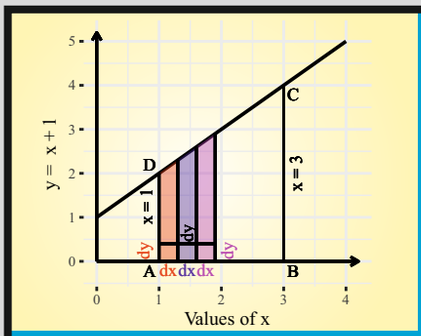
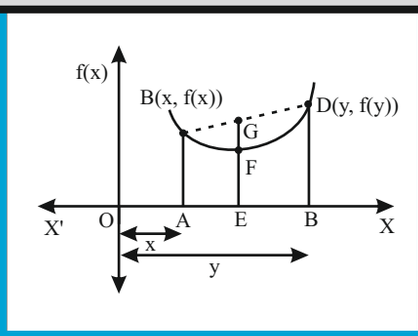
Indira Gandhi National Open University
School of Sciences

MST–011

REAL ANALYSIS, CALCULUS AND GEOMETRY





Even Function: $f(-x) = f(x)$

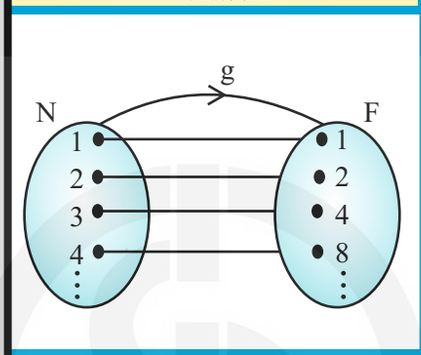
Odd Function: $f(-x) = -f(x)$

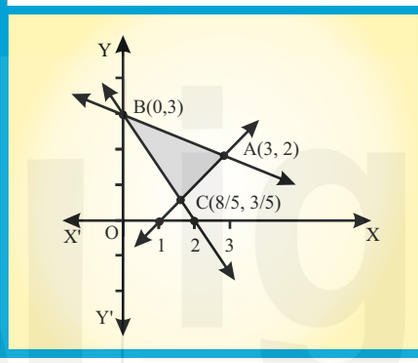
$N \sim E \Rightarrow E$ is enumerable

Countable: either finite or enumerable

μ is a Measure if

$$\mu(\phi)=0,\ \mu\left(\bigcup_{n=1}^{\infty} A_n\right)=\sum_{n=1}^{\infty}\mu(A_n)$$





Sequence $\langle a_n \rangle$ converges to a:

$$|a_n - a| < \varepsilon \quad \forall\ n \geq m$$

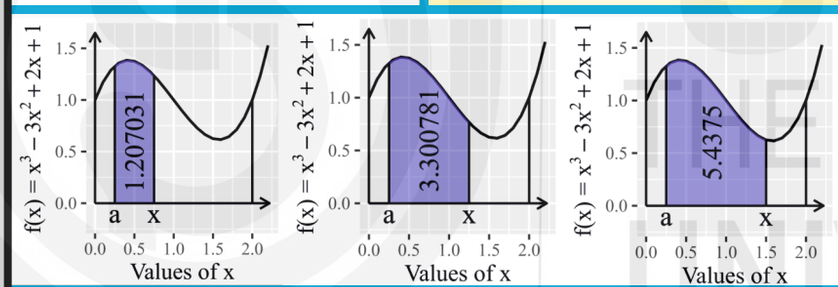
Write it as

$$a_n \to a \text{ as } n \to \infty \text{ or } \lim_{n\to\infty} a_n = a$$

Lower Sum: $L(f,P) = \sum_{i=1}^{n} m_i \delta_i$

Upper Sum: $U(f,P) = \sum_{i=1}^{n} M_i \delta_i$

A B $\quad AB = \int_2^5 dx = [x]_2^5 = 5 - 2 = 3$



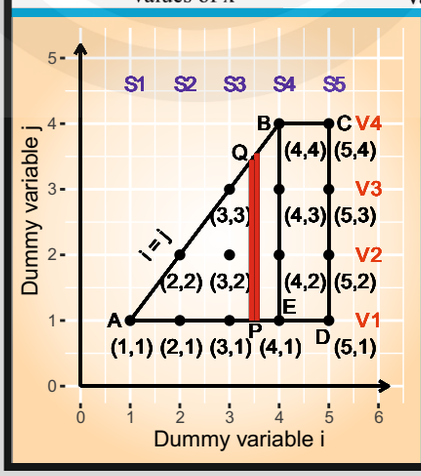
Line segment as a set of points

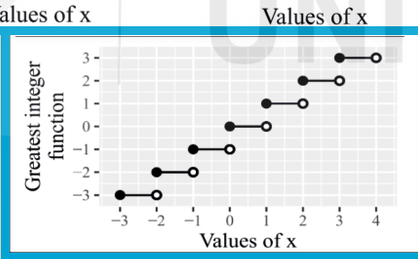
$$\{(1-\lambda)(x_1, y_1)+\lambda(x_2, y_2): 0 \leq \lambda \leq 1\}$$

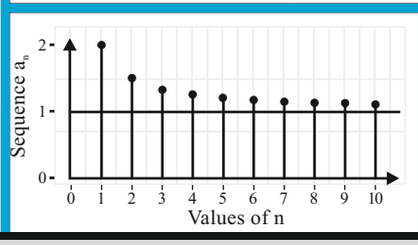
Equation of hyperplane is $\omega^T x + b = 0$

f is convex if $\forall$ x, y$\in$D and $\alpha \in [0,1]$

$$f((1-\alpha)x+\alpha y) \leq (1-\alpha)f(x)+\alpha f(y)$$







Gamma function: $\Gamma n = \int_0^{\infty} x^{n-1} e^{-x} dx,\ n > 0$

Beta function

$$\int_0^1 x^{m-1}(1-x)^{n-1}dx,\ \ m > 0,\ n > 0$$

$$B(m,n) = \frac{\Gamma m\ \Gamma n}{\Gamma m+n},\quad \Gamma\frac{1}{2} = \sqrt{\pi}$$

$$\sum_{k=1}^{n} a_k = \sum_{k=m}^{n+m-1} a_{k-m+1}$$

BLOCK 1: REAL ANALYSIS

BLOCK 2: HYPERPLANE AND CALCULUS

"शिक्षा मानव को बन्धनों से मुक्त करती है और आज के युग में तो यह लोकतंत्र की भावना का आधार भी है। जन्म तथा अन्य कारणों से उत्पन्न जाति एवं वर्गगत विषमताओं को दूर करते हुए मनुष्य को इन सबसे ऊपर उठाती है।"

– इन्दिरा गांधी

ignou
THE PEOPLE'S
UNIVERSITY

"Education is a liberating force, and in our age it is also a democratising force, cutting across the barriers of caste and class, smoothing out inequalities imposed by birth and other circumstances."

– Indira Gandhi

Indira Gandhi National Open University
School of Sciences

MST-011
Real Analysis, Calculus and Geometry

BLOCK 1

Real Analysis 3

BLOCK 2

Hyperplane and Calculus 157

REAL ANALYSIS, CALCULUS AND GEOMETRY

BLOCK 1: Real Analysis

Unit 1: Function of a Single Variable

Unit 2: Types, Continuity and Differentiability of Function of a Single Variable

Unit 3: Set Function and Distance Function

Unit 4: Sequences and Series

Unit 5: Riemann Integration

BLOCK 2: Hyperplane and Calculus

Unit 6: Hyperplane

Unit 7: Convex and Concave Function

Unit 8: Beta and Gamma Functions

Unit 9: Change of Order of Summation and Integration

ignou
THE PEOPLE'S
UNIVERSITY

Indira Gandhi National Open University
School of Sciences

MST-011
Real Analysis, Calculus and Geometry

Block

1

REAL ANALYSIS

UNIT 1
Function of a Single Variable 9

UNIT 2
Types, Continuity and Differentiability of Function of a Single Variable 45

UNIT 3
Set Function and Distance Function 69

UNIT 4
Sequences and Series 91

UNIT 5
Riemann Integration 121

Curriculum and Course Design Committee

Prof. Sujatha Varma
Former Director, SOS
IGNOU, New Delhi

Prof. Diwakar Shukla
Department of Mathematics and Statistics
Dr. H. S. Gaur Central University, Sagar (MP)

Prof. Gulshan Lal Taneja
Department of Mathematics
M. D. University, Rohtak (HR)

Prof. Gurprit Grover
Department of Statistics
University of Delhi, New Delhi

Prof. H. P. Singh
Department of Statistics
Vikram University, Ujjan (MP)

Prof. Rahul Roy
Mathematics and Statistics Unit
Indian Statistical Institute, New Delhi

Prof. Rajender Prasad
Division of Design of Experiments,
IASRI, Pusa, New Delhi

Prof. Rakesh Srivastava
Department of Statistics
M. S. University of Baroda, Vadodara (GUJ)

Prof. Sanjeev Kumar
Department of Statistics
Banaras Hindu University, Varanasi (UP)

Prof. Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur (UP)

Prof. V. K. Singh (Retd.)
Department of Statistics
Banaras Hindu University, Varanasi (UP)

Prof. Manish Trivedi, SOS, IGNOU

Dr. Taruna Kumari, SOS, IGNOU

Dr. Neha Garg, SOS, IGNOU

Dr. Rajesh, SOS, IGNOU

Dr. Prabhat Kumar Sangal, SOS, IGNOU

Dr. Gajraj Singh, SOS, IGNOU

Course Preparation Team

Course Editor

Prof. Pankaj Kumar (**Units 1-9**)
Department of Mathematics,
Guru Jambheshwar University of Science and Technology, Hisar, Haryana

Course Writer

Dr. Rajesh (**Units 1-9**)
Assistant Professor in Statistics,
School of Sciences,
Indira Gandhi National Open University,
New Delhi, Delhi

Formatted and CRC Prepared by Dr. Rajesh and Ms Preeti, SOS, IGNOU

Course Coordinator: Dr. Rajesh

Programme Coordinators: Dr. Neha Garg and Dr. Prabhat Kumar Sangal

Print Production

Mr. Rajiv Girdhar
Assistant Registrar
MPDD, IGNOU, New Delhi

Mr. Hemant Parida
Section Officer
MPDD, IGNOU, New Delhi

Acknowledgement: I gratefully acknowledge my colleague Dr. Prabhat Kumar Sangal for thoroughly reading the course material and providing his valuable suggestions to improve the course.

April;, 2023

© Indira Gandhi National Open University, 2023

ISBN-978-81-266-

All rights reserved. No part of this work may be reproduced in any form, by mimeograph or any other means, without permission in writing from the Indira Gandhi National Open University

Further information on the Indira Gandhi National Open University may be obtained from the University's Office at Maidan Garhi, New Delhi-110068 or visit University's website http://www.ignou.ac.in

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi by the Director, School of Sciences.

REAL ANALYSIS, CALCULUS AND GEOMETRY: COURSE INTRODUCTION

This is the first course of the M.Sc. Applied Statistics programme. In this course, some concepts of mathematics which are required to understand the mathematics behind some tools and techniques of statistics are explained. Visualisation of a concept makes understanding the underlying concept easy and fast. Keeping this in view and the applied nature of the programme lots of graphs have been included in the course. Most of these graphs are made by using the ggplot2 package of R. In this course, you will study some concepts of Real Analysis, Calculus and Geometry. So, this course is divided into two blocks where Block 1 is devoted to explain some concepts of Real Analysis and Block 2 is devoted to explain some concepts of Geometry, especially Hyperplane and Calculus.

Block 1 on real analysis explains ideas of function, set function, sequences, series and Riemann integration. This block is further divided into five units – Unit 1 to Unit 5. **Unit 1** discusses some terms related to a function like domain, range and co-domain of a function. The definition of a function is explained through the relation of the daughter and mother to understand it clearly without any burden of technical words. Domain, range and graphical behaviour of many functions which will be used in different courses of this programme have been discussed and explained through examples. **Unit 2** goes one step ahead of Unit 1 and explains types of functions like 1-1, onto and one-one correspondence. The important idea of a countable set is also discussed in this unit. Geometrical interpretation of types of functions has also been given. A review of the idea of continuity and differentiability of a function at a point has also been given in this unit. **Unit 3** explains the idea of a set function and some special set functions like a measure of a set, the length of an interval, etc. Unit 3 also explains the idea of distance function and a special distance function known as metric through lots of examples. Different types of distance functions like Minkowski, Manhattan, Euclidean, etc. have also been explained through examples and visualisation. The objective of **Unit 4** is to explain the idea of the convergence of sequences and series. Convergence and divergence of some sequences and series have been discussed using some common tests which are also mentioned in this unit. In statistics, you have to work in both worlds: discrete and continuous. **Unit 5** of this block is devoted to explain an important idea of how the role of summation in the discrete world is played by integration in the continuous world. To evaluate the probability of a continuous random variable we have to use the idea of integration. So, Riemann integration and some of its important properties have been discussed in this unit.

Block 2 on hyperplane and calculus discusses equations of a line in 2-dimension, 3-dimension, equations of the plane in 3-dimension and higher dimensions. This block also discusses the convex set, convex function and some of their properties. Beta, gamma functions and change of order of double summation and integral are also discussed in this block. Block 2 is further divided into four units – Unit 6 to Unit 9. **Unit 6** discusses equations of a line in various forms. Equations of planes in 3-dimension and in higher dimensions have been discussed. A brief overview of solutions of inequations in one and two variables is also given. **Unit 7** introduces the idea of a convex set and convex function together with some of their properties. To study beta and gamma distribution in the course MST-012 we need an understanding of beta and gamma functions. Keeping that in view **Unit 8** is devoted to discuss the graphical behaviour of beta and gamma functions. Some properties of beta and gamma functions are also proven and applied to solve problems based on beta and gamma functions. In statistics especially in probability theory many

times you have to change the order of double summation and integration. **Unit 9** is especially devoted to explain this important idea of how we can change the order of double summation and double integration. To explain this idea we have visualised it through many graphs.

The material has been developed for self-study. Many efforts are made to make the material self-explained and self-contained. IGNOU material is known for its characteristic of a built-in teacher in the material. We have done an attempt to meet that objective. However, if you want to refer to some additional books on these topics to know more about the content discussed in this course you may refer to any available books on these topics or any other source. Some books on the topics covered in this course are suggested as follows.

Suggested Further Reading

1. Thomas', G. B., Weir, M. D., Hass, J. & Giordano, F.R. (2004). Thomas' Calculus. Addison Wesley.
2. Tom, M. A. (2004). Mathematical Analysis. Pearson Education.
3. Goldberg, R. R. (1976). Methods of Real Analysis. John Wiley & Sons.
4. Rana, I. K. (2002). An Introduction to Measure and Integration. American Mathematical Society Providence, Rhode Island.
5. Rockafellar, R. T. (1997). Convex Analysis. Princeton Landmarks in Mathematics and Physics Series.

Expected Learning Outcomes

After completing this course, you should be able to:

- explain meanings of function, set function, and the convergence of sequences and series;
- evaluate integral using two definitions of Riemann integration;
- write equations of line, plane and hyperplane in different forms;
- explain the idea of the convex set and convex function. You will be also able to state and prove some properties of the convex set and convex function;
- solve problems based on beta and gamma functions; and
- change the order of double summation and integration.

Course Preparation Team

BLOCK 1 REAL ANALYSIS

In this block, you will get a graphical analysis of many common functions of a single variable which will help you to understand many silent features of these functions. What we mean by a set function and a distance function are also explained in this block. Ideas of convergence and divergence of a sequence and series are also discussed in this block. The last unit of this block will make you familiar with an important idea of how integration plays the role of summation in a continuous world.

This block is divided into five units, namely, Units 1- 5.

Unit 1: The purpose of this unit is to present a graphical analysis of the function of a single variable. This unit discusses the definition of a function and explains how it can be easily understood by comparing the association of daughter and mother. In this unit domain, range and graphical behaviour of many common functions which will be used in different courses of this programme are discussed.

Unit 2: To understand the difference between discrete and continuous variables you should have a good understanding of the world countable. But to understand the meaning of countable in mathematics you should have an understanding of the one-one correspondence between two sets. This unit explains all these ideas with lots of examples and visualisation of the ideas.

Unit 3: In the course MST-012 you will learn to associate probabilities with the events of your interest. This association of probability is a kind of job like defining a set function. The idea of set function is introduced in this unit. In the course MST-026 you have to use some distance functions in some machine learning algorithms. The idea of distance function is also discussed in this unit.

Unit 4: In some courses of this programme you will talk about sequence, series and their convergence. In this unit, the ideas of convergence of a sequence and series have been introduced and explained through many examples.

Unit 5: In the course MST-012 you will study both discrete and continuous probability distributions. To obtain some measures (e.g. mean, variance, etc.) of a discrete probability distribution you apply summation while in continuous probability distributions role of summation is played by integration. You will get an explanation of how the role of summation is played by integration. To obtain the probability of continuous distributions you apply integration. The idea of integration, Riemann integration is also explained through a lot of examples and visualisation.

This material has been developed for self-study. However, if you are interested in studying the content in greater depth/more detail, it would be useful for you to attempt related exercises given in the relevant chapters of the books listed at the end of the course introduction. We hope you will enjoy studying this block.

Expected Learning Outcomes

After completing this block, you should be able to:

- explain meanings of function, set function, convergence and divergence of sequences and series; and
- evaluate integral using two definitions of Riemann integration.

Block Preparation Team

ignou THE PEOPLE'S UNIVERSITY

Notations and Symbols

$:$	:	such that
$\forall$	:	for all
$\mathbb{N}$	:	the family of natural numbers
$\mathbb{Z}$	:	the family of integers numbers
$\mathbb{Q}$	:	the family of rational numbers
$\mathbb{R}$	:	the family of real numbers
$\mathbb{R}^*$	:	the family of extended real numbers
$\mathbb{C}$	:	the family of complex numbers
σ	:	sigmoid function and population standard deviation
sgn	:	signum function
$[x]$	:	greatest integer function or floor function
$\lceil x \rceil$	:	ceiling function
fog	:	composition of functions f and g
Ω	:	universal set
$\mathcal{F}$	:	Field and σ-field of subsets of the universal set Ω
μ	:	set function
λ	:	length function
d	:	distance function
d(x, y)	:	distance between points x and y
$< a_n >$	:	The sequence where a_n is its nth term
L(f, P)	:	the lower sum of f over the partition P
U(f, P)	:	the upper sum of f over the partition P
P[a, b]	:	set of all partition over [a, b]
R[a, b]	:	set of all Riemann integrable functions over [a, b]

Letters A to Z in the calligraphic font are:

$\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}, \mathcal{H}, \mathcal{I}, \mathcal{J}, \mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{O}, \mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$

Letters A to Z in black board font are:

$\mathbb{A}, \mathbb{B}, \mathbb{C}, \mathbb{D}, \mathbb{E}, \mathbb{F}, \mathbb{G}, \mathbb{H}, \mathbb{I}, \mathbb{J}, \mathbb{K}, \mathbb{L}, \mathbb{M}, \mathbb{N}, \mathbb{O}, \mathbb{P}, \mathbb{Q}, \mathbb{R}, \mathbb{S}, \mathbb{T}, \mathbb{U}, \mathbb{V}, \mathbb{W}, \mathbb{X}, \mathbb{Y}, \mathbb{Z}$

UNIT 1

FUNCTION OF A SINGLE VARIABLE

Structure

1.1 Introduction

Expected Learning Outcomes

1.2 Constant and Variable Quantity

1.3 Interval and its Types

1.4 Definition, Notations and Pictorial Presentation of a Function

1.5 Some Special Functions having Applications in Statistics, Probability and Machine Learning

1.6 Summary

1.7 Terminal Questions

1.8 Solutions/Answers

1.1 INTRODUCTION

In statistics, we use parameters for population characteristics such as mean, variance, etc. Constants and variables are also used in statistics. So, it becomes important to understand the distinction between fixed constant, parameter and variable which is discussed in Sec. 1.2. Another important thing that we are going to use on many occasions in different courses is the interval. In Sec. 1.3 different types of intervals are discussed. To understand many concepts of probability and statistics such as probability mass function, probability density function, cumulative distribution function, likelihood function, etc. you should have a good understanding of function, therefore, a brief overview of the function is given in Sec. 1.4. There are some special functions which have their applications in statistics, probability and machine learning such as linear, polynomial, logarithm, exponential, modulus functions, logistic sigmoid function, etc. which are discussed in Sec. 1.5. In the course MST-024 on machine learning you will also use many other functions such as loss function, distance function (Manhattan distance, Euclidean distance, Chebyshev distance), etc. The distance function is discussed in Unit 3.

What we have discussed in this unit is summarised in Sec. 1.6. Self Assessment Questions (SAQs) have been given in some sections which are generally based on the content discussed in that section. But to give you a good practice of what we have discussed in this unit, some more questions

based on the entire unit are given in Sec. 1.7 under the heading Terminal Questions. One of the characteristics of IGNOU study material is that it is self-contained. To make it self-contained regarding solutions of all the SAQs and Terminal Questions we have given solutions of all the SAQs and Terminal Questions of this unit in Sec. 1.8. To best utilise this section, it is recommended that after going through all the concepts discussed in the unit first you should try solutions of all the SAQ's and Terminal Questions yourself. In the case, your answer does not match with the answer given in the solution of the same you may give it another try if still not matching you may refer to the given solution and identify the concept where you need to improve.

In the next unit, you will study types of functions, countable sets, continuity and differentiability of a function.

Expected Learning Outcomes

After completing this unit, you should be able to:

- define different types of constants, variables, and intervals, and can give examples of each;
- define a function and obtain its domain and range;
- evaluate the value of a function at given points; and
- plot a graph of a function and identify its graphical behaviour.

1.2 CONSTANT AND VARIABLE QUANTITY

Suppose someone wants to know about you then perhaps he/she should know your name, the background of the family you belong to, your educational qualifications, how good as a human being you are, other achievements, etc. Similarly, if you want to know/study the subject statistics then it requires that first, you should have an understanding of some rules of mathematics. Understanding these rules of mathematics will make the journey of studying tools/techniques of the subject of statistics easy as well as interesting. Understanding which tools/techniques of the subject statistics will apply in a given situation is a long journey. Purpose of different courses of this master's degree programme M.Sc. Applied Statistics is to start this journey smoothly and in a sequence. Understanding the mathematics rules required to study statistics is not hard if you start from the basics. In fact, when you start from basic, learning any subject becomes interesting and creates a hunger to learn more and more. Like your name, the subject that we are going to discuss here also has its name called mathematics. Unlike your one family learning mathematics starts from knowing about its many families called families of numbers. You know about these families of numbers from your school mathematics. Let us recall these well-known families of numbers mentioned as follows.

