2 = \frac{nS^2}{\sigma^2}$$

is a χ^2-distributed with $(n-1)$ degrees of freedom, where S^2 is the sample variance, σ is the population variance and n is the size of the sample.

We choose any positive number $\chi^2_{\epsilon_1}$ and determine $\chi^2_{\epsilon_2}$ such that

$$\begin{aligned} & P(\chi^2_{\epsilon_1} < \chi^2 < \chi^2_{\epsilon_2} = 1-\epsilon \\ \Rightarrow\ & P\left(\chi^2_{\epsilon_1} < \frac{nS^2}{\sigma^2} < \chi^2_{\epsilon_2}\right) = 1-\epsilon \\ \Rightarrow\ & P\left(S\sqrt{\frac{n}{\chi^2_{\epsilon_2}}} < \sigma < S\sqrt{\frac{n}{\chi^2_{\epsilon_1}}}\right) = 1-\epsilon \end{aligned}$$

Therefore, $\left(S\sqrt{\dfrac{n}{\chi^2_{\epsilon_2}}}, S\sqrt{\dfrac{n}{\chi^2_{\epsilon_1}}}\right)$ is the confidence interval for σ having confidence coefficient $1-\epsilon$.

In practice, $\chi^2_{\epsilon_1}$ and $\chi^2_{\epsilon_2}$ are given by

$$P(\chi^2 > \chi^2_{\epsilon_1}) = 1 - \frac{1}{2}\epsilon \quad \text{and} \quad P(\chi^2 > \chi^2_{\epsilon_2}) = \frac{1}{2}\epsilon. \tag{19.4.4}$$

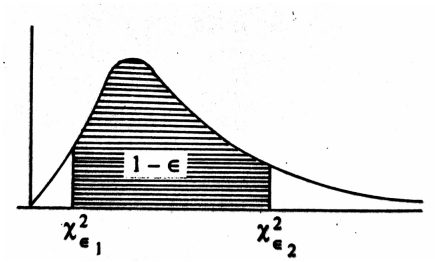


Figure 19.4.3

Example 19.4.1. A sample 2.3, -0.2, -0.4, -0.9 is taken from a normal population with variance 9. Find a 95% confidence interval for the population mean. (Given $P(U) > 1.960) = 0.025$, where U is a normal $(0,1)$ variate.

Solution: With usual notation, we have

$$\overline{x} = \frac{2.3 + (-0.2) + (-0.4) + (-0.9)}{4} = 0.2$$

Also, $n = 4$ and $\sigma^2 = 9,\ \epsilon = 0.05,\ u_\epsilon = 1.96$. Hence, the confidence interval for mean when σ is known is

$$\begin{aligned} & \left(\overline{x} - 1.96\frac{\sigma}{\sqrt{m}}, \overline{x} + 1.96\frac{\sigma}{\sqrt{m}}\right) \\ &= \left(0.2 - 1.96 \times \frac{3}{2},\ \ 0.2 + 1.96 \times \frac{3}{2}\right) \\ &= (-2.74,\ 3.14) \end{aligned}$$

Example 19.4.2. The mean and variance of a sample of size 400 from a normal population are found to be 18.35 and 3.25 respectively. Given $P(U > 1.96) = 0.025$, U being a standard normal variate, find 95% confidence interval for the population mean.

Solution: From the given data

$$\overline{x} = 18.35,\ S^2 = 3.25,\ n = 400,\ \epsilon = 0.05$$

Now

$$s^2 = \frac{n}{n=1} S^2 = \frac{400}{399} \times 3.25 = 3.26]$$

Hence the confidence interval for mean when σ is unknown is

$$\begin{aligned}
&\left(\overline{x} - 1.96\frac{s}{\sqrt{n}}, \overline{x} + 1.96\frac{s}{\sqrt{n}}\right) \\
&= \left(18.35 - 1.96 \times \frac{1.80}{20},\ 18.35 + 1.96 \times \frac{1.80}{20}\right) \\
&= (18.17,\ 18.53)
\end{aligned}$$

Example 19.4.3. Obtain 99% confidence interval of the population standard deviation (σ) on the basis of the data $\sum_{i=1}^{10} x_i = 620$ and $\sum_{i=1}^{10} x_i^2 = 39016$. (It is given that $\chi^2_{0.005,9} = 23.59$ and $\chi^2_{0.995,9} = 1.74$)