Family of natural numbers: $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$ … (1.1)

Family of whole numbers: $W = \{0, 1, 2, 3, 4, \ldots\}$ … (1.2)

Family of integers: $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ … (1.3)

Family of rational numbers: $\mathbb{Q} = \left\{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0, (p, q) = 1\right\}$ … (1.4)

where (p, q) represents the greatest common divisor (gcd) or highest common factor (hcf) of p and q

Family of irrational numbers: A number on the real line which cannot be expressed in p/q form where p and q are integers and q is not equal to zero is called an irrational number. In books on mathematics generally, no notation is specified for the set of all irrational numbers like $\mathbb{Q}$ is used for the set of all rational numbers. … (1.5)

Family of real numbers: $\mathbb{R} = \mathbb{Q} \cup \mathrm{I}$, where I represents the set of all irrational numbers. … (1.6)

Family of complex numbers: $\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$, where the symbol i is read as iota and denotes the square root of – 1, i.e., $i = \sqrt{-1}$, with $i^2 = -1$. (1.7)

Note 1: In higher mathematics family of natural, integer, rational, real and complex numbers are denoted by the symbols $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$ respectively.

After getting the idea about these different families of numbers, the next thing that you should be familiar with is the meaning of constant and variable. Actually, you have used them so many times in school mathematics. Here, we are discussing them because we want to further classify them. In fact, they are a particular type of quantity so the story begins by defining the quantity itself as follows.

Quantity

In mathematics and statistics, quantity means those things on which four basic mathematical operations: addition, subtraction, multiplication and division can be applied. For example: (i) 2, 5 (ii) population mean (μ) (iii) number of children per family (iv) all real numbers between 2 and 3 including 2 and 3, etc. all are examples of different types of quantity which you will come across during your study of different courses of M.Sc. Applied Statistics programme.

A quantity can be classified as a constant or a variable quantity. Constant quantity can be further classified as fixed constant and arbitrary constant/parameter. Variable quantity can be further classified as discrete and continuous. All these classifications are shown in Fig. 1.1.

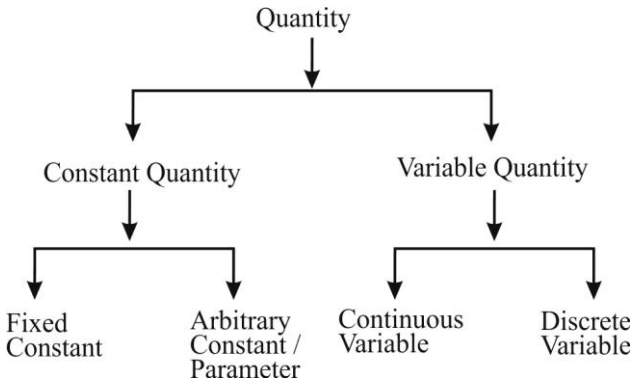


Fig. 1.1: Types of quantity in mathematics and statistics

Constant Quantity

A quantity which remains same (unchanged) throughout a particular problem or distribution is known as a constant quantity. A constant quantity can be further classified as a fixed constant and arbitrary constant (also known as a parameter) which are defined as follows:

Fixed Constant: Those types of constants which always remain the same (unchanged) independent of the place and time are known as fixed constants.

For example: 2, $\frac{3}{2}$, $\sqrt{13}$, $-\sqrt{17}$, π, etc. are fixed constants because the number 2 will remain the same whether you use it in the morning or evening or in India or in any other country. Similarly, the number π which is equal to the quotient of the circumference of a circle ($2\pi r$) and its diameter ($2r$) is a fixed constant. Value of π does not depend on the length of the radius of the circle. Also, remember 22/7 is not the exact value of π it is just an approximate value.

$$\text{Exact value of } \pi = \frac{\text{Length of circumference of the circle}}{\text{Length of diameter of the same circle}}. \qquad \text{... (1.8)}$$

Arbitrary Constant or Parameter: Those types of constants which remain the same in one problem but may vary from problem to problem or distribution to distribution are known as arbitrary constants or parameters and are generally denoted by a, b, c, l, m, n, α, β, γ, μ, σ^2, ρ, etc.

For example, from school mathematics, you know that the equation $y = x + c$ represents a straight line which makes an angle of 45^0 ($\because$ slope is 1 = coefficient of x if we compare it with $y = mx + c$) with a positive direction of the x-axis. For each real value of c, this equation gives a different straight line refer to Fig. 1.2. Here, c behaves like an arbitrary constant because for a particular line, its value is a fixed constant but its value varies from line to line. Similarly, in normal distribution mean (μ) and variance (σ^2) are parameters because they remain fixed constant in a particular normal distribution but vary from normal distribution to normal distribution refer to Fig. 1.3.

Variable

A quantity which may change its value even in a particular problem is known

as a variable. For example, in the quadratic equation $x^2 - x - 12 = 0$, x is a variable because it takes two values 4 and – 3.

$$\left[\because x^2 - x - 12 = 0 \;\Rightarrow (x-4)(x+3) = 0 \;\Rightarrow x = 4 \text{ or } -3\right]$$

A variable can be further classified as discrete and continuous explained as follows.

Discrete Variable: If the nature of the quantity is such that its possible values can be written in a sequence then such a quantity is known as **discrete in nature** or **discrete variable**. For example, the number of children in a family in a country may be 0, 1, 2, 3, … . Here, you know that the minimum possible number of children in a family is 0 and after 0 the next possible value for the

quantity, the number of children in a family is 1. There is no possibility between 0 and 1 for the quantity, the number of children in a family. Similarly, after 1 the next possible value for the quantity, the number of children in a family is 2 and so on. Here, you can write the possible values in a sequence (0, 1, 2, 3, …) so the quantity, the number of children in a family is a discrete quantity and so is an example of a discrete variable. Other examples of discrete variables are the number of accidents at a particular site on a road, the number of defective items in a packet of 100 items, the outcome of a die when it is thrown once, etc.

Continuous Variable: If the nature of the quantity is such that it can take any possible value between two certain limits, then such a quantity is known as **continuous in nature** or **continuous variable**. In other words, if a quantity is continuous in nature, then you cannot write all its possible values in a sequence. For example, you cannot write all real numbers from 2 to 3 in a

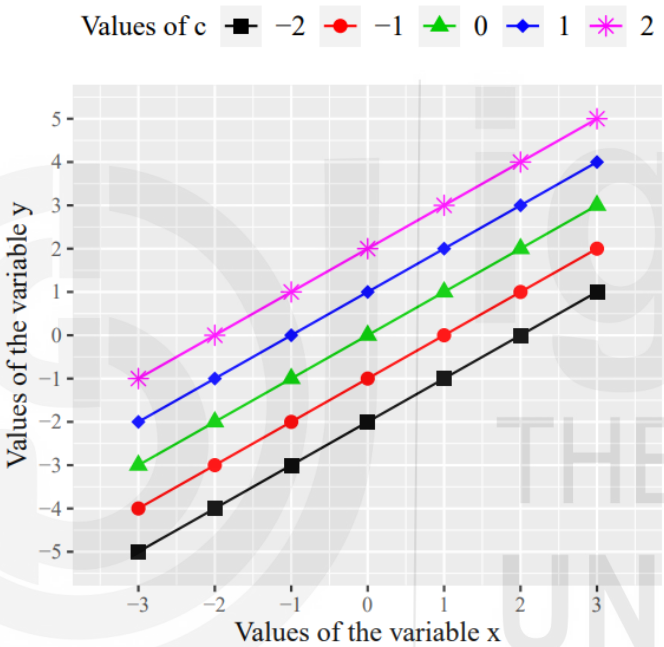


Fig. 1.2: Represents graph of straight lines y = x + c for different values of c

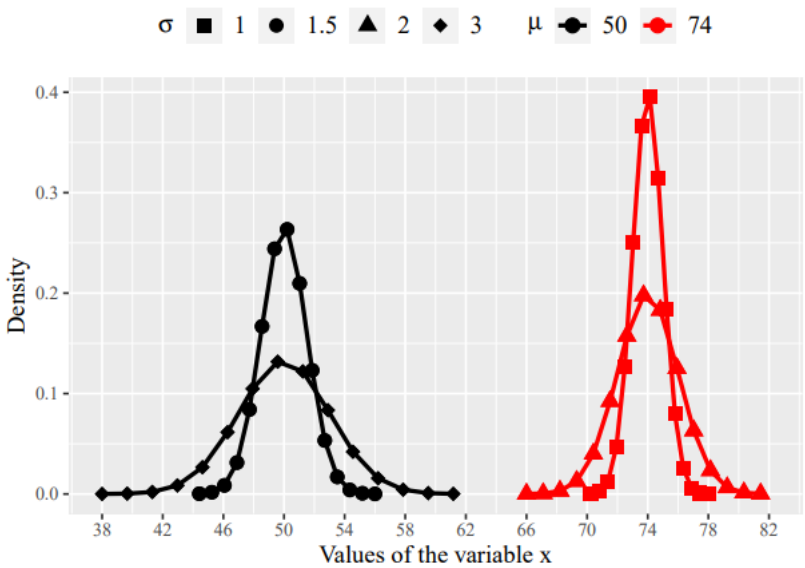


Fig. 1.3: Represents graph of probability density functions of the normal distribution for different values of μ and σ

sequence. In fact, you even cannot tell what is the next possible real number just after 2. If you say 2.1 is the next real number after 2, no it is not the next real number after 2 because 2.01 is between 2 and 2.1. If you say 2.01 is the next real number after 2, no it is not the next real number after 2 because 2.001 is between 2 and 2.01. Similarly, if you say 2.001 is the next real number after 2, no it is not the next real number after 2 because 2.0001 is between 2 and 2.001 and so on. The more advanced mathematical distinction between continuous and discrete quantities can be understood after getting an idea of countable and uncountable sets discussed in Sec. 2.3 of Unit 2 of this course. Other examples of continuous variables are temperature, height, weight, age, etc.

1.3 INTERVAL AND ITS TYPES

Interval

In Sec. 1.2 you have become familiar with well-known families of numbers keep that familiarity in your mind we will recall those families by their names as and when will need.

In this section, we will study interval and its types. Interval makes sense only in the family of real numbers and is defined as follows:

Let $\mathbb{R}$ be the set of all real numbers. Then a set $I \subseteq \mathbb{R}$ is said to be an **interval** if whenever $a, b \in I$ and $a < x < b$ then $x \in I$. For example, the set of all real numbers satisfying $2 \le x \le 3$ is an interval where x can take any real value between 2 and 3 including 2 and 3. … (1.9)

Now, you have gone through the definition of interval so we can explain the point why interval makes sense only in the family of real numbers. For example, consider the family of rational numbers: as $1, 2 \in \mathbb{Q}$ then we cannot form an interval in the family of rational numbers with endpoints 1 and 2 including 1 and 2 because you know that value of $\sqrt{2}$ up to three decimal places is 1.414 and so $1 < \sqrt{2} < 2$ but $\sqrt{2}$ is an irrational number, not a rational number and hence $\sqrt{2} \notin [1, 2]$ because we are assuming the interval [1, 2] in the family of rational numbers. That is why interval makes sense only in the family of real numbers.

Types of Intervals

Intervals are classified on the basis of two criteria:

(a) Whether it contains its endpoints or does not contain its endpoints

(b) Whether its length is finite or infinite

Based on these criteria different types of intervals are defined as follows:

(a) Different Types of Intervals based on the Criteria of Whether it contains its endpoints or does not contain its endpoints

On the basis of endpoints criterion, an interval may be classified as an open or a closed or one side open other side closed interval defined as follows:

Open Interval: An interval $I \subseteq \mathbb{R}$ with end points a and b (a < b) is called an open interval if it contains all real numbers between a and b but does not contain end points a and b. An open interval with endpoints a and b is denoted by (a, b) and mathematically is defined as follows:

$$(a,b) = \{x \in \mathbb{R} : a < x < b\} \qquad \ldots (1.10)$$

For example, open interval (2, 5) contains each real number lying between 2 and 5 but does not contain 2 and 5. To represent an open interval on real line we make open dots at both endpoints. Fig. 1.4 (a) indicates it on the real line.

i.e., $x \in (2, 5) \Rightarrow 2 < x < 5$

Closed Interval: An interval $I \subseteq \mathbb{R}$ with endpoints a and b (a < b) is called closed interval if it contains all real numbers between a and b and also contains endpoints a and b. A closed interval with endpoints a and b is denoted by [a, b] and mathematically is defined as follows:

$$[a, b] = \{x \in \mathbb{R} : a \le x \le b\} \qquad \ldots (1.11)$$

For example, closed interval [2, 5] contains each real number lying between 2 and 5 including 2 and 5. To represent a closed interval on a real line we make solid dots at both endpoints. Fig. 1.4 (b) indicates it on the real line.

i.e., $x \in [2, 5] \Rightarrow 2 \le x \le 5$

Left Open and Right Closed Interval: An interval $I \subseteq \mathbb{R}$ with endpoints a and b (a < b) is called a left open and right closed interval if it contains all real numbers between a and b and also includes only one endpoint b. A left open and right closed interval with endpoints a and b is denoted by (a, b] and mathematically is defined as follows:

$$(a,b] = \{x \in \mathbb{R} : a < x \le b\} \qquad \ldots (1.12)$$

For example, left open and right closed interval (2, 5] contain each real number lying between 2 and 5 including end point 5 but does not contain end point 2. To represent it on real line we have to make an open dot at 2 and a solid dot at 5. Fig. 1.4 (c) indicates it on the real line.

i.e., $x \in (2,5] \Rightarrow 2 < x \le 5$

Left Closed and Right Open Interval: An interval $I \subseteq \mathbb{R}$ with endpoints a and b (a < b) is called a left closed and right open interval if it contains all real numbers between a and b and also includes only one end point a. A left closed and right open interval with endpoints a and b is denoted by [a, b) and mathematically is defined as follows:

$$[a, b) = \{x \in \mathbb{R} : a \le x < b\} \qquad \ldots (1.13)$$

For example, left closed and right open interval [2, 5) contain each real number lying between 2 and 5 including endpoint 2 but does not contain endpoint 5. To represent it on a real line we have to make a solid dot at 2 and an open dot at 5. Fig. 1.4 (d) indicates it on the real line.

i.e., $x \in [2, 5) \Rightarrow 2 \le x < 5$

(b) Different Types of Intervals based on the Criteria Whether its length is finite or infinite

On the basis of the criterion of length of an interval it may be classified as finite or infinite interval defined as follows:

Length of an Interval: Length of each of the intervals (a, b), [a, b], (a, b], [a, b) is defined as $b-a,\ a<b$.

i.e., length of the interval = difference of the end points … (1.14)

For example, if I = (2, 7) then $\ell(\mathrm{I}) = 7-2 = 5$, , where $\ell(\mathrm{I})$ denotes the length of the interval I.

Finite Interval: An interval is said to be finite if its length is finite. (1.15)

For example, if I = (– 3, 5) then $\ell(\mathrm{I})$ = 5 – (– 3) = 5 + 3 = 8 which is finite.

$\therefore$ interval I = (– 3, 5) is finite.

Infinite Interval: An interval is said to be infinite interval if its length is not finite. … (1.16)

For example,

(i) The set $\{x \in \mathbb{R} : x > a,\ a \in \mathbb{R}\}$ is an infinite interval and is denoted by (a, ∞). Fig. 1.4 (e) Indicates the infinite interval $(2, \infty)$ on real line.

(ii) The set $\{x \in \mathbb{R} : x < a,\ a \in \mathbb{R}\}$ is an infinite interval and is denoted by $(-\infty, a)$. Fig. 1.4 (f) indicates the infinite interval $(-\infty, 5)$ on real line.

Similarly, infinite intervals $[a, \infty)$, $(-\infty, a]$ are defined as

$[a, \infty) = \{x \in R : x \geq a\}$, where 'a' is a fixed real number

$(-\infty, a] = \{x \in R : x \leq a\}$, where 'a' is a fixed real number

Also, $(-\infty, \infty) = \{x \in R : -\infty < x < \infty\}$, where 'a' is a fixed real number

Fig. 1.4 (g) to (i) indicate the infinite intervals $(-\infty, \infty)$, $[2, \infty)$, and $(-\infty, 5]$ respectively on real line.

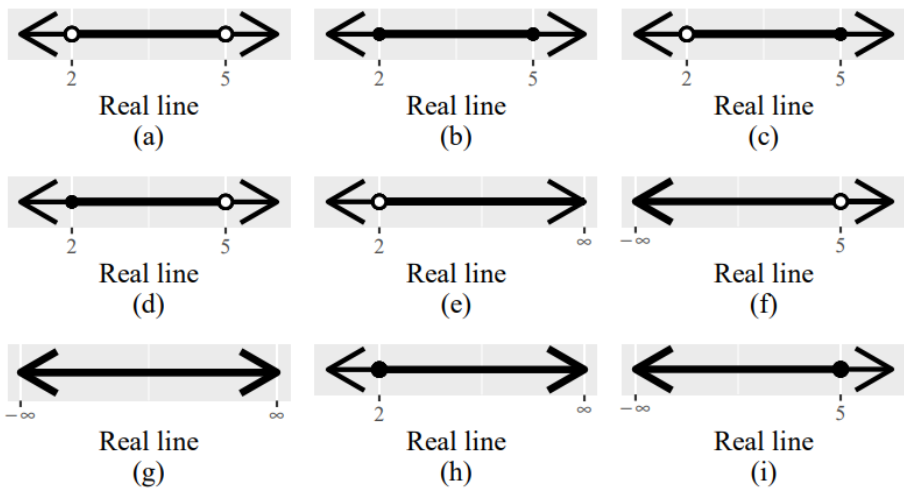


Fig. 1.4: Visualisation of (a) open interval (b) closed interval (c) left open and right closed (d) left closed and right open (e to i) infinite intervals

Remark 1

- Each interval contains infinitely many elements.
- Each interval is an infinite set but an infinite set may or may not be an interval. For example, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ are infinite sets but are not intervals.
- A set may or may not be an interval. For example, {2, 7} is a set but not an interval.
- $\mathbb{R}$, set of all real numbers, is an infinite interval given by $\mathbb{R} = (-\infty, \infty)$.
- Remember that $-\infty$ and ∞ are not included in the set of real numbers. Also, if extreme value is ∞ or $-\infty$, then open bracket is used on that side of the interval having ∞ or $-\infty$. In Unit 3 of this course we will extend set of real numbers by including two special numbers $-\infty$ and $+\infty$ and we will denote this new set by $\mathbb{R}^* = (-\infty, \infty) \cup \{-\infty, +\infty\} = \mathbb{R} \cup \{-\infty, +\infty\}$.
- When we say that x is a finite number (real number) it means that $-\infty < x < \infty$.

Now, we are in a position to define a function.

1.4 DEFINITION, NOTATIONS AND PICTORIAL PRESENTATION OF A FUNCTION

Definition of a Function

You know about function from school mathematics. You can treat function as a machine if it works in such a way that (i) corresponding to each input it gives output and (ii) output is unique for each input. Mathematically, a function is defined as follows:

Let X and Y be two non-empty sets. Then a rule which associates each element of X to a unique element of Y is called a function from X to Y. X is called **domain** of the function. Y is called **co-domain** of the function, and the set of only those values of Y for which function is defined is called **range** of the function. That is, the subset $\{y \in Y : y = f(x) \text{ for some } x \in X\}$ of Y is called range of the function. … (1.17)

If we treat function as a machine then the set of all inputs is called domain of the function and the set of all outputs is called range of the function. If we define a function $f : \mathbb{Z} \to \mathbb{N}$ by

$$y = f(x) = x^2 + 1, \quad x \in \mathbb{Z}$$

then its visualisation as a machine is shown in Fig. 1.5.

Usual Notations used for Defining a Function

(i) A function is generally denoted by f, g, h, ϕ, ψ, etc., in the case of above definition we write f: $X \to Y$ and read as f is a function from X to Y.

(ii) A function f: $X \to Y$ is generally described by writing $y = f(x), \ x \in X$, where $f(x)$ is an expression in terms of x.

For example, the function used in Fig. 1.5 has expression $f(x) = x^2 + 1$.

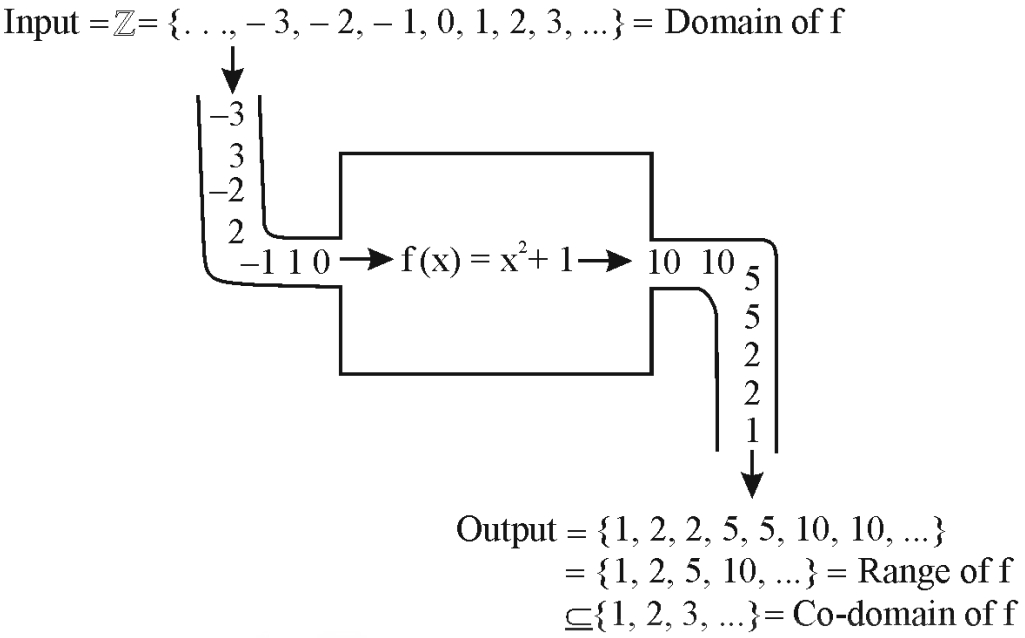


Fig. 1.5: Visualisation of treating function as a machine

Pictorial Presentation of a Function

There are two conditions for a rule to be a function.

(i) Each element of X must be associated to some element of Y.

(ii) There is unique element of Y corresponding to each element of X.

Visualisation of a concept helps in understanding the concept in a better way. Following the same strategy, we are using pictorial presentation of a function which is an attempt to explain the idea of the function through diagram known as pictorial presentation. The box given in the margin of this page specifies the two requirements for a rule to be a function. These two requirements are also satisfied by the association of daughter with her mother. So, to explain the idea of the function through pictorial presentation we will continue with example of daughters and mothers.

Let X = D = set of daughters and Y = M = set of mothers. Then consider the following situations given in Fig. 1.6 (a) to (d) with the help of the diagrams.

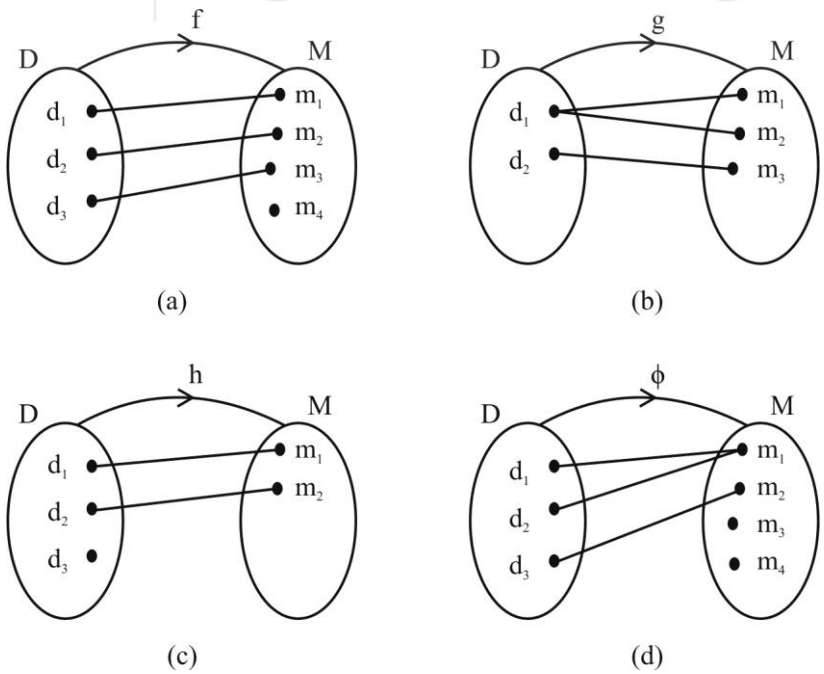


Fig. 1.6: (a) to (d) Pictorial presentation of rules f, g, h, ϕ

The rule f shown in Fig. 1.6 (a) is a function because each daughter has unique mother. [i.e., both conditions mentioned in the box are satisfied] and domain of the function $f =$ set of all daughters $= D = \{d_1, d_2, d_3\}$

co-domain of f = set of all mothers = $M = \{m_1, m_2, m_3, m_4\}$

range of $f =$ set of those mothers who have at least one daughter

$= \{m_1, m_2, m_3\}$

Note 2: One point which may come in your mind is that if m_4 is a mother then there should be at least one daughter of the mother m_4. But as we know that to become a mother it is not necessary that there should be a daughter. A mother may have only one son or only two sons or more than two sons without a daughter.

The rule g shown in Fig. 1.6 (b) is not a function because daughter d_1 has two mothers m_1 and m_2 which is not possible. [i.e., condition (ii) given in the box is not satisfied]

The rule h shown in Fig. 1.6 (c) is not a function because daughter d_3 has no mother. If the daughter d_3 came in this world, then there should be some mother of the daughter d_3. [i.e., condition (i) given in the box is not satisfied]

The rule ϕ shown in Fig. 1.6 (d) is a function because each daughter has unique mother. We see that both daughters d_1, d_2 have the same mother, no problem it is possible. Further mothers m_3, m_4 have no daughters again no problem it is also possible. In this case:

domain of the function $f = \{d_1, d_2, d_3\} = D =$ set of all daughters,

co-domain of the function $\phi = \{m_1, m_2, m_3, m_4\} = M =$ set of all mothers and range of the function $f = \{m_1, m_2\} =$ set of only those mothers who have at least one daughter.

Remark 2

- If $f : X \to Y$ is a function given by $y = f(x)$ then x is known as **pre image** of y and y is known as **image** of x.
- If $y = f(x)$ is a function then values of y depend on values of x. So, y is known as **dependent variable** and x is known as **independent variable**.
- In descriptive statistics we also use dependent and independent variables, especially in correlation and regression. In biostatistics terminology that is generally followed is: independent variable is known as **regressor** and dependent variable is known as **response variable**. In the course MST-024 on machine learning independent variable is known as **feature** and dependent variable is known as **target variable**. So, keep this terminology alert in your mind.

Some Examples of Functions

Example 1: Let $f : \mathbb{N} \to \mathbb{N}$ be defined by

$f(n) = 3n, \quad n \in \mathbb{N}$

Express the function diagrammatically. Also write domain, range and co-domain of the function. Also make the graph of this function. Further, if instead of from $\mathbb{N}$ to $\mathbb{N}$ suppose f is from $\mathbb{R}$ to $\mathbb{R}$ draw its graph.

Solution: $f : \mathbb{N} \to \mathbb{N}$ be defined by

$f(x) = 3n, \quad n \in \mathbb{N}$

$\therefore f(1) = 3, \; f(2) = 6, \; f(3) = 9$ and so on.

Pictorial presentation of the function is shown in Fig. 1.7.

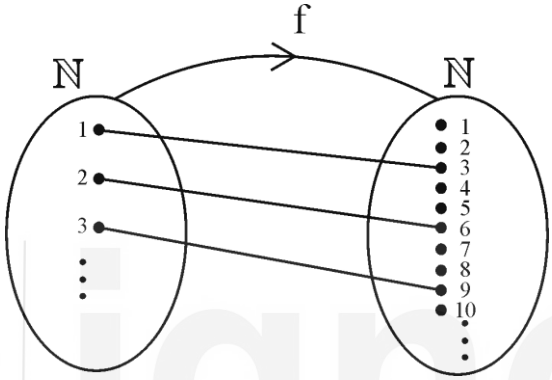


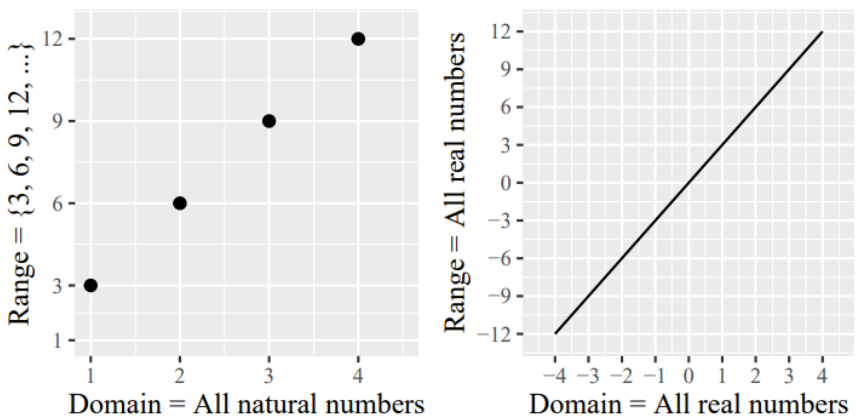
Fig. 1.7: Pictorial presentation of the function

Domain of the function $f = \{1, 2, 3, \ldots\} = \mathbb{N}$

Range of the function f = Set of only those values for which function is defined
= {3, 6, 9, …}

Co-domain $= \{1, 2, 3, \ldots\} = \mathbb{N}$

Graph of the function when both domain and co-domain are set of natural numbers is shown in Fig. 1.8 (a). Graph of the function when both domain and co-domain are set of real numbers is shown in Fig. 1.8 (b). Note the effect of domain and range on the graph of a function. Here mathematical expression of the function is the same but domain and range are different. In Fig. 1.8 (a) only points at (1, 1), (2, 2), (3, 3), (4, 4), … represent its graph while in Fig. 1.8 (b) complete line represent its graph.



(a) (b)

Fig. 1.8: Graph of the function f(x) = 3x (a) when f is from $\mathbb{N}$ to $\mathbb{N}$ (b) when f is from $\mathbb{R}$ to $\mathbb{R}$

Example 2: Find the domain of the function $f:\mathbb{R}\to\mathbb{R}$, defined by

$$f(x)=\sqrt{(x-3)(5-x)}\ ,\quad x\in\mathbb{R}$$

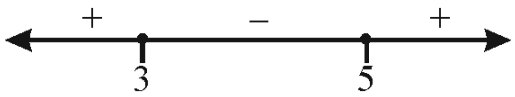
Also evaluate f(3), f(4), f(5), f(7).

Solution: Given function is

$$f(x)=\sqrt{(x-3)(5-x)},\quad x\in\mathbb{R}$$

For $f(x)$ to be real, the quantity under the square root should be non-negative and hence

$$(x-3)(5-x)\geq 0$$



$$\Rightarrow -(x-3)(x-5)\geq 0$$

$$\Rightarrow (x-3)(x-5)\leq 0 \quad \left[\begin{array}{l}\because \text{ when we multiply by } -1 \text{ on both sides of an}\\ \text{inequation then its inequality changes. For more}\\ \text{detail you may refer Sec. 6.7 of Unit 6 of this course}\end{array}\right]$$

Now, if we choose x less than 3, the LHS of this inequality comes out to be (–ve) (–ve) = + ve hence does not satisfy the inequality.

Also, if we choose x greater than 5, the LHS of this inequality comes out to be (+ve) (+ve) = +ve again does not satisfy this inequality.

But, if we take 3 < x < 5, the LHS becomes (+ve) (–ve) = –ve and hence satisfies the inequality. Also, x = 3 and x = 5 satisfy inequality. So, required domain of the given function is [3, 5].

Now, the number 7 does not lie in the domain of f so f(7) cannot be obtained for this function. However, 3, 4 and 5 all lie in the domain of f so values of f at 3, 4 and 5 can be obtained and are given as follows.

$$f(3)=\sqrt{(3-3)(5-3)}\ =\ \sqrt{0\times 2}\ =\ \sqrt{0}=0$$

$$f(4)=\sqrt{(4-3)(5-4)}\ =\ \sqrt{1\times 1}\ =\ 1$$

$$f(5)\ =\ \sqrt{(5-3)(5-5)}\ =\ \sqrt{2\times 0}=\sqrt{0}=0$$

Now, you can try the following Self Assessment Question.

SAQ 1

Find the domain and range of the function $f:\mathbb{R}\to\mathbb{R}$ given by

$f(x)=\dfrac{1}{x-2}$, $x\in\mathbb{R}$. Also, evaluate $f(1)$, $f(3)$, $f(-5)$.

1.5 SOME SPECIAL FUNCTIONS HAVING APPLICATIONS IN STATISTICS, PROBABILITY AND MACHINE LEARNING

In the previous section you have seen that function is a rule (satisfying two conditions mentioned in the box at page number 18) which associates the

elements of one set to the elements of another set. In different courses of the programme M.Sc. Applied Statistics (MSCAST) you will used many functions. Some well-known and commonly used functions are discussed in this section. Let us start with the easiest function known as constant function.

Constant Function

Suppose you have a special coin having head on both sides. Then whenever you will throw it, each time its outcome will be head. Ignore the possibility of landing on edge like Sholay film. If X is a variable having value 1 if head appears and 0 otherwise then X will always take value 1. Such type of behaviour is represented by a constant function in mathematics and **degenerate random variable** in probability theory. In mathematics constant function is defined as follows:

Let $X, Y \subseteq \mathbb{R}$, then a function $f : X \to Y$ is said to be a constant function if it associates each element of its domain to a single number in its co-domain and in general is defined as

$$f(x) = a, \quad \forall\, x \in \mathbb{R}, \text{ where a is a real constant.} \qquad \ldots (1.18)$$

i.e., a function is constant if its range is a singleton set.

For example,

(i) $f : \mathbb{N} \to \mathbb{N}$ defined by

$f(x) = 3, \qquad n \in \mathbb{N}$

is a constant function because all elements of the domain are associated to the single element 3 and its pictorial presentation is shown in the Fig. 1.9 while graphical presentation of this function is shown in Fig. 1.10 (a).

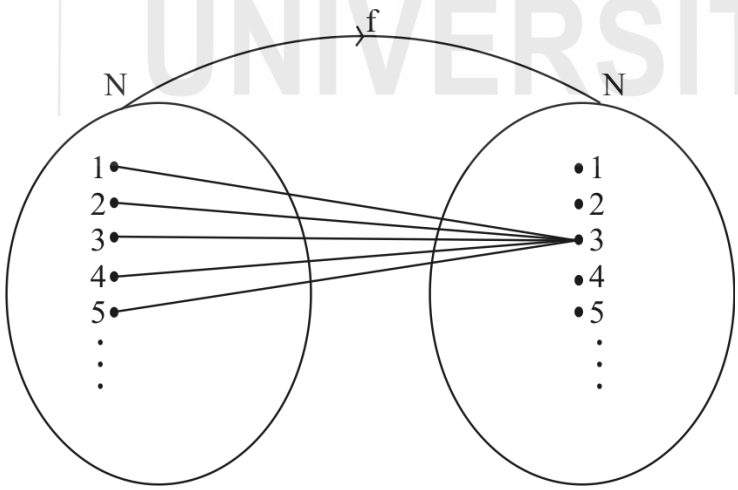
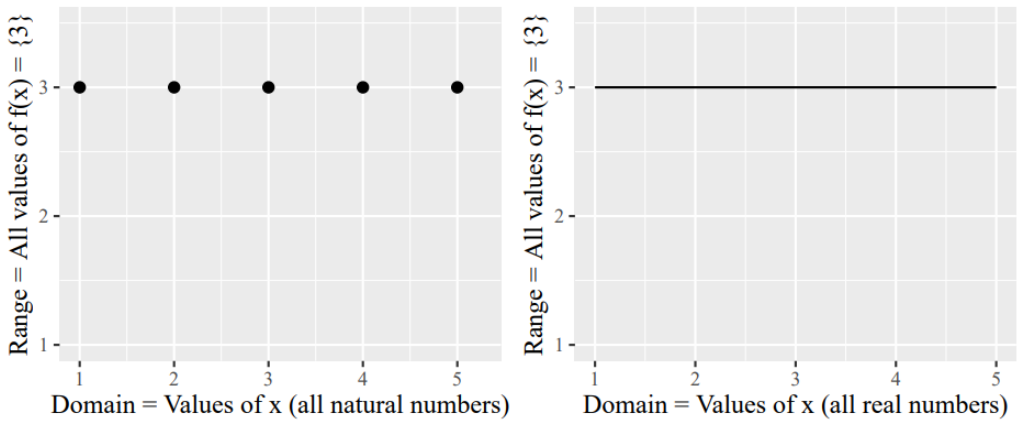


Fig. 1.9: Pictorial presentation of the constant function f(x) = 3

(ii) $f : \mathbb{R} \to \mathbb{R}$ defined by

$f(x) = 3, \; x \in \mathbb{R}$

is also a constant function and its graph is shown in Fig. 1.10 (b).



(a) **(b)**

Fig. 1.10: Graph of the constant function f(x) = 3 (a) when f is from $\mathbb{N}$ to $\mathbb{N}$ (b) when f is from $\mathbb{R}$ to $\mathbb{R}$

Remark 3

If you connect this function with the example of daughters and mothers then you find that all daughters have 3 as their mothers. But then mother 3 has infinitely many daughters. In real life you will not find a mother who has infinite daughters. So, keep in mind that example of daughters and mothers is given just for better understanding of the two conditions for a rule to become a function. So, limit that example only for understanding the two conditions required for a rule to be a function. The next function that you are going to study is identity function. You will see that identity function associates each element with itself. But in real life a daughter cannot be mother of herself. So, again the same reason example of daughters and mothers is given just for better understanding of the two conditions for a rule to become a function. Keep this point in mind.

Identity Function

You know from school mathematics that when you multiply a real number with 1 outcome will be the same number itself. That is, for any real number a, we have a.1 = 1.a = a. In the universe of functions identity function plays the same role as 1 plays in the family of real numbers. In mathematics identity function is defined as follows:

Let $X \subseteq \mathbb{R}$, a function $f : X \to X$ is said to be an identity function if it associates each element with itself and in general it is defined as

$$f(x) = x, \qquad x \in X \qquad \ldots (1.19)$$

For example,

(i) $f : \mathbb{N} \to \mathbb{N}$ defined by

$$f(n) = n, \qquad n \in N \qquad \ldots (1.20)$$

is an identity function because it associates each natural number with itself. Pictorial presentation of this function is shown in the Fig. 1.11.

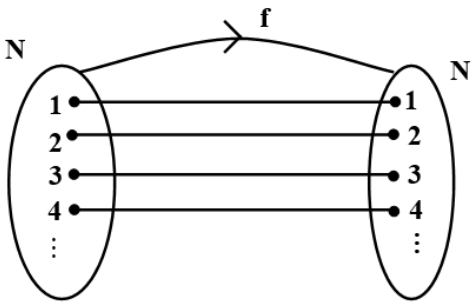


Fig. 1.11: Pictorial presentation of identity function in the world of natural numbers

(ii) Instead of set of natural numbers if you want to define identity function in the set of real numbers then it is defined the same way just replace $\mathbb{N}$ by $\mathbb{R}$ as follows:

$f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = x, \quad x \in \mathbb{R} \qquad \ldots (1.21)$$

Graphs of identity function in the family of natural numbers and in the family of real numbers are shown in Fig. 1.12 (a) and (b) respectively.

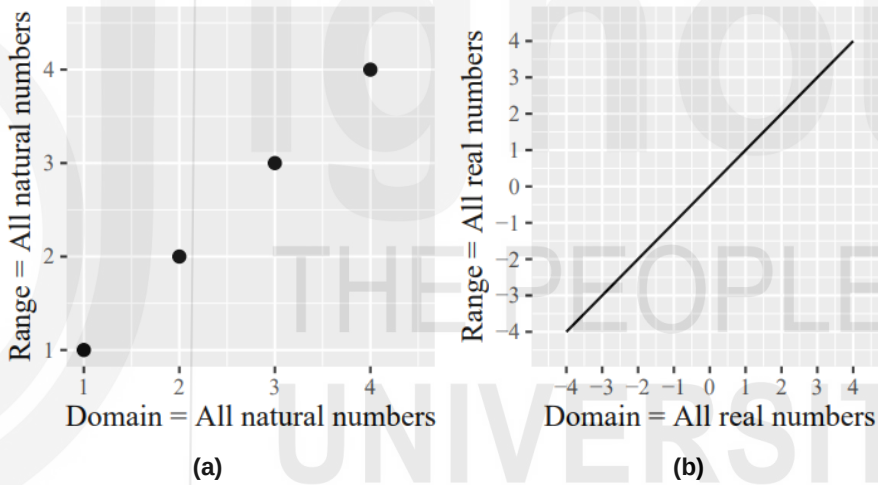


Fig. 1.12: Graph of the identity function f(x) = x (a) when f is from $\mathbb{N}$ to $\mathbb{N}$ (b) when f is from $\mathbb{R}$ to $\mathbb{R}$

Polynomial Function

A function $f : \mathbb{R} \to \mathbb{R}$ is said to be a polynomial function of degree n (only possible values of n are 0, 1, 2, 3, 4, …) if it is defined as

$$f(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n, x \in \mathbb{R} \qquad \ldots (1.22)$$

where $a_0, a_1, a_2, \ldots, a_{n-1}, a_n \in \mathbb{R}$, $a_0 \neq 0$ are constants

e.g., f(x) = $2x^3 + x^2 - x + 5$, is a polynomial function of degree 3. But the expression $f(x) = 2x^{3/2} + x^2$ is not a polynomial because the exponent of x is 3/2 in the first term which is not allowed in a polynomial. Similarly, the expression $f(x) = x + \frac{1}{x}$ is not a polynomial because in the second term exponent of x is -1 which is not allowed in a polynomial. Possible values of exponent in a polynomial are from family of whole numbers, i.e., 0, 1, 2, 3, ….

Zero of a Polynomial: If f(x) is a polynomial then those values of x which makes value of polynomial zero are known as zeros of the polynomial. That is if x = a is such that f(a) = 0 then we say that x = a is a zero of the polynomial.

For example, if $f(x) = x^2 - x - 12$ then $f(4) = 4^2 - 4 - 12 = 0$. So, x = 4 is a zero of the polynomial $f(x) = x^2 - x - 12$.

Multiplicity of a Zero of a Polynomial: Multiplicity of a zero tells us about the behaviour of the polynomial near zero. If $f(x) = (x - a)^m(x - b)^n(x - c)$ then we say that multiplicity of zeros a, b, c are m, n, 1 respectively.

For example, if $f(x) = (x - 3)^7(x - 5)^9(x - 13)$ then multiplicity of zeros 3, 5 and 13 are 7, 9 and 1 respectively.