Solution:

$$\begin{aligned}
\text{Sample variance } S^2 &= \frac{1}{10}\sum_{i=1}^{10} x_i^2 - \left(\frac{1}{10}\sum_{i=1}^{10} x_i\right)^2 \\
&= \frac{39016}{10} - \left(\frac{620}{10}\right)^2 \\
&= 3901.6 - 3844 \\
&= 57.6
\end{aligned}$$

For 99% confidence interval

$$1 - \epsilon = 0.99,\ i.e.,\ \epsilon = 0.01 \text{ and } \frac{1}{2}\epsilon = 0.005$$

Here, $n = 10$. We know that confidence interval of σ is

$$\begin{aligned}
\left(S\sqrt{\frac{n}{\chi^2_{\epsilon_2}}}, S\sqrt{\frac{n}{\chi^2_{\epsilon_1}}}\right) &= \left(\sqrt{\frac{57.6 \times 10}{23.59}},\ \sqrt{\frac{57.6 \times 10}{1.74}}\right) \\
&= (4.94,\ 18.19)
\end{aligned}$$

19.5 Statistical Hypothesis

To make decision regarding a statistical population on the basis of sample observation is called a *Statistical Hypothesis*. It is an assertion or conjecture about the distribution of one or more random variables.

There are two types of hypothesis, viz. simple and composite. When a statistical hypothesis completely specifies the population distribution, it will be called a simple hypothesis and when it will not completely specify the population distribution, it will be called a composite hypothesis. In the case of composite hypothesis the number of unspecified parameters is called the degrees of freedom of the composite hypothesis.

As an illustration, let us consider a Normal (m, σ) distribution and let m_0 and σ_0 be taken to be two given values of m and σ respectively. Then

(i) Hypothesis $m = m_0$ is simple if σ is known and m is unknown.

(ii) Hypothesis $\sigma = \sigma_0$ is simple if m is known and σ is unknown.

(iii) Hypothesis $m = m_0$ is composite if both m and σ are unknown and its degrees of freedom is 1.

19.6 Null Hypothesis and Alternative Hypothesis

Let a population has only one parameter θ. Then a hypothesis about the parameter θ which we want to test is called Null Hypothesis. This is generally written as

$$H_0 : \theta = \theta_0$$

Any other hypothesis about the parameter θ against which we wish to test the null hypothesis is called *Alternative Hypothesis* and this is written as

$$H_1 : \theta = \theta_1$$

Generally, the hypothesis wishing to be rejected by the test is taken as null hypothesis. Say we have two alternatives i.e., either $\theta = \theta_0$ or $\theta = \theta_1$ and we have a priori reason to be more inclined to believe the second hypothesis, then we take the hypothesis $H_0 : \theta = \theta_0$ as null hypothesis.

19.7 Critical Region

Any sample $x_1, x_2, \ldots, x_n$ of size n may be considered to be a point in n-dimensional space and it will be called a *sample point.* All such sample points corresponding to various random samples of size n constitute a sample space S, Fig. 19.7.1, so every sample is a point in S.

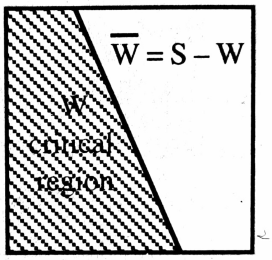


Figure 19.7.1

Let us divide the sample space S into two disjoint parts W and $\overline{W}(= S - W)$. Let us assume that we reject the null hypothesis $H_0 : \theta = \theta_0$ if the observed sample point falls in W and in this case we accept $H_1 : \theta = \theta_1$. On the other hand we accept H_0 if the point fall in $\overline{W}$. Technically the region W, i.e., the region of rejection of the null hypothesis H_0 is called the critical region or region of rejection.

19.8 Two Types of Errors

The decision whether the null hypothesis to be reject or accepted is taken on the basis of the information supplied by the observed sample observations. The conclusion drawn on the basis of a particular sample may

not be always true in respect of the population. The following two cases are called Type I and Type II errors.

Type I Error: When the null hypothesis H_0 is rejected i.e., H_1 is accepted but H_0 is true, the error arising in this situation is called Type I Error. If α be the probability of Type I Error, then

$$\begin{aligned} \alpha &= \text{Probability of Type I Error} \\ &= \text{Probability of rejecting } H_0 \text{ where } H_0 \text{ is true} \\ &= P(x \in W/H_0 = \theta_0), \text{ where } x = (x_1, x_2, \ldots, x_n) \end{aligned} \tag{19.8.1}$$