You know that a polynomial of degree 0, 1, 2, 3, 4, 5 are known as **constant**, **linear**, **quadratic**, **cubic**, **quartic** and **quintic polynomial**. In statistics we use polynomials in regression analysis, time series analysis, etc. The shape of a polynomial is dictated by three things (i) degree of the polynomial (ii) sign of the leading coefficient (iii) multiplicity of the zeros of the polynomial. Let us visualise the effect of these three things through Fig. 1.13 (a) to (f) and Fig. 1.14 (a), (b) to Fig. 1.16 (a), (b) which will help you to identify the suitable polynomial which fits the given data. Fig. 1.13 (a) to (f) shows respectively graphs of the following polynomial functions:

$$f(x) = 2,\quad f(x) = -2,\quad f(x) = x + 1,\quad f(x) = -x + 1,$$
$$f(x) = (x + 1)(x - 1) = x^2 - 1,\quad f(x) = -x^2 - 1$$

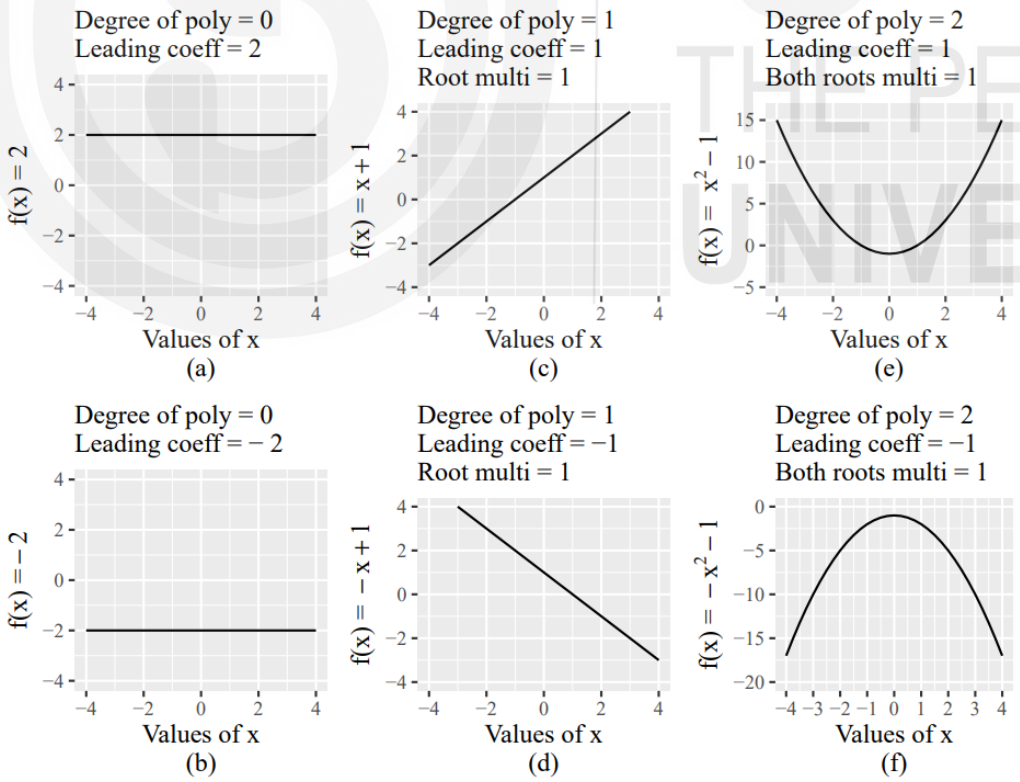


Fig. 1.13: Graphs of the different polynomials of degree 0, 1 and 2 with positive leading coefficient in (a), (c), (e) and negative leading coefficient in (b), (d), (f) and having root(s) multiplicity one

Polynomials shown in Fig. 1.14 (a), (b) to Fig. 1.16 (a), (b) are mentioned in the graphs. From the graphs of these polynomials note that:

- **Even Degree and Positive Leading Coefficient**: End behaviour of the curve on both sides is upward. Refer Fig. 1.16 (a). … (1.23)

- **Even Degree and Negative Leading Coefficient**: End behaviour of the curve on both sides is downward. Refer Fig. 1.16 (b). Note that in the case of even degree either both ends will be upward or downward. … (1.24)

- **Odd Degree and Positive Leading Coefficient**: End behaviour of the curve will be left downward but right upward. Refer Fig. 1.14 (a) and Fig. 1.15 (a). … (1.25)

- **Odd Degree and Negative Leading Coefficient**: End behaviour of the curve will be left upward but right downward. Refer Fig. 1.14 (b) and Fig. 1.15 (b). Note that in the case of odd degree both ends behave opposite to each other, i.e., if one end is upward then another will be downward. … (1.26)

- **Effect of Even Multiplicity of a Zero**: If multiplicity of a zero is even then graph of the polynomial touches the x-axis at that point. Refer Fig. 1.15 (a) and (b) both zeros at x = – 2 and 2 are of even multiplicity. So, graphs touch x-axis at both points. In Fig. 1.16 (a) and (b) two zeros at x = 2 and 3 are of even multiplicity. So, graphs touch x-axis at these points. … (1.27)

- **Effect of Odd Multiplicity of a Zero**: If multiplicity of a zero is odd then graph of the polynomial crosses the x-axis at that point. Further, if multiplicity is 1 then it crosses x-axis at that zero like a straight line (refer Fig. 1.14 (a) and (b) it crosses x-axis at three zeros. Also, in Fig. 1.16 (a) and (b) it crosses x-axis at x = 3.5 but if multiplicity is odd other than 1 like 3, 5, 7, etc. then it crosses x-axis at that zero such that it tries to flat at zero. Refer Fig. 1.16 (a) and (b) at the zero x = 1. … (1.28)

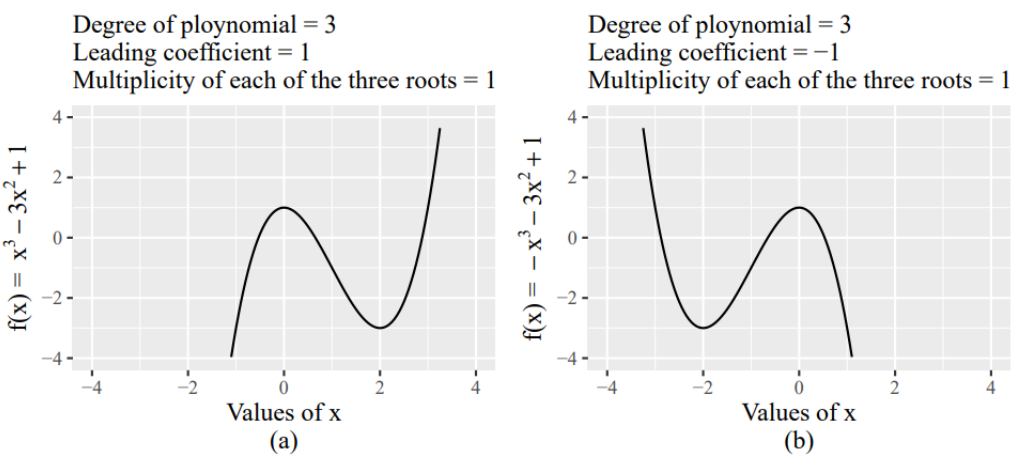


Fig. 1.14: Graphs of the cubic polynomials (a) with positive leading coefficient, and (b) negative leading coefficient

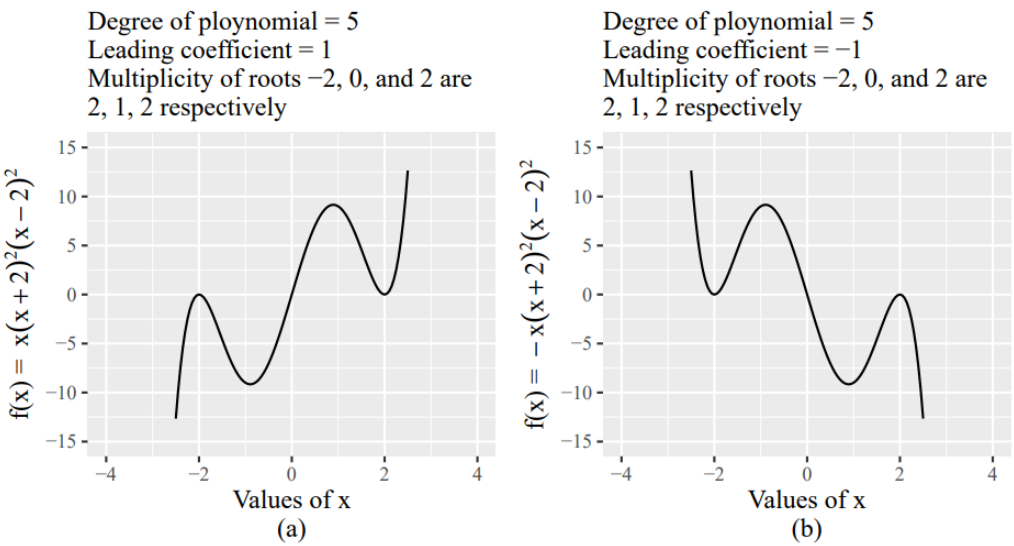


Fig. 1.15: Graphs of the polynomials of degree 5 (a) with positive leading coefficient, and (b) negative leading coefficient

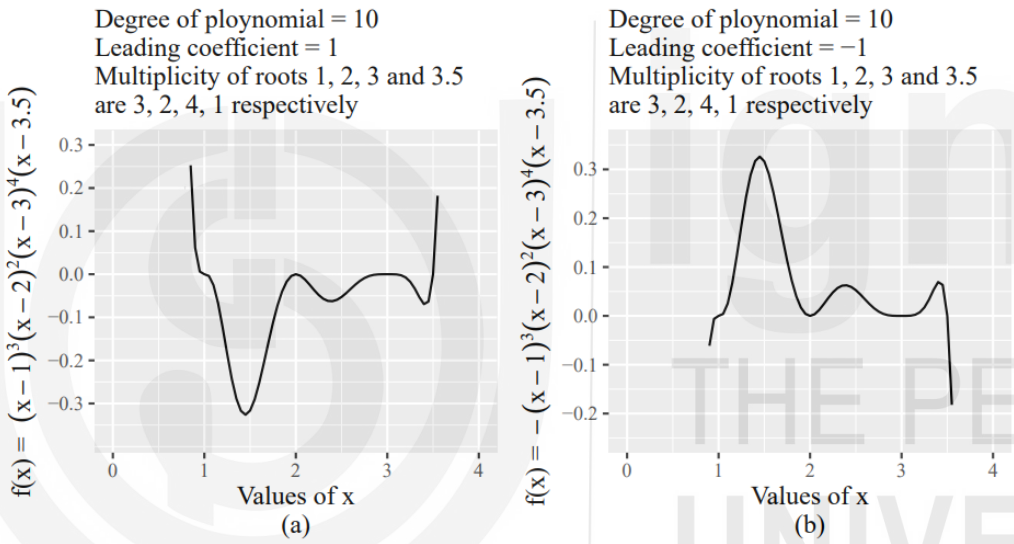


Fig. 1.16: Graphs of the 10-degree polynomials having four roots 1, 2, 3 and 3.5 with multiplicities 3, 2, 4, 1 respectively (a) with positive leading coefficient, and (b) negative leading coefficient

Exponential Function

From school mathematics you know that if you invest Rs P at r% per annum for a period of n years compounded annually then amount A after n years is given by

$$A = P\left(1+\frac{r}{100}\right)^n$$

Suppose P = Rs 1, r = 100% per annum then formula of amount reduces to

$A = 2^n$, where n is time and 2 represents growth rate.

So, if you invest Rs 1 for a period of 4 years and growth rate is 2 then after 4 years your amount will be $A = 2^4 = \text{Rs } 16$. Such type of functions where we have a constant in the base and variable (here time) in the power are known

as exponential functions. Mathematically, an exponential function is defined as follows:

A function $f:\mathbb{R}\to\mathbb{R}$ defined by

$$f(x)=a^x,\quad x\in\mathbb{R},\qquad a>0,\ a\neq 1 \qquad \ldots (1.29)$$

is called an exponential function.

i.e., the nature of an exponential function is of the form = $(\text{constant})^{\text{variable}}$

For example, $f(x)=2^x$, $g(x)=10^x$, $h(x)=\left(\frac{1}{2}\right)^x$, etc. all are exponential functions.

Domain and Range of the Exponential Function: In exponential function variable x (say) is sitting in exponent of a constant so function is defined for all real values of the variable x and hence **domain** of an exponential function is the set of all real numbers. Suppose $f(x)=2^x$ then note that values of this function will never be negative and zero for every real value of x

$$\left[\begin{array}{l}\because 2^0=1,\ 2^1=2,\ 2^2=4,\ 2^3=8,\ \text{etc.}\\ 2^{-1}=\frac{1}{2}=0.5,\ 2^{-2}=\frac{1}{2^2}=\frac{1}{4}=0.25,\ 2^{-3}=\frac{1}{8}=0.125,\ \text{etc.}\end{array}\right]$$

and hence its range is the set of positive real numbers $\mathbb{R}^+=(0,\infty)$. This is not only true for this particular exponential function but it is true in general for all the exponential functions so **range** of an exponential function is $\mathbb{R}^+=(0,\infty)$.

$$\ldots (1.30)$$

Graphical Behaviour of Exponential Function: Steepness of an exponential function is dictated by its rate. Further, if rate is < 1 then with increase in time its value will decay, and when its rate is > 1 then its value will increase with time. Fig. 1.17 (a) shows decay as a = 1/2 < 1 while Fig. 1.17 (b) shows growth as a = 2 > 1. Note that value of each exponential function at x = 0 is 1. So, graph of every exponential function will pass through the point (0, 1) so this special point is also indicated in the graph.

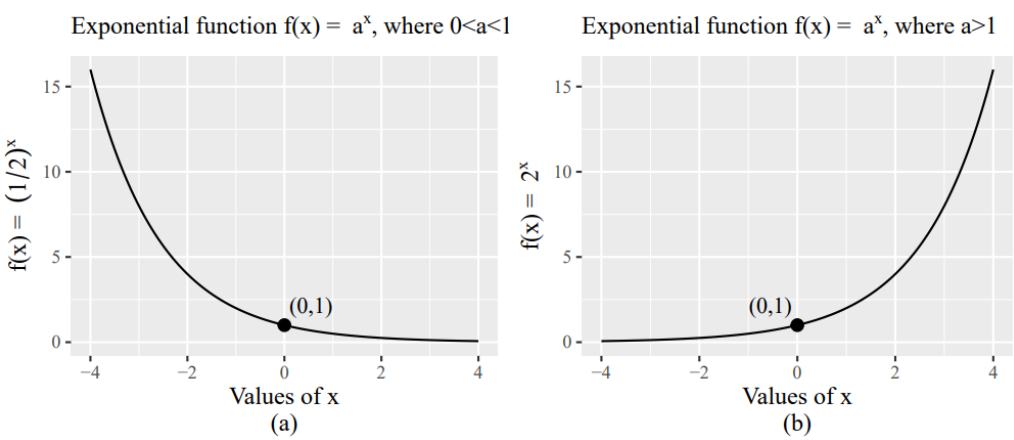


Fig. 1.17: Graph of exponential function (a) f(x) = $(1/2)^x$ (b) f(x) = 2^x

Example 3: In the following exponential functions name the rate, time and amount.

(i) $A = 3^n$ (ii) $B = 10^m$ (iii) $a^k = x$ (iv) $8^{(4/3)} = 16$

Solution: We know that an exponential function can be interpreted as
$\text{Amount} = (\text{rate})^{\text{time}}$

Therefore, we have

(i) Amount = A, rate = 3, time = n

(ii) Amount = B, rate = 10, time = m

(iii) Amount = x, rate = a, time = k

(iv) Amount = 16, rate = 8, time = 4/3

Logarithm Function

In Sec. 1.4 you have seen that function can be treated as a machine (refer Fig. 1.5) which gives unique output corresponding to each input. You know that in an exponential function if rate is 10 and time is 3 then output (amount) will be 1000. If you are interested in another machine (function) which works in such a way that if output of an exponential function is given as input in this new machine, then its output should be corresponding input of the exponential function such a machine (function) is known as logarithm function. In this new machine rate of exponential function is known as base of the logarithm function. Table 1.1 explains this inverse relationship between exponential and logarithm functions and Fig. 1.18 shows both functions as machines simultaneously.

Table 1.1: **Exponential and logarithm functions are inverse of each other**

Exponential Function			**Logarithm Function**		
	Function as a machine			Function as a machine	
Rate	Input	Output	Base	Input	Output
2	5	$2^5 = 32$	2	32	$\log_2 32 = 5$
10	3	$10^3 = 1000$	10	1000	$\log_{10} 1000 = 3$
1/2	4	$(1/2)^4 = 1/16$	1/2	1/16	$\log_{1/2}(1/16) = 4$
2	-5	$2^{-5} = 1/32$	2	1/32	$\log_2(1/32) = -5$
10	-3	$10^{-3} = 1/1000$	10	1/1000	$\log_{10}(1/1000) = -3$
1/2	-4	$(1/2)^{-4} = 16$	1/2	16	$\log_{1/2} 16 = -4$
2	0	$2^0 = 1$	2	1	$\log_2 1 = 0$
e	x	e^x	e	e^x	$\log e^x = x$

Note that when output of an exponential function is given as input to the logarithm function then output of the logarithm function matches with corresponding input of the exponential function and vice versa. In the last row of the Table 1.1 the number e is a special number in mathematic. You have studied expansion of e^x in school mathematics and the same is given as follows.

$$e^x = 1 + \frac{x}{\lfloor 1} + \frac{x^2}{\lfloor 2} + \frac{x^3}{\lfloor 3} + \frac{x^4}{\lfloor 4} + \frac{x^5}{\lfloor 5} + \ldots \qquad \ldots (1.31)$$

So, value of e can be obtained just replacing x by 1 in this expansion of e^x and the same is given by the infinite series (you will study about infinite series in Sec. 4.4 of Unit 4 of this course)

$$e = 1 + \frac{1}{\lfloor 1} + \frac{1}{\lfloor 2} + \frac{1}{\lfloor 3} + \frac{1}{\lfloor 4} + \frac{1}{\lfloor 5} + \ldots \qquad \ldots (1.32)$$

You know that value of e up-to 3-decimal place is 2.718. In fact, it is an irrational number. In the course MSTL-001 you will see that this value of e can be obtained in R by typing exp(1) on R console and hitting enter.

Mathematically, logarithm function is defined as follows:

A function $f : \mathbb{R}^+ \to \mathbb{R}$ is said to be logarithm function if it is defined as

$$y = f(x) = \log_a x, \quad x \in \mathbb{R}^+ = \text{set of all positive real numbers} \qquad \ldots (1.33)$$

where a > 0 and a $\neq 1$

You have seen that if $a^m = n$ then in terms of logarithm we write it as $\log_a n = m$.

Domain and Range of the Logarithm Function: Since logarithm function is inverse of exponential function so range of exponential function will be domain of logarithm function and domain of exponential function will be range of the logarithm function. Hence, domain of the logarithm function is $\mathbb{R}^+ = (0, \infty)$ and range of the logarithm function is $\mathbb{R} = (-\infty, \infty)$. … (1.34)

Graphical Behaviour of Logarithm Function: If base is < 1 then logarithm function is decreasing and when its base is > 1 then it is an increasing function. Note that $\log_{10} 10 = 1, \log_{10} 100 = 2, \log_{10} 1000 = 3, \log_{10} 10000 = 4.$ See the difference between the input numbers (10, 100 = 10^2, 1000 = 10^3, 10000 = 10^4) and output numbers (1, 2, 3, 4). We observe that variation between the numbers 10, 100, 1000 and 10000 is huge compare to the variation of the output numbers 1, 2, 3 and 4. This is one of the silent features of logarithm function and due to this it facilitates graphical represents of the numbers with large variations. For example, if you want to compare the sizes of the populations of all Asian countries in the same plot then you should plot values of population after taking logarithm of population sizes of each country. Without logarithm plot will not be more informative due presence of two large populations figures of India and China. Fig. 1.19 (a) and (b) shows graphical behaviour of logarithm function when a = 1/2 < 1 and when a = 2 > 1. In the first case logarithm function is decreasing and in the second case it is

increasing. Like an exponential function there is a special point in the graph of logarithm function. Refer second last row of the Table 1.1 which says that $\log_2 1 = 0$. There is nothing special in base 2 you can take any base according to the definition of logarithm function its value at 1 will always be 0. So, graph of every logarithm function will pass through the point (1, 0) which is also shown in the graph.

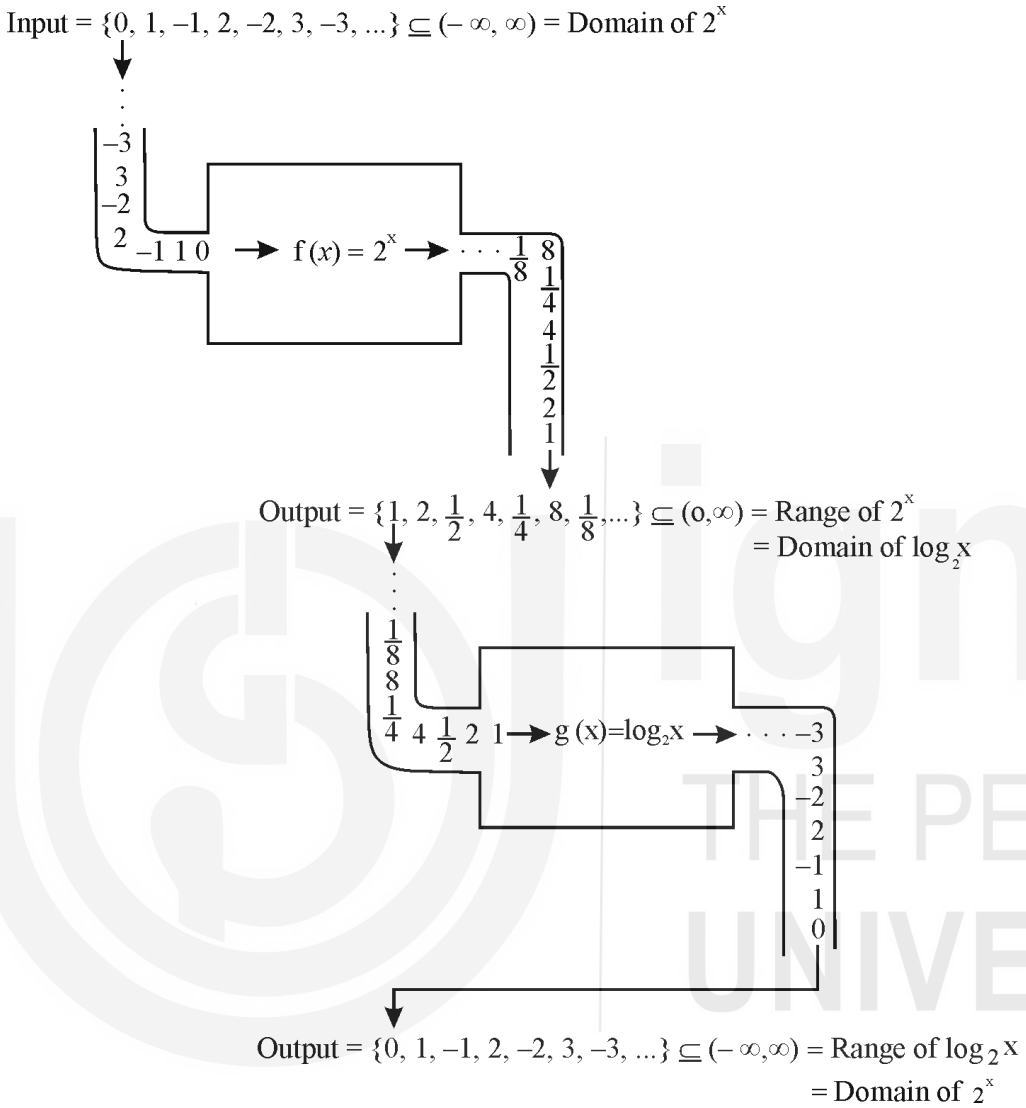


Fig. 1.18: Exponential and logarithm functions are machines which undo effect of each other

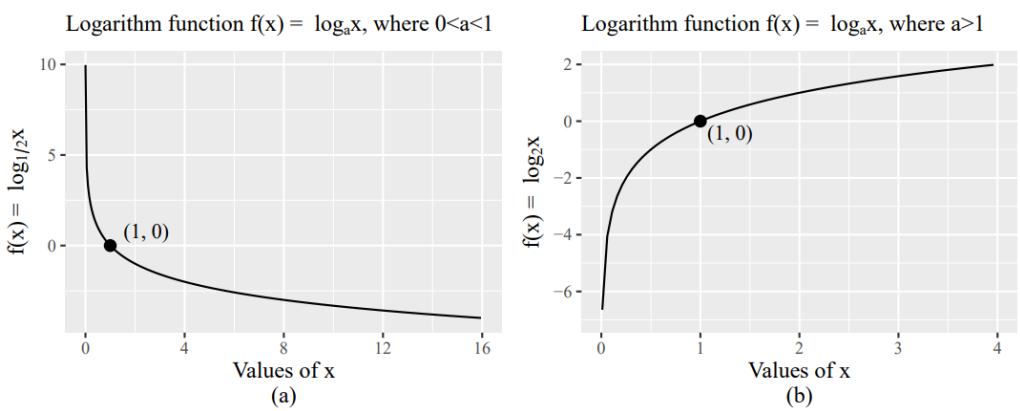


Fig. 1.19: Graph of logarithm function (a) f(x) = $\log_{1/2} x$ **(b) f(x) =** $\log_2 x$

Finally, following are some important laws of logarithm function

1. $\log_a mn = \log_a m + \log_a n$ … (1.35)

2. $\log_a \frac{m}{n} = \log_a m - \log_a n$ … (1.36)

3. $\log_a m^n = n\log_a m$ … (1.37)

4. $a^{\log_a m} = m$ … (1.38)

5. $\log_a a = 1$ … (1.39)

6. $\log_a b = \frac{1}{\log_b a}$ … (1.40)

7. $\log_a b = \frac{\log_n b}{\log_n a}$ this is known as base change formula, in fact we can take any base in place of n. … (1.41)

Remark 4

- If base of the logarithm is 10 then it is known as **common logarithm**.
- If base of the logarithm is e then it is known as **natural logarithm** and some time is written as ln x instead of log x.
- When we write log x it means base is e. That is, in most of the cases base is mentioned only when it is other than e.

Example 4: Solve for x in each part:

(i) $\log_2 64 = x$ (ii) $\log_4 64 = x$ (iii) $\log_8 64 = x$ (iv) $\log_{1/4} 64 = x$

(v) $\log_8 16 = x$ (vi) $\log_{\sqrt{3}} 81 = x$ (vii) $\log_{16} x = \frac{3}{2}$ (viii) $\log_2 x = -3$

Solution:

(i) $\log_2 64 = x \Rightarrow 2^x = 64 \quad \left[\because \log_a b = m \Rightarrow a^m = b\right]$
$\Rightarrow 2^x = 2^6 \Rightarrow x = 6 \quad \left[\because a^m = a^n \Rightarrow m = n\right]$

Similarly,

(ii) $\log_4 64 = x \Rightarrow 4^x = 64 \Rightarrow 4^x = 4^3 \Rightarrow x = 3$

(iii) $\log_8 64 = x \Rightarrow 8^x = 64 \Rightarrow 8^x = 8^2 \Rightarrow x = 2$

(iv) $\log_{1/4} 64 = x \Rightarrow (1/4)^x = 64 \Rightarrow (1/4)^x = 4^3$
$\Rightarrow (1/4)^x = (1/4)^{-3} \Rightarrow x = -3$

(v) $\log_8 16 = x \Rightarrow 8^x = 16 \Rightarrow (2^3)^x = 2^4 \Rightarrow 2^{3x} = 2^4 \Rightarrow 3x = 4 \Rightarrow x = \frac{4}{3}$

(vi) $\log_{\sqrt{3}} 81 = x \Rightarrow (\sqrt{3})^x = 81 \Rightarrow 3^{x/2} = 3^4 \Rightarrow \frac{x}{2} = 4 \Rightarrow x = 8$

(vii) $\log_{16} x = \frac{3}{2} \Rightarrow (16)^{3/2} = x \Rightarrow (4^2)^{3/2} = x \Rightarrow 4^3 = x \Rightarrow x = 64$

(viii) $\log_2 x = -3 \Rightarrow 2^{-3} = x \Rightarrow \frac{1}{2^3} = x \Rightarrow \frac{1}{8} = x \Rightarrow x = \frac{1}{8}$

Sigmoid Logistic Function

In the course MST-012 you will study about probability and random variables. One of the silent features of probability is that its value lies between 0 and 1. The value zero means impossible event and 1 means sure event. Output of sigmoid logistic function also lies between 0 and 1. So whenever there is a requirement to predict value as probability, sigmoid function is one of candidates you can think about. Because of this nature, in the course MST-026 we will use it in neural network as an activation function. Mathematically, sigmoid logistic function $\sigma : \mathbb{R} \to \mathbb{R}$ is defined by

$$\sigma(x) = \frac{1}{1+e^{-x}}, \quad x \in \mathbb{R} \qquad \ldots (1.42)$$

Domain and Range of the Sigmoid Logistic Function: Here x is sitting in exponent of e (constant) so function is defined for all real values of the variable x and hence **domain** of sigmoid logistic function is the set of all real numbers. That is domain is $(-\infty, \infty)$. Also note that e^{-x} being exponential function will always take positive values for all real values of x and hence the value of the expression $\frac{1}{1+e^{-x}}$ or $\frac{e^x}{e^x+1}$ will always lie between 0 and 1, i.e., $0 < \frac{1}{1+e^{-x}} < 1$ or $0 < \frac{e^x}{e^x+1} < 1$ and hence its range is $(0, 1)$. Range suggest that whenever you have to predict value as probability p (say) you can think about this function. … (1.43)

Graphical Behaviour of Sigmoid Logistic Function: Note that value of this function is very close to zero when x is – 5 or less. It approaches to 0 as x approaches to $-\infty$. At x = 0 its value is 0.5. It approaches to 1 as x approaches to ∞. Due to similarity of its shape like S it is known as Sigmoid function.

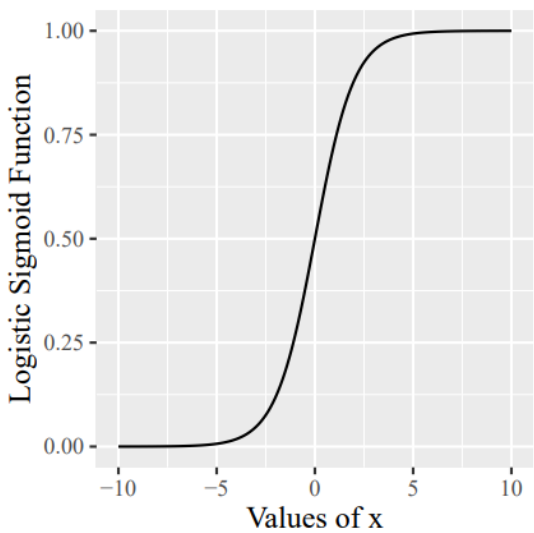


Fig. 1.20: Graph of sigmoid logistic function

Piecewise Function

Before defining it first we have to understand what we mean by partition of a set.

Partition of a Set: Let X be a non-empty set then a collection $C = \{A_1, A_2, A_3, \ldots, A_n\}$ of non-empty subsets of X is called a partition of X if it satisfies following two conditions:

(i) $X = A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n$

(ii) $A_i \cap A_j = \phi$ for $i \neq j$ … (1.44)

For example, if X = {1, 3, 7, 9, 12, 13} then the collection of the non-empty subsets $A_1 = \{1, 3, 7\}$, $A_2 = \{9\}$, $A_3 = \{12, 13\}$ of X form a partition of X. Another partition of X may be formed by the non-empty subsets $B_1 = \{1, 3,\}$, $B_2 = \{7, 9, 12, 13\}$ of X. Similarly, more partitions can be formed.

Partition of Domain of a Function: Let $D \subseteq \mathbb{R}$ and $f : D \to \mathbb{R}$ be a function. If the collection of sub-domains $D_1, D_2, D_3, \ldots, D_n$ satisfies $D = \bigcup_{i=1}^{n} D_i$ and $D_i \cap D_j = \phi$ for $i \neq j$ then we say that the collection containing the n sub-domains $D_1, D_2, D_3, \ldots, D_n$ partition the domain D of the function f. (1.45)

For example, suppose $f : [0, 5] \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x+2, & 0 \le x < 2 \\ 2x, & 2 \le x < 3 \\ x-1, & x = 3 \\ x^2, & 3 < x \le 5 \end{cases}$$

then $D = [0, 5]$, $D_1 = [0, 2)$, $D_2 = [2, 3)$, $D_3 = \{3\}$, $D_4 = (3, 5]$

The function defined like this in pieces (here we have four pieces) is known as **piecewise function**. Often sub-domains are mutually disjoint but, in some cases if $D_i \cap D_j \neq \phi$ for some i and j then values of ith and j^{th} pieces of the function at the common point is the same. For example, if in the present example D_1 also includes 2, i.e., $D_1 = [0, 2]$ then $D_1 \cap D_2 = \{2\}$. Now value of f obtained using the first piece and the second piece at x = 2 should match. This can be seen as follows: … (1.46)

Using the first piece: $f(2) = 2 + 2 = 4$ $\left[\because 2 \in [0, 2] \text{ so } f(x) = x + 2\right]$

Using the second piece: $f(2) = 2(2) = 4$ $\left[\because \text{ also } 2 \in [2, 3) \text{ so } f(x) = 2x\right]$

For Indian citizens having age less than 60 years and taxable income Rs x (rounded in rupees) then calculation of income tax for FY 2020-21 under old tax regime is a piecewise function given as follows:

$$f(x)=\begin{cases}0, & x\le 250000\\ 5\%\text{ of }(x-250000), & 250001\le x\le 500000\\ 12500+20\%\text{ of }(x-500000), & 500001\le x\le 1000000\\ 112500+30\%\text{ of }(x-1000000), & x\ge 1000001\end{cases}$$

There are some piecewise functions which are used in mathematics and statistics. Here, we will discuss only four of them namely: modulus function, unit step function (both having two pieces), Signum function (having three pieces), and greatest integer function (having infinite pieces).

Absolute Value Function or Modulus Function

If a machine behaves in such a way that its output is exactly the same number as was the input in the case its input is zero or positive number, and its output is negative times the input in the case input is a negative number, then behaviour of such a machine matches with absolute value function or modulus function. Mathematically, modulus function is defined as follows:

A function $f:\mathbb{R}\to\mathbb{R}^{+}\cup\{0\}$ defined by

$$f(x)=|x|=\begin{cases}x, & x\ge 0\\ -x, & x<0\end{cases},\ x\in R \qquad \ldots (1.47)$$

is called absolute value function.

Note that this function has two pieces one piece corresponding to non-negative values of x and other corresponding to negatives values of x.

Domain and Range of the Modulus Function: In modulus function neither x appears in denominator nor under the square root sign so it is defined for all real numbers and hence its domain is $\mathbb{R}=(-\infty,\infty)$. But this function does not give negative number as output so range of the modulus function is $\mathbb{R}^{+}\cup\{0\}=[0,\infty)$. … (1.48)

Graphical Behaviour of Modulus Function: Graph of modulus function is always of v shape. Graph of the function of the form |x – a| is always symmetrical about the vertical line x = a and vertex of the v is at x = a on the x-axis. Since values of modulus function is always ≥ 0 so its graph will never go below x-axis. Graphs of three modulus functions |x + 8|, |x| and |x – 8| are shown in Fig. 1.21.

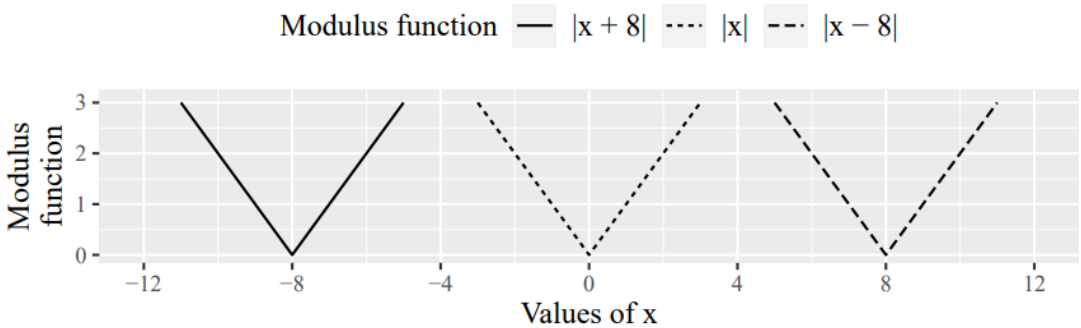


Fig. 1.21: Graph of three modulus functions |x + 8|, |x| and |x – 8|

Unit Step Function

If a machine behaves in such a way that its output is 1 in the case its input is zero or positive number, and its output is 0 in the case input is a negative number, then behaviour of such a machine matches with unit step function. Mathematically, unit step function is defined as follows:

A function $f : \mathbb{R} \to \{0, 1\}$ defined by

$$f(x) = U(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}, \quad x \in R \qquad \ldots (1.49)$$

is called unit step function.

Note that like modulus function this function also has two pieces one piece corresponding to non-negative values of x and other corresponding to negative values of x.

Domain and Range of the Unit Step Function: Unit step function is defined for all real values of x hence its domain is $\mathbb{R} = (-\infty, \infty)$. But this function takes only two values 0 and 1 so range of the unit step function is $\{0, 1\}$ … (1.50)

Graphical Behaviour of Unit Step Function: Unit step function takes only two values 0 and 1 so its graph is represented by two horizontal straight lines y = 0 and y = 1 and is shown in Fig. 1.22. Note the point why we have made open dot at the position (0, 0) and solid dot at the position (0, 1). Refer Fig. 1.4 where we have said about it.

Fig. 1.22: Graph of unit step function U(x)

Example 5: If $f(x) = |x| - |1 - x| - 3|2x - 5|$ then evaluate $f(2)$, $f(-2)$, $f(3)$, $f(-3)$, $f(1/2)$.

Solution:

$$f(2) = |2| - |1 - 2| - 3|2(2) - 5| = 2 - |-1| - 3|-1| = 2 - 1 - 3 = -2$$

$$\left[\begin{array}{l} \because 2 > 0, \text{ so by definition of modulus function } |2| = 2 \\ \text{Also } -1 < 0, \text{ so } |-1| = -(-1) = 1 \end{array}\right]$$

Similarly,

$$f(-2) = |-2| - |1 + 2| - 3|2(-2) - 5| = 2 - |3| - 3|-9| = 2 - 3 - 27 = -28$$

$$f(3) = |3| - |1-3| - 3|2(3)-5| = 3 - |-2| - 3|1| = 3 - 2 - 3 = -2$$

$$f(-3) = |-3| - |1+3| - 3|2(-3)-5| = 3 - |4| - 3|-11| = 3 - 4 - 33 = -34$$

$$f\left(\frac{1}{2}\right) = \left|\frac{1}{2}\right| - \left|1-\frac{1}{2}\right| - 3\left|2\left(\frac{1}{2}\right)-5\right| = \frac{1}{2} - \left|\frac{1}{2}\right| - 3|-4| = \frac{1}{2} - \frac{1}{2} - 12 = -12$$

Example 6: Solve the equation $|1-2x| = 5$.

Solution:

$|1-2x| = 5$

$\Rightarrow 1-2x = \pm 5$ $\left[\because \text{ if } |x| = a \Rightarrow x = \pm a \text{ refer (6.92) in Unit 6}\right]$

$\Rightarrow 1-2x = 5 \quad \text{or} \quad 1-2x = -5$

$\Rightarrow -2x = 5-1 \quad \text{or} \quad -2x = -5-1$

$\Rightarrow -2x = 4 \quad \text{or} \quad -2x = -6$

$\Rightarrow x = -2 \quad \text{or} \quad x = 3$

Signum Function

The piecewise function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \text{sgn}(x) = \begin{cases} -1, & x < 0 \\ 0, & x = 0 \\ 1, & x > 0 \end{cases} \qquad x \in \mathbb{R} \qquad \ldots (1.51)$$

is called signum function.

Sometimes this function is also defined as follows:

$$f(x) = \text{sgn}(x) = \begin{cases} \dfrac{|x|}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases} \qquad \ldots (1.52)$$

Note that this function has three pieces one piece corresponding to negative values of x, one piece corresponding to x = 0 and third piece corresponding to positive values of x.

Domain and Range of the Signum Function: This function is defined for all real values of x and hence its domain is $\mathbb{R} = (-\infty, \infty)$. But this function takes only three values -1, 0 and 1 so range of the signum function is $\{-1, 0, 1\}$. … (1.53)

Graphical Behaviour of Signum Function: Being piecewise function it has three pieces two of them are lines parallel to x-axis and third piece is a single point at origin. Its graph is shown in Fig, 1.23 (a). Note that here we have made two open dots one at the points (0, 1) and other at the point (0, -1) and one solid dot at the point (0, 0), i.e., at the origin because at x = 0 function takes value 0 not 1 or $-$ 1.

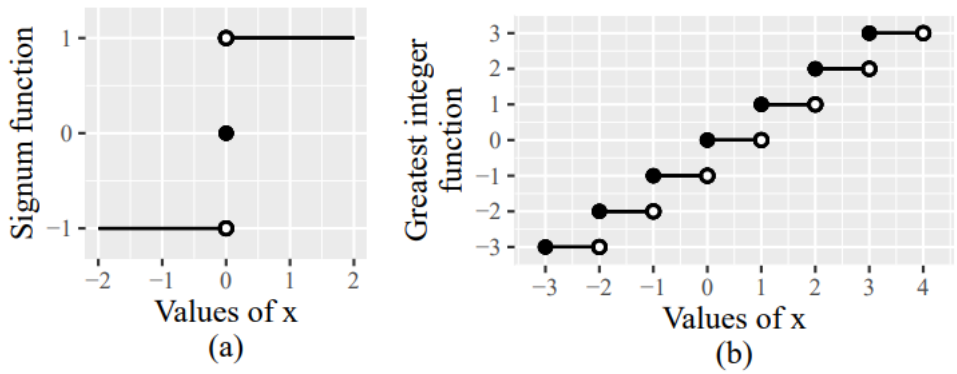


Fig. 1.23: Graph of (a) Signum function (b) Greatest integer function

Greatest Integer Function

If for a given real number x, we have a machine which gives output as the greatest integer less than or equal to x then such a machine is known as greatest integer function in mathematics. Greatest integer function is denoted by [x]. Thus, the piecewise function $f:\mathbb{R}\to\mathbb{R}$ defined by

$$f(x)=[x]=\begin{cases}\vdots & \\ -3, & -3\le x<-2\\ -2, & -2\le x<-1\\ -1, & -1\le x<0\\ 0, & 0\le x<1\\ 1, & 1\le x<2\\ 2, & 2\le x<3\\ 3, & 3\le x<4\\ \vdots & \end{cases} \qquad \ldots (1.54)$$

is called greatest integer function.

Note that this function has infinite pieces where value in each piece is an integer. In general, if n is an integer and n ≤ x < n + 1 then [x] = n.

Domain and Range of the Greatest Integer Function: This function is defined for all real values of x and hence its domain is $\mathbb{R}=(-\infty,\infty)$. Since this function takes only integral values so range of the greatest integer function is $\mathbb{Z}=\{\ldots,-3,-2,-1,0,1,2,3,\ldots\}$. … (1.55)

Graphical Behaviour of Greatest Integer Function: We have seen that it is a piecewise function having infinite number of pieces. Its graphical behaviour is shown in Fig. 1.23 (b). Consider the interval [1, 2], note that here we have made solid dot at the point (1, 1) because at x = 1 function takes value 1. But at the point (2, 1) we have made an open dot because at x = 2 function takes value 2 not 1 hence solid dot at the point (2, 2) and open dot at the point (2, 1). Similarly, graph is plotted in other intervals.

Note 3: In Unit 4 of the course MST-012 you will study cumulative distribution function (cdf) of a discrete random variable. There you will see that graphical behaviour of cdf of a discrete random variable is similar to this function. Height at each step will be dictated by the probability at the corresponding value of

the random variable. This is one of the reasons we discussed this function here so that understanding of cdf becomes easy for you.

Remark 5

- In the course MSTL-011 you will study **floor()** function in R. The floor() function in R is nothing but is simply greatest integer function that is discussed here.
- In the course MSTL-011 you will study another function namely **ceiling()** function in R. The ceiling() function in R gives us the smallest integer greater than or equal to x. It is generally denoted by $\lceil x \rceil$. For example, $\lceil 4.75 \rceil = 5$, $\lceil 4.999 \rceil = 5$, $\lceil 4.0001 \rceil = 5$, $\lceil -4.75 \rceil = -4$, $\lceil -0.15 \rceil = 0$, $\lceil -100.75 \rceil = -100$.

Example 7: Evaluate: (i) $\text{sgn}(-1.76)$ (ii) $\text{sgn}(3.89)$ (iii) $\text{sgn}(0)$ (iv) $\text{sgn}(5)$ (v) $[-1.76]$ (vi) $[3.89]$ (vii) $[0]$ (viii) $[5]$ (ix) $\lceil -1.76 \rceil$ (x) $\lceil 3.89 \rceil$ (xi) $\lceil 0 \rceil$ (xii) $\lceil 5 \rceil$

Solution: We know that

$$\text{sgn}(x) = \begin{cases} -1, & x < 0 \\ 0, & x = 0 \\ 1, & x > 0 \end{cases}, \quad [x] = \begin{cases} \vdots & \\ -3, & -3 \le x < -2 \\ -2, & -2 \le x < -1 \\ -1, & -1 \le x < 0 \\ 0, & 0 \le x < 1 \\ 1, & 1 \le x < 2 \\ 2, & 2 \le x < 3 \\ 3, & 3 \le x < 4 \\ \vdots & \end{cases}, \quad \lceil x \rceil = \begin{cases} \vdots & \\ -2, & -3 < x \le -2 \\ -1, & -2 < x \le -1 \\ 0, & -1 < x \le 0 \\ 1, & 0 < x \le 1 \\ 2, & 1 < x \le 2 \\ 3, & 2 < x \le 3 \\ 4, & 3 < x \le 4 \\ \vdots & \end{cases}$$

Therefore,

(i) $\text{sgn}(-1.76) = -1$ (ii) $\text{sgn}(3.89) = 1$ (iii) $\text{sgn}(0) = 0$ (iv) $\text{sgn}(5) = 1$

(v) $[-1.76] = -2$ (vi) $[3.89] = 3$ (vii) $[0] = 0$ (viii) $[5] = 5$

(ix) $\lceil -1.76 \rceil = -1$ (x) $\lceil 3.89 \rceil = 4$ (xi) $\lceil 0 \rceil = 0$ (xii) $\lceil 5 \rceil = 5$

Even and Odd Functions

Even Function: If a function is such that it takes the same value at x and – x then it is known as even function. So, even function is defined as follows.

A function $f(x)$ is said to be even function if it satisfies

$f(-x) = f(x)$, for all points x of the domain of the function f. … (1.56)

For example,

(i) $f(x) = x^6 + x^4 + x^2$ is an even function since

$$f(-x) = (-x)^6 + (-x)^4 + (-x)^2 = x^6 + x^4 + x^2 = f(x)$$

(ii) $f(x) = |x|$

$f(-x) = |-x| = |(-1)(x)| = |(-1)||x| = (1)|x|$ as $|-1| = -(-1) = 1$

$= |x| = f(x)$

$\therefore$ modulus function is an even function

(iii) $f(x) = x^2 + x^3$

$f(-x) = (-x)^2 + (-x)^3 = x^2 - x^3 \neq f(x)$

$\therefore$ it is not an even function.

Odd Function: A function $f(x)$ is said to be odd function if it satisfies $f(-x) = -f(x)$, for all points x of the domain of the function f, i.e., the value of the function at – x becomes negative times the value of the function at x.

… (1.57)

For example,

(i) $f(x) = x^5 + x^3 + x$ is an odd function since

$$f(-x) = (-x)^5 + (-x)^3 + (-x) = -x^5 - x^3 - x = -\left(x^5 + x^3 + x\right) = -f(x)$$

(ii) $f(x) = \dfrac{1}{x^3}$ is also an odd function since

$$f(-x) = \frac{1}{(-x)^3} = \frac{1}{-x^3} = -\frac{1}{x^3} = -f(x)$$

(iii) $f(x) = x^2 + x^3$ is not an odd function since

$f(-x) = (-x)^2 + (-x)^3 = x^2 - x^3$

$-f(x) = -\left(x^2 + x^3\right) \neq f(-x)$

We see that if $f(x) = x^2 + x^3$ then neither $f(-x) = f(x)$ nor $f(-x) = -f(x)$

$\therefore f(x) = x^2 + x^3$ is neither even nor odd function.