Type II Error: When the null hypothesis H_0 is accepted i.e., H_1 is rejected but H_0 is false, the error arising in this situation will be called Type II error. If β be the probability of Type II Error, then

$$\begin{aligned} \beta &= \text{Probability of Type II Error} \\ &= \text{Probability of accepting } H_0 \text{ where } H_0 \text{ is false} \\ &= P(x \in \overline{W}/H_1 = \theta_1), \text{ where } x = (x_1, x_2, \ldots, x_n) \end{aligned} \tag{19.8.2}$$

19.9 Level of Significance

The probability of Type I Error, α, is called *level of significance* of the test. It is also called the *size of the critical region.*

19.10 Power of the test

If β be the probability of Type II Error, then $1 - \beta$ is defined as the *power function* of the test hypothesis. The graph obtained by plotting power on the y-axis against various values of the parameter θ on the x-axis on a graph paper is called a *power curve.* The value of the power function at a parameter point is called the *power of the test* at that point.

Example 19.10.1. A random sample of size 10 is taken from a normal population and the following values were calculated for the variable (x) under study:

$$\sum_{i=1}^{10} x_i = 620, \quad \sum_{i=1}^{10} x_i^2 = 39016.$$

Test the null hypothesis $H_0 : \sigma = 8$ against $H_1 : \sigma > 8$ on the basis of the above data. Use $\alpha = 0.05$ as level of significance. (Given $\chi^2_{0.05}$ for 9 degrees of freedom = 16.92)

Solution: Here

$$n = 10, \quad \sum_{i=1}^{10} x_i = 620, \quad \sum_{i=1}^{10} x_i^2 = 39016.$$

Then

$$\begin{aligned} S^2 &= \frac{\sum_{i=1}^{10} x_i^2}{n} - \left(\frac{\sum_{i=1}^{10} x_i}{n} \right)^2 \\ &= \frac{39016}{10} - \left(\frac{620}{10} \right) \\ &= 3901.6 - 3844 \\ &= 57.6 \end{aligned}$$

We make the null hypothesis $H_0 : \sigma = 8$ against $H_1 : \sigma > 8$. Again

$$\begin{aligned}\chi^2 &= \frac{nS^2}{\sigma^2} \\ &= \frac{10 \times 57.6}{(8)^2} \\ &= 9\end{aligned}$$

Since $\chi^2_{observed} = 9 < \chi^2_{0.05,9} = 16.92$, so, H_0 is accepted and thus we conclude that the value of σ may be taken as 8 at 95% level of significance.

Example 19.10.2. For a large lot of freshly minted coins a random sample of size 50 is taken. The mean weight of coins in the sample is found to be 28.57 gm. Assuming that the population standard deviation of weight is 1.25 gm., will it it be reasonable to suppose that the population mean is 28 gm ?

Solution: The size of the sample is 50 and so $n = 50$. Population mean $m = 28$ gm and population s.d. $\sigma = 1.25$ gm. Let the null hypothesis H_0 and the alternative hypothesis H_1 be given by

$$\begin{aligned}&H_0 : m = 28 \\ &H_1 : m \neq 28 \\ &\text{S.E. of } \overline{x} = \frac{\sigma}{\sqrt{n}} = \frac{1.25}{\sqrt{50}} = \frac{1.25}{7.071} = 0.18\end{aligned}$$

Let us consider the statistic

$$z = \frac{\overline{x} - m}{\text{S.E. of } (\overline{x})}$$

which is standard normal. Therefore,

$$z = \frac{28.57 - 28}{0.18} = 3.17$$

Since the observed value of z exceeds 1.64, thus z falls in the critical region at 5% level of significance and so the null hypothesis H_0 is rejected at 5% level of significance. So it will not be reasonable to suppose that the population mean is 28 gm. at 5% level of significance.

Example 19.10.3. The mean life time of a sample of 100 electric bulbs produced by a manufacturing company is estimated to be 1570 hours with a standard deviation of 120 hours. If μ be the mean life time of all the bulbs produced by the company, test the hypothesis $\mu = 1600$ hours against the alternative hypothesis $\mu \neq 1600$ hours, using level of significance 0.05.

Solution: Here $n =$ size of the sample = 100, population mean $\mu = 1570$ and population S.D. $\sigma = 120$. We test the null hypothesis $H_0 = \mu = 1600$ against the alternative hypothesis $H_1 : \mu \neq 1600$ at 5% level of significance.

$$\text{Here } \overline{x} = 1570 \text{ and S.E. of } \overline{x} = \frac{\sigma}{\sqrt{n}} = \frac{120}{\sqrt{100}} = 12$$

Therefore,

$$\begin{aligned}z &= \frac{\overline{x} - \mu}{\text{S.E. of } (\overline{x})} \\ &= \frac{1570 - 1600}{12} \\ &= -2.5\end{aligned}$$

Hence z falls in the critical region at 5% level of significance and so we reject the null hypothesis.