Graphical Behaviour of Even and Odd Functions: Graph of an even function is symmetrical about y-axis or a line parallel to y-axis. Graph of odd function is symmetrical about origin and if you rotate its graph by 180^0 then it remains unchanged. Fig. 1.24 (a), (b) show graphs of even and odd functions respectively.

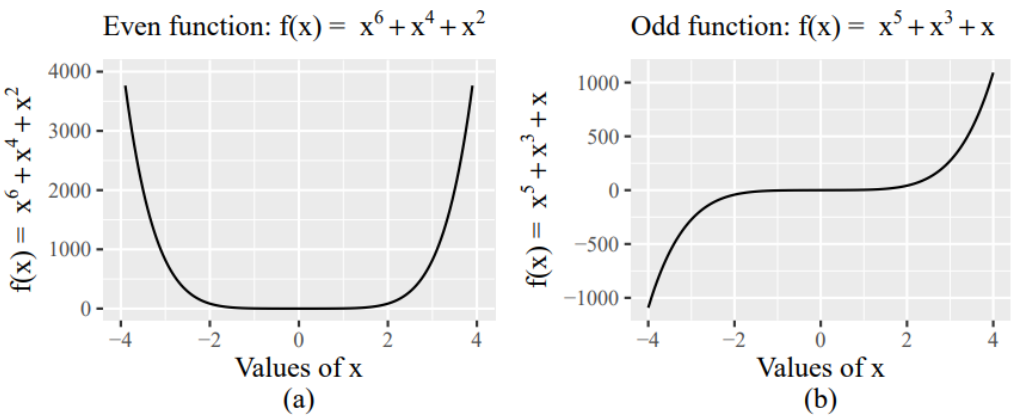


Fig. 1.24: (a) Graph of an even function (b) Graph of an odd function

Now, you can try the following Self-Assessment Question.

SAQ 2

If $f(x) = 5 - |x-3| + |x+1|$ then evaluate f(2), f(–2), f(6), f(–5), f(12).

1.6 SUMMARY

A brief summary of what we have covered in this unit is given as follows:

- In mathematics and statistics **quantity** means those things on which four basic mathematical operations addition, subtraction, multiplication and division can be applied.
- A quantity which remains same (unchanged) throughout a particular problem or distribution is known as a **constant quantity**. Those types of constants which always remain the same (unchanged) independent of the place and time are known as **fixed constant**. Those types of constants which remain same in one problem but may vary from problem to problem or distribution to distribution are known as **arbitrary constants** or **parameters.**
- A quantity which may change its value even in a particular problem is known as **variable**. If the nature of the quantity is such that its possible values can be written in a sequence then such a quantity is known as discrete in nature or **discrete variable**. If the nature of the quantity is such that it can take any possible value between two certain limits, then such a quantity is known as continuous in nature or **continuous variable**.
- Let $\mathbb{R}$ be the set of all real numbers. Then a set $I \subseteq \mathbb{R}$ is said to be an **interval** if whenever $a, b \in I$ and $a < x < b$ then $x \in I$.
- An interval $I \subseteq R$ with end points a and b (a < b) is called **open interval** if it contains all real numbers between a and b but does not contain end points a and b. If the interval also includes end points a and b then it is called **closed interval**. If interval I includes only one end point b not the other end point a then it is known as **left open and right closed interval**. If interval I includes only one end point a not the other end point b then it is known as **left closed and right open interval**.
- **Length** of each of the intervals (a, b), [a, b], (a, b], [a, b) is defined as $b - a,\ a < b$.
- An interval is said to be **finite** if its length is finite. An interval is said to be **infinite** if its length is not finite.
- Let X and Y be two non-empty sets. Then a rule which associates each element of X to a unique element of Y is called a **function** from X to Y. X is called **domain** of the function. Y is called **co-domain** of the function and the set of only those values of Y for which function is defined is called **range** of the function. That is, subset

$\{y \in Y : y = f(x)$ for some $x \in X\}$ of Y is called range of the function.

- If $f : X \to Y$ is a function given by $y = f(x)$ then x is known a **pre image** of y and y is known as **image** of x.
- If $y = f(x)$ is a function then values of y depend on values of x. So, y is known as **dependent variable** and x is known as **independent variable**.
- In different courses of the programme M.Sc. Applied Statistics (MSCAST) you will use many functions. Some well-known and commonly used functions are discussed in this unit such as constant, identity, polynomial, exponential, logarithm, sigmoid, modulus, signum, greatest integer, even and odd functions.

1.7 TERMINAL QUESTIONS

1. Find the domain and range of the function $f : \mathbb{R} \to \mathbb{R}$ given by

   $f(x) = \dfrac{1}{\sqrt{x+5}}$, $x \in \mathbb{R}$. Also, evaluate $f(0), f(4), f(-1), f(-5)$ if possible.

2. If $5^4 = 625$ then express it in logarithm form.

3. If $f(x) = \dfrac{1}{1+e^{-x}}$ then evaluate $f(0), f(-\log x), f(\log x)$.

4. If $f(x) = (x+4)^3(x-3)^4(x+6)^5(x-5)^6$ then which xero has multiplicity equal to 5.

5. If $f(x) = (x+4)^3(x-3)^4(x+6)^5(x-5)^6$ then which xero has multiplicity equal to 4.

1.8 SOLUTIONS/ANSWERS

Self-Assessment Questions (SAQs)

1. If we put x = 2 in the given function then denominator becomes zero and hence function is not defined at x = 2. Other than 2 it is defined for all real values of x. That is function takes real values for all real values of x except x = 2. Hence domain of the function f is $\mathbb{R} \setminus \{2\}$.

   We note that f(x) cannot be zero for any value of x in its domain. However, other than 0, f(x) can assume any real value so it's range is $\mathbb{R} \setminus \{0\}$

   Now, $f(1) = \dfrac{1}{1-2} = \dfrac{1}{-1} = -1$, $f(3) = \dfrac{1}{3-2} = \dfrac{1}{1} = 1$ and

   $f(-5) = \dfrac{1}{-5-2} = \dfrac{1}{-7} = -\dfrac{1}{7}$

2. $f(x) = 5 - |x-3| + |x+1|$

$$f(2) = 5 - |2-3| + |2+1| = 5 - |-1| + |3|$$
$$= 5 - (-(-1)) + 3 \quad [\because -1 < 0, \text{ so } |-1| = -(-1), \text{ but } 3 > 0, \text{ so } |3| = 3]$$
$$= 5 - 1 + 3 = 7$$

Similarly,

$$f(-2) = 5 - |-2-3| + |-2+1| = 5 - |-5| + |-1|$$
$$= 5 - (-(-5)) + (-(-1)) = 5 - 5 + 1 = 1$$

$$f(6) = 5 - |6-3| + |6+1| = 5 - |3| + |7| = 5 - 3 + 7 = 9$$

$$f(-5) = 5 - |-5-3| + |-5+1| = 5 - |-8| + |-4|$$
$$= 5 - (-(-8)) + (-(-4)) = 5 - 8 + 4 = 1$$

$$f(12) = 5 - |12-3| + |12+1| = 5 - |9| + |13| = 5 - 9 + 13 = 9$$

Terminal Questions

1. First of all x + 5 is under the square root so value of $\sqrt{x+5}$ will be real if $x + 5 \geq 0$. But $\sqrt{x+5}$ lies in the denominator so it cannot be zero. Hence, given function will assume real values only when x + 5 > 0. So, domain of the given function is $(-5, \infty)$.

   We know that value of a square root function is always ≥ 0. So, value of $\sqrt{x+5}$ will be ≥ 0. Also note that the function $\frac{1}{\sqrt{x+5}}$ cannot assume zero value for any value of x in its domain. Hence, range of the function is $(0, \infty)$.

   Now,

$$f(0) = \frac{1}{\sqrt{0+5}} = \frac{1}{\sqrt{5}} = \frac{\sqrt{5}}{5}, \; f(4) = \frac{1}{\sqrt{4+5}} = \frac{1}{\sqrt{9}} = \frac{1}{3}, \; f(-1) = \frac{1}{\sqrt{-1+5}} = \frac{1}{2}$$

   But value of the function at x = − 5 cannot be calculated because it is not a point in the domain of the function.

2. We know that if $a^m = b$ then $\log_a b = m$, therefore $5^4 = 625 \Rightarrow \log_5 625 = 4$

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

$$f(-\log x) = \frac{1}{1+e^{-(-\log x)}} = \frac{1}{1+e^{\log x}} = \frac{1}{1+x} \quad [\because e^{\log f(x)} = f(x)]$$

$$f(\log x) = \frac{1}{1+e^{-\log x}} = \frac{1}{1+e^{\log x^{-1}}} = \frac{1}{1+e^{\log\left(\frac{1}{x}\right)}} = \frac{1}{1+\frac{1}{x}} = \frac{x}{x+1}$$

4. Exponent of the factor (x + 6) is 5, hence required zero is obtained by solving x + 6 = 0 which gives x = − 6. Hence, multiplicity of the zero − 6 is 5.

5. Exponent of the factor $(x - 3)$ is 4, hence required zero is obtained by solving $x - 3 = 0$ which gives $x = 3$. Hence, multiplicity of the zero 3 is 4.

ignou
THE PEOPLE'S
UNIVERSITY

UNIT 2

TYPES, CONTINUITY AND DIFFERENTIABILITY OF FUNCTION OF A SINGLE VARIABLE

Structure

2.1 Introduction

Expected Learning Outcomes

2.2 Types of Function

2.3 Countable Set

2.4 Combining Functions

2.5 Continuity of a Function

2.6 Differentiability of a Function at a Point

2.7 Summary

2.8 Terminal Questions

2.9 Solutions/Answers

2.1 INTRODUCTION

In Unit 1 you have studied different types of intervals, the definition of a function of a single variable and its domain and range. We have also defined many functions by correlating them with daily life examples. A relationship between daughters and mothers explained both the conditions required for a rule to become a function. Here we have further classified functions known as types of a function in Sec. 2.2.

After understanding the meaning of one-one correspondence in Sec. 2.2 we will use this concept to define an important idea of a countable set in Sec. 2.3. Composition of two functions and algebra of functions are discussed in Sec. 2.4. Another two important concepts in mathematics that will be used in some courses of this programme are continuity and differentiability. You have already studied continuity and differentiability in earlier classes. A brief overview of these concepts is given in Secs. 2.5 and 2.6 respectively.

What we have discussed in this unit is summarised in Sec. 2.7. Self-Assessment Questions (SAQs) have been given in some sections which are generally based on the content discussed in that section. But to give you a

good practice of what we have discussed in this unit, some more questions based on the entire unit are given in Sec. 2.8 under the heading Terminal Questions. Due to the reason mentioned in Sec. 1.1 of Unit 1, solutions of all the SAQs and Terminal Questions are given in Sec. 2.9.

In the next unit, you will study the set function and the distance function.

Expected Learning Outcomes

After completing this unit, you should be able to:

- explain one-one, onto and one-to-one correspondence with their geometrical interpretation;
- define the concept of countable and uncountable sets;
- explain the idea of the composition of two functions and the algebra of functions;
- define the continuity of a function at a point and will be able to check whether the given function is continuous or not at a given point; and
- define the differentiability of a function at a point and will be able to check whether the given function is differentiable or not at a given point.

2.2 TYPES OF FUNCTION

In Unit 3 of the course MST-012 you will study the discrete random variable. To define a discrete random variable, we will use the world countable. So, to understand what we mean by the discrete random variable you should know the mathematical meaning of the word countable. But to understand the mathematical meaning of the word countable, you should have a good understanding of one-one correspondence or bijective map or one-one and onto function. This section is devoted to explain all these concepts.

One-One Function

If a machine gives different outputs for different inputs, then such a machine behaves like a one-one function. Mathematically, the one-one function is defined as follows.

A function $f : X \to Y$ is said to be 1-1 or **injective** function if distinct elements of X are associated with distinct elements of Y under f. … (2.1)

i.e., if $x_1, x_2 \in X$ and $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$ … (2.2)

or whenever $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$ … (2.3)

If we compare this definition with an example of daughters and mothers then one-one function means different daughters must have different mothers, i.e., there cannot be two daughters having the same mother for a function to be one-one.

For example, the function f shown in Fig. 2.1 (a) is not one-one because two different elements x_1, x_2 of X have the same image y_1. But the function f shown in Fig. 2.1 (b) is one-one because all the three elements of X have

distinct images in Y.

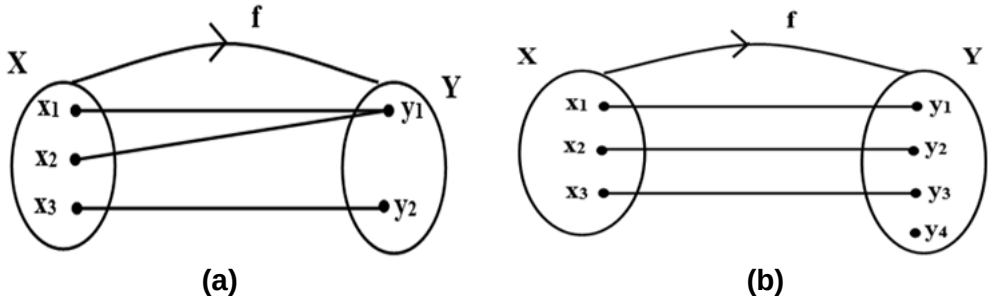


Fig. 2.1 (a) Not one-one function (b) A function which is one-one

Remark 1

If we want to show that a function f is one-one then do the following step. For $x_1, x_2 \in X$ take $f(x_1) = f(x_2)$ and show that $x_1 = x_2$.

For example,

(i) Show that the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 7x + 5$ is 1-1 function.

Solution: Let $x_1, x_2 \in \mathbb{R}$ be such that

$f(x_1) = f(x_2) \Rightarrow 7x_1 + 5 = 7x_2 + 5 \Rightarrow 7x_1 = 7x_2 \Rightarrow x_1 = x_2$

$\Rightarrow f$ is 1-1 function

(ii) Check whether the function $f : R \to R$ defined by $f(x) = x^2$ is 1-1 or not.

Solution: $f(x) = x^2$

Let $x_1 = 2,\ x_2 = -2$ then $x_1 \neq x_2$

But $f(x_1) = f(2) = (2)^2 = 4$ and also $f(x_2) = f(-2) = (-2)^2 = 4$

$\therefore f(2) = f(-2)$, i.e., there exists x_1 and x_2 such that $f(x_1) = f(x_2)$ but $x_1 \neq x_2$

$\Rightarrow f$ is not 1-1 function.

Onto Function

A function $f : X \to Y$ is said to be onto or **surjective** if each element of Y has at least one pre image in X. … (2.4)

i.e., for each $y \in Y$, there exists at least one $x \in X$ such that $f(x) = y$ ... (2.5)

If we compare this definition with example of daughters and mothers then onto function means each mother must have at least one daughter.

For example, the function f shown in Fig. 2.2 (a) is not onto because $y_4 \in Y$ but there is no $x \in X$ such that $y_4 = f(x)$. However, the function f shown in Fig. 2.2 (b) is onto because each element of Y has at least one pre image in X, i.e., y_1 has two pre-images and y_2 has three pre-images.

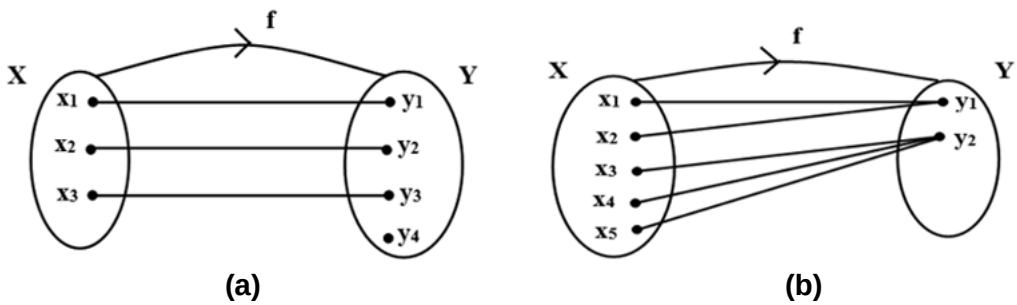


Fig. 2.2 (a) Not onto function (b) A function which is onto

Remark 2

If we want to show that a function $f(x)$ is onto then first we take an element y in Y and we have to show that there exists an element x is X such that $f(x) = y$

For example, show that the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 7x + 5$ is an onto function.

Solution: Here X = $\mathbb{R}$, Y = $\mathbb{R}$

Let $y \in Y = \mathbb{R}$, then $\frac{y-5}{7} \in X = \mathbb{R}$ be such that $\left[\begin{array}{l} \text{Why } \frac{y-5}{7} \text{ because} \\ y = f(x) = 7x + 5 \\ \Rightarrow x = \frac{y-5}{7} \end{array}\right]$

$$f\left(\frac{y-5}{7}\right) = 7\left(\frac{y-5}{7}\right) + 5 = (y-5) + 5 = y$$

Therefore, we have proved that for each $y \in Y = \mathbb{R}$ there exists $\frac{y-5}{7} \in X = \mathbb{R}$ such that $f\left(\frac{y-5}{7}\right) = y$

Hence, f is an onto function.

One-One and Onto Function

A function $f : X \to Y$ is said to be one-one and onto or **bijective** or **one-one correspondence** if f is both one-one as well as onto. … (2.6)

If we compare this definition with example of daughters and mothers then one-one and onto function means each mother has exactly one daughter and there is no mother who does not have any daughter. The function f shown in Fig. 2.3 represents a situation of one-one and onto function from X = {1, 2, 3, 4} to Y = {2, 3, 4, 5} where f(x) = x + 1, x ∈ X.

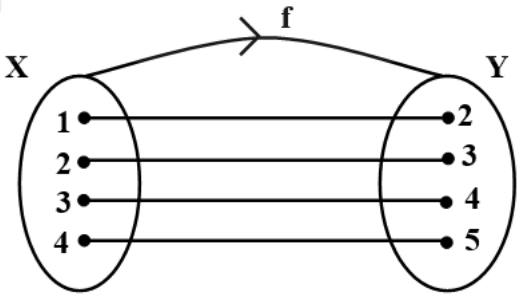


Fig. 2.3: One-one correspondence

Also, the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$f(x) = 7x + 5, \quad x \in \mathbb{R}$

is both one-one and onto (already shown) and hence it is a one-one correspondence between set of all real numbers $\mathbb{R}$ to itself.

Remark 3

Definition of a function f from X to Y says that corresponding to each element of X there is a unique element in Y. But one-one correspondence establishes

this relationship from Y to X also, i.e., corresponding to each element of Y there is a unique element in X. Application of one-one correspondence lies in the fact that once we have proved one-one correspondence between two sets X and Y then:

- you can say that number of elements in the two sets X and Y are equal in the sense that either both are finite or both are infinite sets.
- you know that multiplicative inverse of a non-zero number a is 1/a. For example, multiplicative inverse of 5 is 1/5. Another advantage of one-one correspondence is that we can talk about inverse of the function. Recall that exponential and logarithm functions are inverse of each other explained in Sec. 1.5 refer (1.29) and (1.33) of the previous unit.

Geometrical Meaning of Injective, Surjective and Bijective Functions

To explain these concepts geometrically we will use graphs of the following four functions:

The function $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x)=-x^2$ its graph is shown in Fig. 2.4 (a).

The function $f:\mathbb{R}_0^+\to\mathbb{R}_0$ defined by $f(x)=\frac{1}{x}$ its graph is shown in Fig. 2.4 (b), where $\mathbb{R}_0$ represents set of all real numbers other than zero and $\mathbb{R}_0^+$ represents set of all positive real numbers.

The function $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x)=x$ its graph is shown in Fig. 2.4 (c).

The function $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x)=(x-1)^2(x-3)$ its graph is shown in Fig. 2.4 (d).

Identification of One-One Function using Horizontal Line Test

If a function is 1-1 then geometrically it will satisfy the following condition.

Each horizontal line either does not intersect the graph of the function at all or if it intersects it will intersect exactly at one point. In other words, there should be no horizontal line which intersects the graph of the function at two or more than two points for a function to be one-one. … (2.7)

For example,

(i) Graph shown in Fig. 2.4 (b) is the graph of a one-one function because there is no horizontal line which intersects the graph at more than one point. In fact, each horizontal line above the x-axis intersects the graph at exactly one point. X-axis itself is also a horizontal line and it will not intersect the graph of this function, no problem, because this function will never take zero value. So, it is a one-one function over its domain.

(ii) Graph shown in the Fig. 2.4 (c) is also the graph of a one-one function because each horizontal line intersects the graph exactly at one point.

(iii) But the graph shown in the Fig. 2.4 (a) is not the graph of a one-one function because if we draw any horizontal line below the x-axis, then it will intersect the graph at two points. Similarly, graph shown in Fig. 2.4 (d) is also not the graph of a one-one function because there exist horizontal lines which can intersect its graph at more than one points.

For example, horizontal line $f(x) = -1/2$ will intersect its graph at three points.

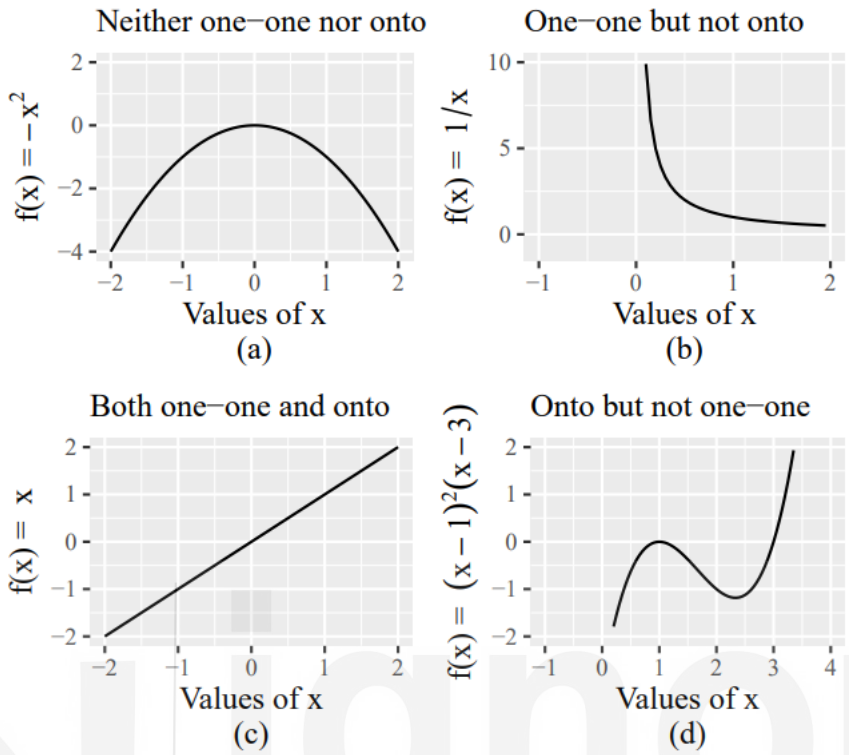


Fig. 2.4: Graphs of some functions to explain idea of one-one and onto functions geometrically

Identification of Onto or Surjective Function using Horizontal Line Test

If a function is onto then geometrically it will satisfy the following condition.

Horizontal line corresponding to each point of the codomain must intersect the graph of the function at least at one point. … (2.8)

For example,

(i) Graph shown in the Fig. 2.4 (d) is the graph of an onto function because each horizontal line intersects the graph of the function at least at one point. Graph shown in Fig. 2.4 (c) is also onto function because each horizontal line intersects the graph exactly at one point.

(ii) The function having graph shown in the Fig. 2.4 (b) is not an onto function because codomain of this function is whole real line but if we draw any horizontal line below the x-axis then it will not intersect the graph at all.

(iii) Similarly, the function having graph shown in the Fig. 2.4 (a) is not an onto function because codomain of this function is whole real line but if we draw any horizontal line above the x-axis then it will not intersect the graph at all.

Identification of Bijective Function or One-One Correspondence using Horizontal Line Test

If a function is both one-one and onto then it is known as one-one correspondence or bijective function. Therefore, by combining geometrical conditions of both one-one and onto functions we can say that a function will

be bijective if **each horizontal line** corresponding to each point of the codomain must intersect the graph of the function exactly at one point. … (2.9)

For example,

(i) The function having graph shown in the Fig. 2.4 (c) is the graph of a one-one and onto function because each horizontal line intersects the graph of the function exactly at one point. So, it is both one-one as well as onto function and therefore it is a bijective function or one-one correspondence.

(ii) The functions having graph shown in the Figs. 2.4 (a) and (b) are not the graph of bijective functions because they are not onto. We have already discussed why they are not onto.

(iii) Similarly, the function having graph shown in the Fig. 2.4 (d) is not the graph of a bijective function because it is not one-one. We have already discussed why it is not one-one.

Total Number of 1-1, Onto and Bijective Functions

Suppose $A, B \subset \mathbb{R}$ are two finite subsets where n(A) = m, n(B) = n then let us try to count total number of functions from A to B and how many of them are 1-1, onto and bijective functions.

Total Number of Functions: By definition of a function each element of A must be associated to a unique element of B. So, each element of A has n options to associate an element in B. Therefore, by fundamental principle of multiplication in counting we have n^m functions in total from A to B. … (2.10)

Total Number of One-One Functions: By definition of one-one function different elements of A must be associated to different elements of B. So, to define one-one function from A to B first of all we should have m ≤ n because if m > n then one-one function cannot be defined. Hence, in the case m > n total number of 1-1 functions from A to B is zero. In the case m ≤ n the first element of A can be associated to any of the n elements so has n options. Second element of A has n – 1 option available in B to associate. Similarly, third element has n – 2 options. Finally, m^{th} element of A has n – m + 1 options to get associate in B. Hence, using fundamental principle of multiplication in counting we have n(n – 1)(n – 2)…(n – m + 1) one-one functions in total from A to B. In the case m = n the total number of one-one functions from A to B will be n(n – 1)(n – 2) … 3.2.1, i.e., n!. … (2.11)

Total Number of Onto Functions: By definition of onto function each element of B has at least one pre-image in A. So, to define onto function from A to B first of all we should have m ≥ n because if m < n then onto function cannot be defined. Hence, in the case m < n total number of onto functions from A to B is zero. In the case m ≥ n instead of counting number of onto functions we will count the number of functions which are not onto and then subtract it from the total number of functions to get the number of onto functions. That is here we are going to use the fundamental principle of subtraction for counting. Suppose $B = \{b_1, b_2, b_3, \ldots, b_n\}$ then the function will not be onto in the cases

- either any one of b_i's does not have pre-image or

- any two b_i's do not have pre-images or
- any three b_i's do not have pre-images and so on either
- any n – 1 b_i's do not have pre-images in set A.

But the cases any one b_i does not have pre-images also includes the cases where any two b_i's do not have pre-images in A. Similarly, the cases any two b_i do not have pre-images also includes the cases where any three b_i's do not have pre-images in A and so on. This is exactly the situation of principle of inclusion-exclusion which you have studied for two and three sets in set theory in school mathematics. This principle for n sets states that

$$n(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} n(A_i) - \sum_{i \neq j}^{n} n(A_i \cap A_j) + \sum_{i \neq j \neq k}^{n} n(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n-1} n(A_1 \cap A_2 \cap \ldots \cap A_n)$$

Using this principle in our case, number of functions which are not onto are given as follows.

$$\binom{n}{1}(n-1)^m - \binom{n}{2}(n-2)^m + \binom{n}{3}(n-3)^m - \binom{n}{4}(n-4)^m + \ldots + (-1)^{n-1}\binom{n}{n-1}(1)^m$$

Explanation of the first term is: Any one of $b_1, b_2, b_3, \ldots, b_n$ does not have pre-image in A implies there are $\binom{n}{1}$ ways. Now, total number of functions from A to B\{b_i} are $(n-1)^m$ so total number of functions where one b_i's does not have pre-image is $\binom{n}{1}(n-1)^m$. Similarly, other terms can be understood. Now, we use fundamental principle of subtraction of counting:

$$\begin{pmatrix}\text{Total number of onto}\\ \text{functions A to B}\end{pmatrix} = \begin{pmatrix}\text{Total number of}\\ \text{functions A to B}\end{pmatrix} - \begin{pmatrix}\text{Total number of functions}\\ \text{which are not onto from A to B}\end{pmatrix}$$

So, finally, total number of functions from A to B which are onto are given by

$$n^m - \left[\binom{n}{1}(n-1)^m - \binom{n}{2}(n-2)^m + \binom{n}{3}(n-3)^m - \binom{n}{4}(n-4)^m + \ldots + (-1)^{n-1}\binom{n}{n-1}(1)^m\right]$$

$$= n^m - \binom{n}{1}(n-1)^m + \binom{n}{2}(n-2)^m - \binom{n}{3}(n-3)^m + \binom{n}{4}(n-4)^m - \ldots - (-1)^{n-1}\binom{n}{n-1}(1)^m$$

$$= \binom{n}{0}n^m - \binom{n}{1}(n-1)^m + \binom{n}{2}(n-2)^m - \binom{n}{3}(n-3)^m + \binom{n}{4}(n-4)^m - \ldots - (-1)^{n-1}\binom{n}{n-1}(1)^m$$

$$= \sum_{r=0}^{n-1}\binom{n}{r}(-1)^r(n-r)^m \qquad \ldots (2.12)$$

Total Number of Bijective Functions: By definition of bijective function different elements of A have different images and each element of B has at least one pre-image in A. This is possible only when m = n. Hence, in the case m ≠ n total number of bijective functions from A to B is zero. Like the case of one-one function when m = n total number of bijective functions will be

$$n(n-1)(n-2)\ldots(n-m+1) \qquad \ldots (2.13)$$

But here m = n, so

Total number of bijective functions = $n(n-1)(n-2)\ldots(n-n+1)$

$$= n(n-1)(n-2)\ldots(1) = n! \qquad \ldots (2.14)$$

Now, you can try the following Self-Assessment Question.

SAQ 1

(a) Show that the function $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x)=[x]$ is neither 1-1 nor onto.

(b) Show that the function $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x)=x^2$ is neither 1-1 nor onto.

(c) Show that the function $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x)=x^3$ is both 1-1 and onto.

(d) If n(A) = 5, n(B) = 7 then how many one-one functions are possible from A to B.

2.3 COUNTABLE SET

From school mathematics you know that a set is said to be a **finite set** if either it is an empty set or it has a finite number of elements.

For example, A = {2, 7, 9} is a finite set since it has three elements and 3 is a finite number. In fact, any real number x satisfying $-\infty < x < \infty$ is a finite number.

You also know that a set which is not finite is called an **infinite set**.

For example, B = {5, 10, 15, 20, …} is an infinite set since number of elements in B is not finite.

From earlier classes you know that two finite sets are said to be equivalent if number of elements in the two sets are equal. For example, if A = {a, b, c} and B = {7, 3, m} then $A \sim B$ because n(A) = 3 = n(B), where n(A) denotes cardinality (number of elements in the set) of the set A. But if you have two sets each having infinite number of elements then how you will check whether they are equivalent or not? To answer this question, we will take help of one-one correspondence which has been explained in the previous section of this unit. So, more general definition of equivalents sets in mathematics is given as follows.

Equivalent Sets: Let A and B be two sets either both finite or both infinite then we say that sets A and B are equivalent if either there exists a one-one correspondence from A to B or from B to A and is denoted by $A \sim B$. … (2.15)

For example,

(i) If A = {1, 2, 3, 4} and B = {3, 8, 13, 18}, then $A \sim B$ because there exists a one-one correspondence $f: A \to B$ defined by $f(x) = 5x - 2, \quad x \in A$ between A and B as shown in Fig. 2.5 (a)

(ii) If C = {3, 6, 9, 12, …} and D = {5, 10, 15, 20, …}, then $C \sim D$ because there exists a one-one correspondence $g: C \to D$ defined by $g(x) = \frac{5}{3}x, \quad x \in C$ between C and D as shown in Fig. 2.5 (b)

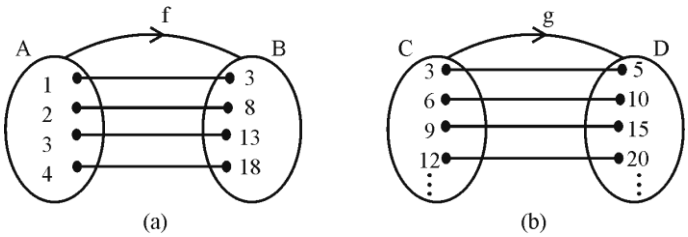


Fig. 2.5: Visualisation of equivalence of two (a) finite sets (b) infinite sets

In the previous unit we have discussed different families of numbers in mathematics. One of those families was the set of natural numbers. This family of natural numbers plays an important role to distinguish between countable and uncountable sets. The next definition specifies this importance.

Enumerable Set: A set E is said to be enumerable, if it is equivalent to the set of natural numbers, i.e., if $\mathbb{N} \sim \mathrm{E}$. … (2.16)
i.e., if there exists a one- one correspondence between $\mathbb{N}$ and E.
An enumerable set is also known as **denumerable** set or **countably infinite set**. … (2.17)

For example,

(i) If $\mathrm{E} = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots\right\}$ then E is enumerable because there exists one-one correspondence $\mathrm{f} : \mathbb{N} \to \mathrm{E}$ defined by

$$\mathrm{f}(\mathrm{n}) = \frac{1}{\mathrm{n}}, \quad \mathrm{n} \in \mathbb{N}$$

The same is shown in Fig. 2.6 (a).

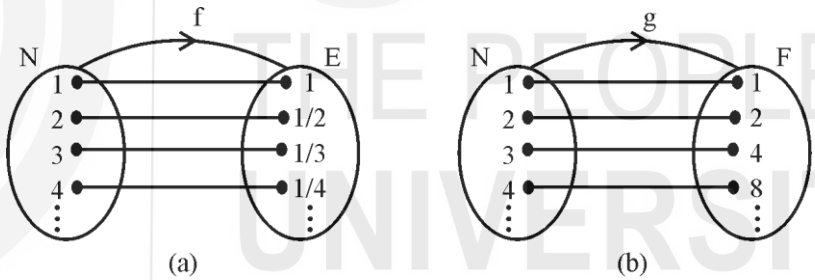


Fig. 2.6: One-one correspondence between (a) $\mathbb{N}$ and E (b) $\mathbb{N}$ and F

(ii) If $\mathrm{F} = \{1, 2, 4, 8, \cdots\}$ then F is enumerable because there exists one-one correspondence $\mathrm{f} : \mathbb{N} \to \mathrm{F}$ defined by

$$\mathrm{f}(\mathrm{n}) = 2^{\mathrm{n}-1}, \quad \mathrm{n} \in \mathbb{N}$$

The same is shown in Fig. 2.6 (b).

You have studied about one-one correspondence and enumerable set so now you can understand the definition of countable set which is given as follows.

Countable Set: A set is said to be countable if either it is finite or enumerable. … (2.18)

For example,

(i) If A = {} = ϕ, then number of elements in the set A is zero and therefore it is a finite set. So, set A is countable.

(ii) If B = {a, b, c}, then the number of elements in the set B is 3 and therefore it is a finite set. So, set B is countable.

(iii) If $C = \left\{1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \cdots\right\}$ then the number of elements in the set C is infinite but this set has one-one correspondence with the set of all natural numbers $\mathbb{N}$, $f : \mathbb{N} \to C$ defined by

$$f(n) = \frac{1}{n^2}, \quad n \in \mathbb{N}$$

Therefore, it is enumerable. Hence, it is a countable set.

Remark 4

- In Unit 3 of the Course MST-012, you will meet the word countable in the definition of the discrete random variable. If you want to enjoy the flavour of discrete world it is recommended that you should keep this definition of countable set in your mind when you go will through the definition of discrete random variable.

- Note that when a set is countable then its elements can be written in a sequence. You will study about sequence in Unit 4 of this course in detail. So, all those variables whose values can be written in a sequence are known as discrete random variable. Example 1 explain it further.

- Subset of a countable set is countable.

Example 1: Does the set containing all possible values of 'number of accidents' at a crossing during 4am to 10pm on a particular day is a countable set?

Solution: Answer of this question is yes and the explanation is given as follows.

Number of accidents during 4am to 10pm on a particular day may be 0 or 1 or 2 or 3 and so on. So, if we write these possible values in the set A then we have A = {0, 1, 2, 3, 4, …}. In real life obviously this set A is finite and so is countable. But if theoretically we assume that set A is infinite then also we can define a function $f : \mathbb{N} \to A$ by

$$f(x) = n - 1, \quad n \in \mathbb{N}$$

This function is both 1-1 and onto and hence $\mathbb{N} \sim A$. So, set A is enumerable and so countable.

Now, you can try the following Self Assessment Question.

SAQ 2

(a) Does the set of all integers is countable?

(b) Does the set of all prime numbers is countable?

Now, we define algebra and composition of functions in the next section.

2.4 COMBINING FUNCTIONS

You know how to add, subtract, multiply and divide in real numbers. You can also perform these operations in the universe of functions known as algebra of

functions which will be discussed here. Further, if function f takes you from set A to set B and function g takes you from set B to set C then you can compose functions g and f to go from set A to set C via set B this is known as composition of functions g and f which is also discussed here. Visualisation of composition of two functions f and g is shown in Figs. 2.7 and 2.8.

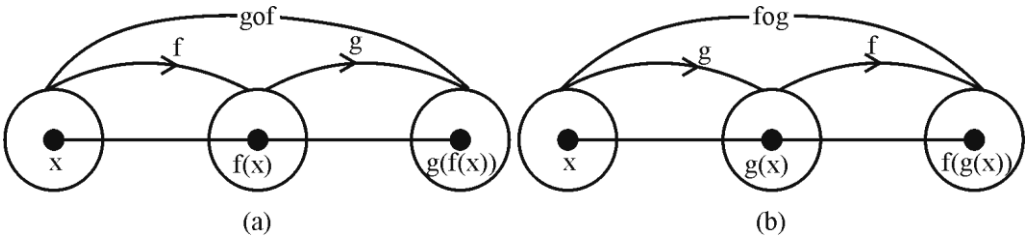


Fig. 2.7: Visualisation of composition of two functions

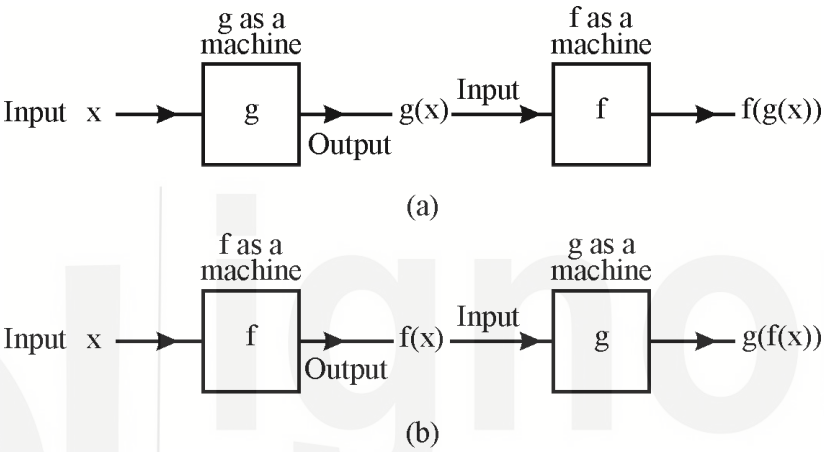


Fig. 2.8: Visualisation of composition of two functions

Let f and g be two functions having domains D_1 and D_2 respectively. Let $D = D_1 \cap D_2$ be their common domain. So, both functions f and g are defined on D hence we can apply both functions on all $x \in D$. This permits us to do algebra on functions. Sum, difference, product and quotient of functions f and g are defined as follows:

Sums of f and g

$$(f+g)(x) = f(x) + g(x), \quad x \in D \qquad \ldots (2.19)$$

Differences of f and g

$$(f-g)(x) = f(x) - g(x), \quad x \in D \qquad \ldots (2.20)$$

Products of f and g

$$(f\,g)(x) = f(x)\,g(x), \quad x \in D \qquad \ldots (2.21)$$

Quotients of f and g

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}, \quad x \in D, \text{ where } g(x) \neq 0 \text{ and} \qquad \ldots (2.22)$$

$$\left(\frac{g}{f}\right)(x) = \frac{g(x)}{f(x)}, \quad x \in D, \text{ where } f(x) \neq 0 \qquad \ldots (2.23)$$

Sometimes we also need to multiply a function by a constant. Let c be a real number then product of constant c with the function f is defined as follows:

$$(c\,f)(x) = c\,f(x), \quad x \in D_1 \qquad \ldots (2.24)$$

One more way of combining functions f and g is known as composition of f and g denoted by $g \circ f$. This makes sense if range of f is subset of domain of g and is defined as follows:

$$(g \circ f)(x) = g(f(x)), \quad x \in D_1, \text{ where } \text{Range}(f) \subseteq D_2 \qquad \ldots (2.25)$$

Similarly, $f \circ g$ makes sense if range of g is subset of domain of f and is defined as follows:

$$(f \circ g)(x) = f(g(x)), \quad x \in D_2, \text{ where } \text{Range}(g) \subseteq D_1 \qquad \ldots (2.26)$$

Let us explain these ideas of combining two functions in Examples 2 and 3 given as follows:

Example 2: Functions f and g are defined by $f(x) = \sqrt{2-x}$, $g(x) = \sqrt{x-1}$. Find domains of f and g. Obtain domains where $f+g$, $f-g$, fg and f/g are defined.

Also obtain (i) $(f+g)(2)$ (ii) $(f-g)(1)$ (iii) $(fg)(3/2)$ (iv) $\left(\frac{f}{g}\right)(3/2)$

(v) $(f+g)(3)$ (vi) $\left(\frac{g}{f}\right)(2)$ (vii) $\left(\frac{g}{f}\right)(1.8)$ (viii) $\left(\frac{f}{g}\right)(1.4)$ (ix) $\left(\frac{f}{g}\right)(1)$.

Solution: Domains of f and g are $(-\infty, 2]$ and $[1, \infty)$ respectively. Now, sums, differences and products are defined on $(-\infty, 2] \cap [1, \infty) = [1, 2]$, while f/g is defined on (1, 2] and g/f is defined on [1, 2).

(i) $(f+g)(2) = f(2) + g(2) = \sqrt{2-2} + \sqrt{2-1} = 0 + 1 = 1$

(ii) $(f-g)(1) = f(1) - g(1) = \sqrt{2-1} - \sqrt{1-1} = 1 - 0 = 1$

(iii) $(fg)(3/2) = f(3/2)g(3/2) = \sqrt{2-3/2}\sqrt{3/2-1} = \sqrt{1/2}\sqrt{1/2} = 1/2$

(iv) $\left(\frac{f}{g}\right)(3/2) = \frac{f(3/2)}{g(3/2)} = \frac{\sqrt{2-3/2}}{\sqrt{3/2-1}} = \frac{\sqrt{1/2}}{\sqrt{1/2}} = 1$

(v) Since domain of f + g is [1, 2] and $3 \notin [1, 2]$, so (f + g)(3) is not defined.

(vi) Since domain of g/f is [1, 2) and $2 \notin [1, 2)$ so (g/f)(2) is not defined.

(vii) $\left(\frac{g}{f}\right)(1.8) = \frac{g(1.8)}{f(1.8)} = \frac{\sqrt{1.8-1}}{\sqrt{2-1.8}} = \frac{\sqrt{0.8}}{\sqrt{0.2}} = \sqrt{\frac{0.8}{0.2}} = \sqrt{4} = 2$

(viii) $\left(\frac{f}{g}\right)(1.4) = \frac{f(1.4)}{g(1.4)} = \frac{\sqrt{2-1.4}}{\sqrt{1.4-1}} = \frac{\sqrt{0.6}}{\sqrt{0.4}} = \sqrt{\frac{0.6}{0.4}} = \sqrt{\frac{6}{4}} = \frac{\sqrt{6}}{2}$

(ix) Since domain of f/g is (1, 2] and $1 \notin (1, 2]$ so (f/g)(1) is not defined.

Example 3: Functions f and g are defined by $f(x) = \sqrt{x+3}$, $g(x) = x - 2$. Obtain (i) $(f \circ g)(x)$ (ii) $(g \circ f)(x)$ (iii) $(f \circ f)(x)$ (iv) $(g \circ g)(x)$ (v) $(f \circ g)(6)$ (vi) $(g \circ f)(6)$ (vii) $(f \circ f)(6)$ (viii) $(g \circ g)(6)$.

Solution: By definition of composition of two functions, we have

(i) $(f \circ g)(x) = f(g(x)) = f(x-2) = \sqrt{x-2+3} = \sqrt{x+1}$

(ii) $(g \circ f)(x) = g(f(x)) = g(\sqrt{x+3}) = \sqrt{x+3} - 2$

(iii) $(f \circ f)(x) = f(f(x)) = f(\sqrt{x+3}) = \sqrt{\sqrt{x+3}+3}$

(iv) $(g o g)(x) = g(g(x)) = g(x-2) = x-2-2 = x-4$

(v) Since $(f o g)(x) = \sqrt{x+1}$ so, $(f o g)(6) = \sqrt{6+1} = \sqrt{7}$

(vi) Since $(g o f)(x) = \sqrt{x+3} - 2$ so, $(g o f)(6) = \sqrt{6+3} - 2 = \sqrt{9} - 2 = 3 - 2 = 1$

(vii) Since $(f o f)(x) = \sqrt{\sqrt{x+3}+3}$ so,

$(f o f)(6) = \sqrt{\sqrt{6+3}+3} = \sqrt{\sqrt{9}+3} = \sqrt{3+3} = \sqrt{6}$

(viii) Since $(g o g)(x) = x - 4$ so, $(g o g)(6) = 6 - 4 = 2$

Now, you can try the following Self Assessment Question.