Thus at 5% level of significance it will not be reasonable to suppose that the mean life of the bulb will be 1600 hours.

Example 19.10.4. In a sample of 600 students of a certain college, 400 are found to use dot pens. In another college from a sample of 900 students 450 were found to use dot pens. Test whether the colleges are significantly different with respect to the habit of using dot pens. (Null and alternative hypothesis should be stated clearly.)

Solution: With usual notations, null hypothesis will be that the the population proportions of the two colleges regarding the habit of using dot pen are equal. So $H_0 : (P_1 = P_2)$ and alternative hypothesis is $H_1 : (P_1 \neq P_2)$.

Here,

$$n_1 = 600, \quad p_1 = \frac{400}{600} = 0.667$$
$$n_2 = 900, \quad p_2 = \frac{450}{900} = 0.5$$

If for the null hypothesis $P_1 = P_2 = P$, then sample estimate of P is

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{600 \times 0.667 + 900 \times 0.5}{600 + 900} = 0.567$$

Now,

$$\begin{aligned} \text{S.E. of } (p_1 - p_2) &= \sqrt{pq\left(\frac{1}{p_1} + \frac{1}{p_2}\right)} \\ &= \sqrt{0.567 \times (1 - 0.567) \times \left(\frac{1}{600} + \frac{1}{900}\right)} \\ &= \sqrt{0.567 \times 0.433 \times (0.0017 + 0.0011)} \\ &= 0.026 \end{aligned}$$

Now,

$$z = \frac{p_1 - p_2}{\text{S.E.}} = \frac{0.667 - 0.5}{0.026} = 6.42$$

At 1% level the critical region is $|z| > 2.58$. So this z falls in the critical region and hence H_0 is rejected. So the two colleges are significantly different with respect to the habit of using dot pens.

Unit 20

Course Structure

- Analysis of variance
- One factor experiments
- Linear mathematical model for ANOVA

20.1 Introduction

Suppose that in an agricultural experiment, four different chemical treatments of soil produced mean wheat yields of 28, 22, 18 and 24 bushels per acre, respectively. Is there a significant difference in these means, or is the observed spread simply due to chance?

Such problem can be solved by using an important technique known as the *analysis of variance,* developed by Fisher. It makes use of the F distribution already considered in previous unit. Basically, in many situations there is a need to test the significance of differences among three or more sample means, or equivalently to test the null hypothesis that the sample means are all equal.

20.2 One-Way Classification or One-Factor Experiments

In a *one-factor experiment* measurements or observations are obtained for a independent groups of samples, where the number of measurements in each group is b. We speak of *a treatments,* each of which has *b repetitions* or *replications.* In the above example, $a = 4$.

The results of a one-factor experiment can be presented in a table having a rows and b columns (Table. 16.1). Here x_{jk} denotes the measurement in the j-th row and k-th column, where $j = 1, 2, \ldots, a$ and $k = 1, 2, \ldots, b$. For example, x_{35} refers to the fifth measurement for the third treatment.

Table 16.1

Treatment 1	x_{11}	x_{12}	$\cdots$	x_{1b}	$\overline{x}_1$
Treatment 2	x_{21}	x_{22}	$\cdots$	x_{2b}	$\overline{x}_2$
$\vdots$			$\vdots$		
Treatment a	x_{a1}	x_{a2}	$\cdots$	x_{ab}	$\overline{x}_a$

We shall denote by $\overline{x}_{j\cdot}$ the mean of the measurements in the j-th row. We have

$$\overline{x}_{j\cdot} = \frac{1}{b}\sum_{k=1}^{b} x_{jk}, \quad j = 1, 2, \ldots, a \tag{20.2.1}$$

The dot in $\overline{x}_{j\cdot}$ is used to show that the index k has been summed out. The values $\overline{x}_{j\cdot}$ are called group means or treatment means or row means. The grand mean or overall mean is the mean of all the measurement in all the groups and is denoted by $\overline{x}$, i.e.,

$$\overline{x} = \frac{1}{ab}\sum_{j,k} x_{jk} = \frac{1}{ab}\sum_{j=1}^{a}\sum_{k=1}^{b} x_{jk}. \tag{20.2.2}$$

20.3 Total Variation, Variation Within Treatments, Variation Between Treatments

We define the total variation, denoted by v, as the sum of the squares of the deviations of each measurement from the grand mean $\overline{x}$, i.e.,

$$\text{Total variation} \;=\; v = \sum_{j,k}(x_{jk} - \overline{x}^2). \tag{20.3.1}$$

By writing the identity,

$$x_{jk} - \overline{x} = (x_{jk} - \overline{x}_j) + (\overline{x}_j - \overline{x}) \tag{20.3.2}$$

and then squaring and summing over j and k, we can show that

$$\sum_{j,k}(x_{jk} - \overline{x})^2 = \sum_{j,k}(x_{jk} - \overline{x}_j)^2 + \sum_{j,k}(\overline{x}_j - \overline{x})^2 \tag{20.3.3}$$

$$\Rightarrow \sum_{j,k}(x_{jk} - \overline{x})^2 = \sum_{j,k}(x_{jk} - \overline{x}_j)^2 + b\sum_{j}(\overline{x}_j - \overline{x})^2 \tag{20.3.4}$$

We call the first summation on the right side of (20.3.4) the variation within the treatments (since it involves the squares of the deviations of x_{jk} from the treatment means $\overline{x}_j$) and denoted it by v_w. Therefore,

$$v_w = \sum_{j,k}(x_{jk} - \overline{x}_j)^2 \tag{20.3.5}$$

The second summation on the right side of (20.3.4) is called the variation between treatments (since it involves the squares of the deviation of the various treatment means $\overline{x}_j$ from the grand mean $\overline{x}$ and is denoted by v_b). Therefore,

$$v_b = \sum_{j,k}(\overline{x}_j - \overline{x})^2 = b\sum_{j}(\overline{x}_j - \overline{x})^2 \tag{20.3.6}$$

Equation (20.3.4) can thus be written as

$$v = v_w + v_b. \tag{20.3.7}$$

20.4 Shortcut Methods for Obtaining Variations

To minimize the labour in computing the above variations, the following forms are convenient:

$$v = \sum_{j,k} x_{jk}^2 - \frac{\tau^2}{ab} \tag{20.4.1}$$

$$v_b = \frac{1}{b}\sum_{j} \tau_j^2 - \frac{\tau^2}{ab} \tag{20.4.2}$$

$$v_w = v - v_b \tag{20.4.3}$$

where τ is the total of all values x_{jk} and $\tau_{j\cdot}$ is the total of all values in the j-th treatment, i.e.,

$$\tau = \sum_{j,k} x_{jk} \quad \tau_{j\cdot} = \sum_{k} x_{jk} \tag{20.4.4}$$

In practice it is convenient to subtract some fixed value from all the data in the table; this has no effect on the final results.

20.5 Linear Mathematical Model for Analysis of Variance

We can consider each row of Table 16.1 as a random sample of size b from the population fro that particular treatment. Therefore, for treatment j we have the independent, identically distributed random variables $X_{j1}, X_{j2}, \ldots, X_{jb}$, which respectively, take on the values $x_{j1}, x_{j2}, \ldots, x_{jb}$. Each of the $X_{jk}(k = 1, 2, \ldots, b)$ can be expressed as the sum of its expected value and a "chance" or "error" term:

$$X_{jk} = \mu_j + \Delta_{jk} \tag{20.5.1}$$

The Δ_{jk} can be taken as independent (relative to j as well as to k), normally distributed random variables with mean zero and variance σ^2. This is equivalent to assuming the the $X_{jk}(j = 1, 2, \ldots, a;\ k = 1, 2, \ldots, b)$ are mutually independent, normal variables with means μ_j and common variance σ^2. Let us define the constant μ by

$$\mu = \frac{1}{a}\sum_{j} \mu_j$$

We can think of μ as the mean for a sort of grand population comprising all the treatment populations. Then (20.5.1) can be rewritten as

$$X_{jk} = \mu + \alpha_j + \Delta_{jk} \quad \text{where} \quad \sum_{j} \alpha_j = 0 \tag{20.5.2}$$

The constant α_j can be viewed as the special effect of the j-th treatment.
The null hypothesis that all treatment means are equal is given by $(H_0 : \alpha_j = 0; j = 1, 2, \ldots, a)$ or equivalently by $(H_0 = \mu_j = \mu;\ j = 1, 2, \ldots, a)$. If H_0 is true, the treatment populations, which by assumption are normal, have a common mean as well as a common variance. Then there is just one treatment population, and all treatments are statistically identical.

20.6 Expected Values of the Variations

The between-treatments variation V_b, the within-treatments variation V_w, and the total variation V are random variables that, respectively, assume the values v_b, v_w, and v as defined in (20.3.6), (20.3.5) and (20.3.1), we can show that

$$E(V_b) = (a-1)\sigma^2 + b\sum_j \alpha_j^2 \tag{20.6.1}$$

$$E(V_w) = a(b-1)\sigma^2 \tag{20.6.2}$$

$$E(V) = (ab-1)\sigma^2 + b\sum_j \alpha_j^2 \tag{20.6.3}$$

From (20.6.2) it follows that

$$E\left[\frac{V_w}{a(b-1)}\right] = \sigma^2 \tag{20.6.4}$$

so that

$$\hat{S}_w^2 = \frac{V_w}{a(b-1)} \tag{20.6.5}$$

is always a best (unbiased estimate of σ^2 regardless of whether H_0 is true or not. On the other hand, from (20.6.1) and (20.6.3), we see that only if H_0 is true will we have have

$$E\left[\frac{V_b}{a-1}\right] = \sigma^2 \qquad E\left[\frac{V}{ab-1}\right] = \sigma^2 \tag{20.6.6}$$

so that only in such case will

$$\hat{S}_b^2 = \frac{V_b}{a-1} \qquad \hat{S}^2 = \frac{V}{ab-1} \tag{20.6.7}$$

provide unbiased estimates of σ^2. If H_0 is not true, however, then we have from (20.6.1)

$$E[\hat{S}_b^2] = \sigma^2 + \frac{b}{a-1}\sum_j \alpha_j^2 \tag{20.6.8}$$

20.7 Distributions of the Variations

Theorem 20.7.1. $\dfrac{V_w}{\sigma^2}$ is chi-square distributed with $a(b-1)$ degrees of freedom.

Theorem 20.7.2. Under the null hypothesis H_0, $\dfrac{V_b}{\sigma^2}$ and $\dfrac{V}{\sigma^2}$ are chi-square distributed with $a-1$ and $ab-1$ degrees of freedom, respectively.

20.8 The F Test for the Null Hypothesis of Equal Means

If the null hypothesis H_0 is not true, i.e., if the treatment means are not equal, we see from (20.6.8) that we can expect $\hat{S}_b^2$ to be greater than σ^2, with the effect becoming more pronounced as the discrepancy between means increases. On the other hand, from (20.6.4) and (20.6.5) we can expect $\hat{S}_w^2$ to be equal to σ^2 regardless of whether the means are equal or not. It follows that a good statistic for testing the hypothesis H_0 is provided by $\dfrac{\hat{S}_b^2}{\hat{S}_w^2}$. If this is significantly large, we can conclude that there is a significant difference between treatment means and thus reject H_0. Otherwise, we can either accept H_0 or reserve judgement pending further analysis.

Theorem 20.8.1. The statistic $F = \dfrac{\hat{S}_b^2}{\hat{S}_w^2}$ has the F distribution with $a-1$ and $a(b-1)$ degrees of freedom.

20.9 Analysis of Variance Tables

The calculations required for the above test are summarized in Table 16.2, which is called an *analysis of variance table*. In practice we would compute v and v_b using either the long method, (20.3.1) and (20.3.6), or the short method, (20.4.1) and (20.4.2), and then compute $v_w = v - v_b$. It should be noted that the degrees of freedom for the total variation, i.e., $ab - 1$, is equal to the sum of the degrees of freedom for the between-treatment and within-treatments variations.

Table 16.2

Variation	Degrees of Freedom	Mean Square	F
Between Treatments, $v_b = b\sum_j (\bar{x}_{j\cdot} - \bar{x})^2$	$a - 1$	$\hat{s}_b^2 = \dfrac{v_b}{a-1}$	$\dfrac{\hat{s}_b^2}{\hat{s}_w^2}$ with $a-1, a(b-1)$ degrees of freedom
Within Treatments, $v_w = v - v_b$	$a(b-1)$	$\hat{s}_w^2 = \dfrac{v_w}{a(b-1)}$	
Total, $v = v_b + v_w = \sum_{j,k}(x_{jk} - \bar{x})^2$	$ab - 1$		

Example 20.9.1. Table 16.3 shows the yields in bushels per acre of a certain variety of wheat grown in a particular type of soil treated with chemicals A, B, or C.

Table 16.3

A	48	49	50	49
B	47	49	48	48
C	49	51	50	50

Find (a) the mean yields for the different treatments, (b) the grand mean for all treatments, (c) the total variation, (d) the variation between treatments, (e) the variation within treatments. Use the long method.

Solution: To simplify the arithmetic, we may subtract some suitable number, say, 45, from all the data without affecting the values of the variations. We then obtain the data of Table 16.4

Table 16.4

3	4	5	4
2	4	3	3
4	6	5	5

(a) The treatment (row) means for Table 16.4 are given, respectively, by

$$\begin{aligned}\overline{x}_1 &= \frac{1}{4}(3+4+5+4) = 4,\\ \overline{x}_2 &= \frac{1}{4}(2+4+3+3) = 3,\\ \overline{x}_3 &= \frac{1}{4}(4+6+5+5) = 5,\end{aligned}$$

Therefore, the mean yields, obtained by adding 45 to these, are 49, 48 and 50 bushels per acre for A, B and C respectively.

$$\text{(b)} \quad \overline{x} = \frac{1}{12}(3+4+5+4+2+4+3+3+4+6+5+5) = 4$$

Therefore, the grand mean for the original set of data is $45 + 4 = 46$ bushels per acre.

(c)

$$\begin{aligned}\text{Total variation} = v &= \sum_{j,k}(x_{jk}-\overline{x})^2\\ &= (3-4)^2+(4-4)^2+(5-4)^2+(4-4)^2\\ &\quad +(2-4)^2+(4-4)^2+(3-4)^2+(3-4)^2\\ &\quad +(4-4)^2+(6-4)^2+(5-4)^2+(5-4)^2\\ &= 14\end{aligned}$$

(d)

$$\begin{aligned}\text{Variation between treatments} = v_b &= b\sum_j(\overline{x}_j-\overline{x})^2\\ &= 4[(4-4)^2+(3-4)^2+(5-4)^2] = 8\end{aligned}$$

(e)

$$\text{Variation within treatments} = v_w = v - v_b = 14 - 8 = 6$$

Example 20.9.2. Referring to Example 20.9.1, find an unbiased estimate of the population variance σ^2 from (a) the variation between treatments under the null hypothesis of equal treatment means, (b) the variation within treatments.

Solution:

(a)

$$\hat{s}_b^2 = \frac{v_b}{a-1} = \frac{8}{3-1} = 4$$

(b)

$$\hat{s}_w^2 = \frac{v_w}{a(b-1)} = \frac{6}{3(4-1)} = \frac{2}{3}$$

Example 20.9.3. Referring to Example 20.9.1, can we reject the null hypothesis of equal means at (a) the 0.05 significance level? (b) the 0.01 significance level? (Given that $F_{0.95,2,9} = 4.26$ and $F_{0.99,2,9} = 8.02$).

Solution: We have

$$F = \frac{\hat{s}_b^2}{\hat{s}_w^2} = \frac{4}{2/3} = 6$$

with $a - 1 = 3 - 1 = 2$ and $a(b-1) = 3(4-1) = 9$ degrees of freedom.

(a) Since $F = 6 > F_{0.95,2,9} = 4.26$, we can reject the null hypothesis of equal means at the 0.05 level.

Table 16.5

Variation	Degrees of Freedom	Mean Square	F
Between Treatments, $v_b = 8$	$a - 1 = 2$	$\hat{s}_b^2 = \frac{8}{2} = 4$	$F = \frac{\hat{s}_b^2}{\hat{s}_w^2} = \frac{4}{2/3} = 6$ with 2, 9 degrees of freedom
Within Treatments, $v_w = v - v_b = 14 - 8 = 6$	$a(b-1) = (3)(3) = 9$	$\hat{s}_w^2 = \frac{6}{9} = \frac{2}{3}$	
Total, $v = 14$	$ab - 1 = (3)(4) - 1 = 11$		

(b) Since $F = 6 > F_{0.99,2,9} = 8.02$, we cannot reject the null hypothesis of equal means at the 0.01 level.

The analysis of variance table for Examples 20.9.1 - 20.9.3 is shown in Table 16.5.

Exercise 20.9.4. Use the shortcut formulas (20.4.1) through (20.4.3) to obtain the results of Example 20.9.1.

20.10 Modifications for Unequal Number of Observations

In case the treatments $1, \ldots, a$ have different numbers of observations equal to $n_1, \ldots, n_a$, respectively, the above results are easily modified. We therefore obtain

$$v = \sum_{j,k}(x_{jk} - \overline{x})^2 = \sum_{jk} x_{jk}^2 - \frac{\tau^2}{n} \tag{20.10.1}$$

$$v_b = \sum_{j,k}(\overline{x}_{j\cdot} - \overline{x})^2 = \sum_{j} n_j(\overline{x}_{j\cdot} - \overline{x})^2 = \sum_{j} \frac{\tau_{j\cdot}^2}{n_j} - \frac{\tau^2}{n} \tag{20.10.2}$$

$$v_w = v - v_b \tag{20.10.3}$$

where $\sum_{j,k}$ denotes the summation over k from 1 to n_j and then over j from 1 to a, $n = \sum_j n_j$ is the total number of observations in all treatments, τ is the sum of all observations, $\tau j\cdot$ is the sum of all values in the j-th treatment, and $\sum_j$ is the sum from $j = 1$ to a. The analysis of variance table for this case is given in Table 16.6.

Example 20.10.1. Table 16.7 shows the lifetimes in hours of samples from three different types of television tubes manufactured by a company. Using the long method, test at (a) the 0.05, (b) the 0.01 significance level whether there is a difference in the three types. (Given that $F_{0.95,2,9} = 4.26$ and $F_{0.99,2,9} = 8.02$).

Solution. It is convenient to subtract a suitable number, say, 400, obtaining Table 16.8. In this table we

Table 16.6

Variation	Degrees of Freedom	Mean Square	F
Between Treatments, $v_b = \sum_j n_j(\bar{x}_{j.} - \bar{x})^2$	$a - 1$	$\hat{s}_b^2 = \dfrac{v_b}{a-1}$	$\dfrac{\hat{s}_b^2}{\hat{s}_w^2}$ with $a-1, n-a$ degrees of freedom
Within Treatments, $v_w = v - v_b$	$n - a$	$\hat{s}_w^2 = \dfrac{v_w}{n-a}$	
Total, $v = v_b + v_w = \sum_{j,k}(x_{jk} - \bar{x})^2$	$n - 1$		

Table 16.7

Sample 1	407	411	409		
Sample 2	404	406	408	405	402
Sample 3	410	408	406	408	

have indicated the row total, the sample or group means, and the grand mean. We then have

$$\begin{aligned} v &= \sum_{j,k}(x_{jk} - \overline{x})^2 = (7-7)^2 + (11-7)^2 + \cdots + (8-7)^2 = 72 \\ v_b &= \sum_{j,k}(\overline{x}_{j.} - \overline{x})^2 = \sum_j n_j(\overline{x}_{j.} - \overline{x})^2 = 3(9-7)^2 + 5(7-5)^2 + 4(8-7)^2 = 36 \\ v_w &= v - v_b = 72 - 36 = 36 \end{aligned}$$

The data can be summarized in the analysis of variance table, Table 16.9. Now, for 2 and 9 degrees of freedom we have $F_{0.95,2,9} = 4.26$ and $F_{0.99,2,9} = 8.02$. Therefore, we can reject the hypothesis of equal means (i.e., there is no difference in the tree types of tubes) at the 0.05 level but not at the 0.01 level.

Table 16.8

						Total	Mean
Sample 1	7	11	9			27	9
Sample 2	4	6	8	5	2	25	5
Sample 3	10	8	6	8		32	8
	$\bar{x}$ = grand mean = $\frac{84}{12} = 7$						

Table 16.9

Variation	Degrees of Freedom	Mean Square	F
$v_b = 36$	$a - 1 = 2$	$\hat{s}_b^2 = \frac{36}{2} = 18$	$\frac{\hat{s}_b^2}{\hat{s}_w^2} = \frac{18}{4}$
$v_w = 36$	$n - a = 9$	$\hat{s}_w^2 = \frac{36}{9} = 4$	$= 4.5$

Exercise 20.10.2. Use the shortcut formulas (20.10.1) through (20.10.3) to obtain the results of Example 20.10.1.

References

1. J. P Tremblay and R. Manohar : Discrete Mathematical Structures with Applications to Computers.
2. J. L. Gersting : Mathematical Structures for Computer Sciences.
3. S. Lepschutz : Finite Mathematics.
4. F. Harary : Graph Theory.
5. S. Sahani : Concept of Discrete Mathematics.
6. J. E. Whitesitt : Boolean Algebra and its Applications.
7. K. L. P. Mishra and N. Chandrasekaran : Automata, Languages, and Computation.
8. G. E. Revesz : Introduction to Formal Languages.
9. K . D. Joshi : Foundation of Discrete Mathematics.
10. Modern Probability Theory: B. R. Bhat.
11. Elementary Probability Theory and Stochastic Processes: K. L. Chung.
12. Linear Statistical Inference and its Applications: C. R. Rao.
13. Life-testing and Reliability Estimation: S. K. Sinha & B. K. Kale.
14. An Outline of Statistical Theory (Vol 1 and 2): A. M. Goon, M. K. Gupta & B. Dasgupta.