SAQ 3

Consider functions f and g as defined in Example 3. Obtain domains of the functions (i) f (ii) g (iii) $(f o g)(x)$ (iv) $(g o f)(x)$ (v) $(f o f)(x)$ (vi) $(g o g)(x)$.

2.5 CONTINUITY OF A FUNCTION

Limit and Continuity at a Point: In school mathematics you have studied what we mean by limit and continuity of a function at a point. You know that a function f is continuous at x = a if left hand limit (LHL), right hand limit (RHL) and value of the function f at x = a exist and, all the three are equal. i.e.,

$$\lim_{x \to a^-} f(x) = \lim_{x \to a^+} f(x) = f(a). \qquad \ldots (2.27)$$

You know the meaning of $x \to a^-$. For example, if $a = 2$, then $x \to 2^-$, means x is coming close and close to 2 but remains less than 2. For understanding purpose, you can suppose that x is taking values 1.9, 1.98, 1.998, 1.9998, 1.99998, and so on. In real mathematical sense this string of 9 is huge beyond your imagination. It can be realised by assuming that suppose someone go on writing 9 in 1.9999… whole life then his/her child do the same job and suppose it continue for 10 (nothing special in 10 just to specify some number for understanding the concept, you can consider another number) generations and tenth generation person put 8 at the end then this number is still less than 2 but very close to 2. Similarly, meaning of $x \to a^+$, can be understood by assuming a = 2 then $x \to 2^+$, means x is coming close and close to 2 but remains greater than 2. For understanding purpose, you can suppose that x is taking values 2.1, 2.01, 2.001, 2.0001, 2.00001, and so on. So, remember $x \to a$ means either x is slightly less than a or slightly greater than a but x ≠ a. You also know that at x = a if LHL = RHL ≠ f(a), then we say that $\lim_{x \to a} f(x)$ exists but function f is not continuous at x = a. Further, if both LHL and RHL exist but LHL ≠ RHL then we say that $\lim_{x \to a} f(x)$ does not exist.

Continuous Function: If a function is continuous at each point in its domain, then we say that it is a **continuous function**. … (2.28)

Geometrical Interpretation of the Continuity of a Function at a Point: Geometrically continuity of a function at a point x = a means when we draw the graph of the function with pen or pencil, we do not need to lift the pen or pencil as we cross the point (a, f(a)) on the graph of that function. … (2.29)

For example, (i) all polynomial functions are continuous refer Figs. 1.13 to 1.16 because there is no whole in their graphs, similarly, (ii) exponential function is continuous refer Fig. 1.17 (iii) logarithm function is continuous refer Fig. 1.19 (iv) sigmoid function is continuous refer Fig. 1.20. On the other hand (i) unit step function is not continuous because there is a whole in its graph at the origin refer Fig. 1.22, similarly, (ii) signum function is not continuous refer Fig. 1.23 (a) (iii) greatest integer function is not continuous refer Fig. 1.23 (b). In fact, greatest integer function is not continuous at all integers. So, it has infinite points where it is not continuous. But if a function is continuous at all points of its domain except at one point where it is not continuous then we say that this function is not continuous. So, to say that a function is not continuous it is enough to find out one point in its domain where it is not continuous.

Continuity on an Interval: We have defined that a function f is continuous at a point x = c if $\lim_{x\to c^-} f(x) = \lim_{x\to c^+} f(x) = f(c)$. Now there may be the situation that $\lim_{x\to c^-} f(x) = f(c) \neq \lim_{x\to c^+} f(x)$ in this case we say that the function f is **left continuous** at x = c. Another situation may be that $\lim_{x\to c^+} f(x) = f(c) \neq \lim_{x\to c^-} f(x)$ in this case we say that function f is **right continuous** at x = c. … (2.30)

If [a, b] lies in the domain of the function f then we say that function f is **continuous on the closed interval** [a, b] if (i) f is continuous at each point in the open interval (a, b) (ii) f is right continuous at x = a, and (iii) f is left continuous at x = b. … (2.31)

Let us now consider one example given as follows.

Example 4: Discuss the continuity of signum function at x = 0.

Solution: By definition of signum function (already discuss in the previous unit also refer Fig. 1.23(a)), we know that

$$f(x) = \text{sgn}(x) = \begin{cases} -1, & x<0 \\ 0, & x=0 \\ 1, & x>0 \end{cases} \qquad x \in \mathbb{R}$$

Now,

$$\text{LHL} = \lim_{x\to 0^-} f(x) = \lim_{x\to 0^-}(-1) \quad \left[\because x \to 0^- \Rightarrow x \text{ is slightly less than 0, so } f(x) = -1\right]$$
$$= -1$$

Similarly, $\text{RHL} = \lim_{x\to 0^+} f(x) = \lim_{x\to 0^+}(1) = 1$

Since, LHL ≠ RHL, so $\lim_{x\to 0} f(x)$ does not exist and hence f is not continuous at x = 0. $\left[\because \text{For continuity limit should exists as well as it must be equal to } f(0)\right]$

Epsilon (ε) and Delta (δ) Definition of Limit of a Function at a Point

Let $D \subseteq \mathbb{R}$ and $c \in \mathbb{R}$ be such that (c – r, c) and (c, c + r) are contained in D for some r > 0, and let $f : D \to \mathbb{R}$. Then $\lim_{x\to c} f(x)$ exists if and only if there is L such that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever

$$|x - c| < \delta, \text{ where } x \in D \Rightarrow |f(x) - L| < \varepsilon \qquad \ldots (2.32)$$

Epsilon (ε) and Delta (δ) Definition of Continuity of a Function at a Point

Let $D \subseteq \mathbb{R}$ then the function $f : D \to \mathbb{R}$ is said to be continuous at a point $c \in D$ if and only if for each $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever

$$|x - c| < \delta, \text{ where } x \in D \Rightarrow |f(x) - f(c)| < \varepsilon \qquad \ldots (2.33)$$

In this definition of continuity $\varepsilon > 0$ depends on both point c as well as $\delta > 0$.

… (2.34)

If in this definition of continuity $\varepsilon > 0$ depends only on $\delta > 0$ not on the point c then we say that function is **uniformly continuous**. … (2.35)

Example 5: Using ε and δ definition show that $\lim_{x \to 5}(2x - 4) = 6$.

Solution: In usual notations we are given f(x) = 2x – 4, c = 5, L = 6.

Now, let $\varepsilon > 0$ be given then

$$|f(x) - L| < \varepsilon \Rightarrow |2x - 4 - 6| < \varepsilon \Rightarrow |2x - 10| < \varepsilon \Rightarrow 2|x - 5| < \varepsilon \Rightarrow |x - 5| < \frac{\varepsilon}{2}$$

Let $\delta = \frac{\varepsilon}{2}$, then we have for given $\varepsilon > 0$, there exists $\delta = \frac{\varepsilon}{2}$, such that

whenever $|x - 5| < \delta = \frac{\varepsilon}{2}$, $\Rightarrow |(2x - 4) - 6| < \varepsilon$, so $\lim_{x \to 5}(2x - 4) = 6$.

Now, you can try the following Self-Assessment Question.

SAQ 4

For given $\varepsilon = 0.1$ find $\delta > 0$ such that whenever $|x - c| < \delta$, then $|f(x) - L| < \varepsilon$, where $f(x) = \sqrt{5x - 1}$, $L = 3$, $c = 2$.

2.6 DIFFERENTIABILITY OF A FUNCTION AT A POINT

In earlier classes you have studied a lot about differentiability and its applications. Let y = f(x) be a function then average rate of change of y with respect to x over the interval [a, a + h] is defined by

$$\frac{f(a+h) - f(a)}{a + h - a} = \frac{f(a+h) - f(a)}{h}$$

If $\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$ exists, then it is known as instantaneous rate of change of y with respect to x at the point a, and is known as derivative of y with respect to x at the point a and is denoted by $f'(a)$. … (2.36)

Example 6: If $f(x) = 2x^2 - 4x + 5$ then find $f'(3)$ using definition.

Solution: By definition of derivative, we know that

$$f'(3) = \lim_{h\to 0}\frac{f(3+h)-f(3)}{h} = \lim_{h\to 0}\frac{2(3+h)^2 - 4(3+h)+5-\left[2(3)^2-4(3)+5\right]}{h}$$

$$= \lim_{h\to 0}\frac{2(9+h^2+6h)-(12+4h)+5-\left[18-12+5\right]}{h} = \lim_{h\to 0}\frac{2h^2+8h}{h}$$

$$= \lim_{h\to 0}(2h+8) = 8$$

Now, you can try the following Self-Assessment Question.

SAQ 5

If $f(x) = |x-5|$ then discuss its differentiability at x = 5.

2.7 SUMMARY

A brief summary of what we have covered in this unit is given as follows:

- A function $f : X \to Y$ is said to be **1-1** or **injective** function if distinct elements of X are associated to distinct elements of Y under f. If a function is 1-1 then **geometrically** it will satisfy the following condition.

  Each horizontal line either does not intersect the graph of the function at all or if it intersects it will intersect exactly at one point. In other words, there should be no horizontal line which intersects the graph of the function at two or more than two points.

- A function $f : X \to Y$ is said to be onto or **surjective** if each element of Y has at least one pre image in X. That is for each $y \in Y$, there exists at least one $x \in X$ such that $f(x) = y$. If a function is onto then **geometrically** it will satisfy the following condition.

  Horizontal line corresponding to each point of the codomain must intersect the graph of the function at least at one point.

- A function $f : X \to Y$ is said to be one-one and onto or **bijective** or **one-one correspondence** if f is both one-one as well as onto. That is if a function is both one-one and onto then it is known as one-one correspondence or bijective function. Therefore, by combining geometrical conditions of both one-one and onto functions we can say that a function will be bijective if each horizontal line corresponding to each point of the codomain must intersect the graph of the function exactly at one point.

- A set is said to be a **finite set** if either it is an empty set or it has a finite number of elements. A set which is not finite is called an **infinite set**.

- Let A and B be two sets either both finite or both infinite then we say that sets A and B are **equivalent** if either there exists a one-one correspondence from A to B or from B to A and is denoted by $A \sim B$.

- A set E is said to be **enumerable**, if it is equivalent to the set of natural numbers, i.e., if $\mathbb{N} \sim E$.

- A set is said to be countable if either it is finite or enumerable.

- Sum, difference, product and quotient of functions f and g having D as their common domain are defined as follows:

  $(f+g)(x) = f(x) + g(x), \quad x \in D$

  $(f-g)(x) = f(x) - g(x), \quad x \in D$

  $(f\,g)(x) = f(x)\,g(x), \quad x \in D$

  $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}, \quad x \in D,$ where $g(x) \neq 0$ and

  $\left(\frac{g}{f}\right)(x) = \frac{g(x)}{f(x)}, \quad x \in D,$ where $f(x) \neq 0$

- Sometimes we also need to multiply a function by a constant. Let c be a real number then product of constant c with the function f is defined as follows:

  $(c\,f)(x) = c\,f(x), \quad x \in D_1$

- Composition of functions f and g is denoted by $g\,o\,f$. This makes sense if range of f is subset of domain of g and is defined as follows:

  $(g\,o\,f)(x) = g(f(x)), \quad x \in D_1,$ where $\text{Range}(f) \subseteq D_2$

  Similarly, $f\,o\,g$ makes sense if range of g is subset of domain of f and is defined as follows:

  $(f\,o\,g)(x) = f(g(x)), \quad x \in D_2,$ where $\text{Range}(g) \subseteq D_1$

- If at x = a, LHL = RHL ≠ f(a) then we say that $\lim_{x\to a} f(x)$ exists but function f is not continuous at x = a.

- A function f is **continuous** at x = a if left hand limit (LHL), right hand limit (RHL) and value of the function f at x = a exist and all the three are equal, i.e., $\lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x) = f(a)$.

- Differentiability of a function y = f(x) with respect to x at the point x = a is denoted by $f'(a)$ and is defined by

  $$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}$$

2.8 TERMINAL QUESTIONS

1 (a) Show that the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = |x|$ is neither 1-1 nor onto.

(b) Show that the function $f : \mathbb{R}\backslash\{5\} \to \mathbb{R}\backslash\{2\}$ defined by $f(x) = \frac{2x-3}{x-5}$ is both 1-1 and onto.

(c) If n(A) = 5, n(B) = 7 then how many functions are possible from A to B.

(d) If n(A) = 5, n(B) = 3 then how many onto functions are possible from A to B.

(e) If n(A) = 5, n(B) = 5 then how many bijective functions are possible from A to B.

2 If a coin is tossed till head appears then write the set of all possible outcomes. Is this set countable?

3. Discuss the continuity of the greatest integer function at x = n, where n is an integer. Also, comment on left and right continuity of the function at n.

4. Discuss the continuity of the greatest integer function at x = m, where m is not an integer.

2.9 SOLUTIONS/ANSWERS

Self-Assessment Questions (SAQs)

1 (a) We have to show that the function $f:\mathbb{R}\to\mathbb{R}$ defined by

$f(x)=[x],\ \ x\in\mathbb{R}$

is neither 1-1 nor onto. By definition of greatest integer function, we have

$f(2.3)=[2.3]=2$ and also $f(2.4)=[2.4]=2$

So, $f(2.3)=f(2.4)$ but $2.3\neq 2.4$ $\Rightarrow$ it is not a one-one function.

Also, $2.7\in\mathbb{R}$ (codomain) but there is no $x\in\mathbb{R}$ (domain) such that

$f(x)=2.7$ $\left[\begin{array}{l}\because \text{ Output of greatest integer function is always}\\ \text{an integer and hence cannot be } 2.7\end{array}\right]$

$\Rightarrow$ it is not an onto function.

Hence, the given function f is neither 1-1 nor onto.

(b) We have to show that the function $f:\mathbb{R}\to\mathbb{R}$ defined by

$f(x)=x^2,\ \ x\in\mathbb{R}$

is neither 1-1 nor onto.

Consider $-2, 2\in\mathbb{R}$ (domain)

Now, $f(-2)=(-2)^2=4$ and $f(2)=(2)^2=4$. So, $f(-2)=f(2)$ but $-2\neq 2$

$\Rightarrow$ it is not a one-one function.

Also, $-2\in\mathbb{R}$ (codomain) but there is no $x\in\mathbb{R}$ (domain) such that

$f(x)=x^2=-2$ $[\because$ Square of a real number can never be equal to $-2]$

$\Rightarrow$ it is not an onto function.

Hence, the given function f is neither 1-1 nor onto.

(c) We have to show that the function $f:\mathbb{R}\to\mathbb{R}$ defined by

$$f(x) = x^3, \quad x \in \mathbb{R}$$

is both 1-1 and onto.

Consider $x_1, x_2 \in \mathbb{R}$ (domain) be such that

$$f(x_1) = f(x_2) \Rightarrow x_1^3 = x_2^3 \Rightarrow x_1 = x_2$$

$\Rightarrow$ it is a one-one function.

Also, for each $y \in \mathbb{R}$ (codomain) there exists $\sqrt[3]{y} \in \mathbb{R}$ (domain) such that

$$f(\sqrt[3]{y}) = \left(\sqrt[3]{y}\right)^3 = y$$

$\Rightarrow$ it is an onto function.

Hence, the given function f is both 1-1 and onto.

(d) Here n(A) = 5 = m and n(B) = 7 = n, so total number of 1-1 functions from A to B are given by

$$n(n-1)(n-2)\ldots(n-m+1) = 7(7-1)(7-2)(7-3)(7-4) = 7.6.5.4.3$$
$$= 2520$$

2 (a) We have to show that the set $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ is countable. Define a function $f : \mathbb{N} \to \mathbb{Z}$ by

$$f(x) = \begin{cases} \dfrac{n-1}{2}, & \text{if } n \text{ is odd} \\ -\dfrac{n}{2}, & \text{if } n \text{ is even} \end{cases} \qquad x \in \mathbb{N}$$

We claim that f is 1-1

Let $n_1, n_2 \in \mathbb{N}$ be such that

$$f(n_1) = f(n_2) \qquad \ldots (2.1)$$

Case I: Both n_1, n_2 are odd, then (2.1) gives

$$\frac{n_1 - 1}{2} = \frac{n_2 - 1}{2} \Rightarrow n_1 - 1 = n_2 - 1 \Rightarrow n_1 = n_2$$

$\Rightarrow$ In this case f is 1-1.

Case II: Both n_1, n_2 are even, then (2.1) gives

$$-\frac{n_1}{2} = -\frac{n_2}{2} \Rightarrow -n_1 = -n_2 \Rightarrow n_1 = n_2$$

$\Rightarrow$ In this case f is 1-1.

Case III: One of n_1, n_2 is even and other is odd. Without loss of generality suppose n_1 is odd and n_2 is even, then (2.1) gives

$\dfrac{n_1 - 1}{2} = -\dfrac{n_2}{2} \Rightarrow n_1 - 1 = -n_2 \Rightarrow n_1 + n_2 = 1$ a contradiction because both n_1 and n_2 are ≥ 1 and hence $n_1 + n_2 \geq 2$.

This implies if $n_1 \neq n_2$ then $f(n_1) \neq f(n_2)$

Hence, in all cases, we have f is 1-1.

Next, we claim that f is onto.

Let n be an integer. Two cases arise:

Case I: n is a negative integer $\Rightarrow$ – 2n is an even positive integer and so $-2n \in \mathbb{N}$ is such that

$$f(-2n) = -\frac{-2n}{2} = n$$

So, in this case f is onto.

Case II: n (≥ 0) is a non-negative integer $\Rightarrow$ 2n + 1 is an odd positive integer.

$\Rightarrow 2n+1 \in \mathbb{N}$ is an odd natural number such that

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

So, in this case f is onto.

Hence, f is onto.

We have proved that f is 1-1 and onto hence $\mathbb{N} \sim \mathbb{Z} \Rightarrow \mathbb{Z}$, the set of all integers is a countable set.

(b) We know that an integer > 1 is called a prime number if it has only two positive divisors 1 and the number itself. Therefore, if we denote the set

of all prime numbers by P then

P = {2, 3, 5, 7, 11, 13, 17, 19, 23, …}

Note that this set P is a subset of $\mathbb{N}$, the set of all natural numbers. We know that subset of a countable set is countable. Hence, P is countable because set of all natural numbers is countable and $P \subseteq \mathbb{N}$.

3. We are given

$$f(x) = \sqrt{x+3},\ \ g(x) = x-2,\ \ (f\,o\,g)(x) = \sqrt{x+1},\ \ (g\,o\,f)(x) = \sqrt{x+3} - 2$$

$$(f\,o\,f)(x)) = \sqrt{\sqrt{x+3}+3},\ \ (g\,o\,g)(x) = x-4$$

(i) Function f will take real values if $x+3 \geq 0 \Rightarrow x \geq -3$. So, domain of f is $[-3, \infty)$.

(ii) Function g takes real values for all real values of x. Hence, domain of g is $\mathbb{R}$.

(iii) Function $(f\,o\,g)(x) = \sqrt{x+1}$ will take real values if $x+1 \geq 0 \Rightarrow x \geq -1$. Hence, domain of $f\,o\,g$ is $[-1, \infty)$.

(iv) Function $(g\,o\,f)(x) = \sqrt{x+3} - 2$ will take real values if $x+3 \geq 0 \Rightarrow x \geq -3$. Hence, domain of $g\,o\,f$ is $[-3, \infty)$.

(v) Function $(f\,o\,f)(x) = \sqrt{\sqrt{x+3}+3}$ will take real values if $x+3 \geq 0 \Rightarrow x \geq -3$. Hence, domain of $f\,o\,f$ is $[-3, \infty)$.

(vi) Function $(g o g)(x) = x - 4$ takes real values for all real values of x. Hence, domain of $g o g$ is $\mathbb{R}$.

4. We are given $f(x) = \sqrt{5x-1}$, $L = 3$, $\varepsilon = 0.1$, $c = 2$

Now, $|f(x) - L| < \varepsilon \Rightarrow \left|\sqrt{5x-1} - 3\right| < 0.1$

$\Rightarrow -0.1 < \sqrt{5x-1} - 3 < 0.1 \quad \left[\because |x| < a \Rightarrow -a < x < a, \text{ refer } (6.95)\right]$

$\Rightarrow 2.9 < \sqrt{5x-1} < 3.1 \Rightarrow (2.9)^2 < 5x - 1 < (3.1)^2$

$\Rightarrow 8.41 + 1 < 5x < 9.61 + 1 \Rightarrow 9.41 < 5x < 10.61 \Rightarrow 1.882 < x < 2.122$

Now, $\delta = \min\{2 - 1.882,\ 2.122 - 2\} = \min\{0.118,\ 0.122\} = 0.118$

Hence, $\delta = 0.118$. In fact, all values between 0 and 0.118 excluding 0 will work as values of δ for given $\varepsilon = 0.1$.

5. We are given $f(x) = |x - 5|$

We have already studied about absolute value function or modulus function in Unit 1 of this course refer (1.47). Here, first we have to find out left hand derivative and right hand derivative.

$$\text{LHD}_{\text{at } x=5} = \lim_{h \to 0^-} \frac{f(5+h) - f(5)}{h} = \lim_{h \to 0^-} \frac{|5+h-5| - |5-5|}{h} = \lim_{h \to 0^-} \frac{|h| - |0|}{h}$$

$$= \lim_{h \to 0^-} \frac{-h - 0}{h} \quad \left[\begin{array}{l} \because h \to 0^- \text{ means h is slightly less than 0,} \\ \text{so } |h| = -h. \text{ Also, } |0| = 0, \text{ refer (1.47)} \end{array}\right]$$

$$= \lim_{h \to 0^-} \frac{-h}{h} = \lim_{h \to 0^-} (-1) = -1$$

$$\text{RHD}_{\text{at } x=5} = \lim_{h \to 0^+} \frac{f(5+h) - f(5)}{h} = \lim_{h \to 0^+} \frac{|5+h-5| - |5-5|}{h} = \lim_{h \to 0^+} \frac{|h| - |0|}{h}$$

$$= \lim_{h \to 0^+} \frac{h - 0}{h} \quad \left[\begin{array}{l} \because h \to 0^+ \text{ means h is slightly greater than 0,} \\ \text{so } |h| = h. \text{ Also, } |0| = 0, \text{ refer (1.47)} \end{array}\right]$$

$$= \lim_{h \to 0^+} \frac{h}{h} = \lim_{h \to 0^+} (1) = 1$$

Since, $\text{LHD}_{\text{at } x=5} = -1 \neq \text{RHD}_{\text{at } x=5} = 1$. Hence, f is not differentiable at x = 5.

Terminal Questions

1 (a) We have to show that the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$f(x) = |x|, \quad x \in \mathbb{R}$

is neither 1-1 nor onto. By definition of modulus function, we know that

$$f(x) = |x| = \begin{cases} x, & x \geq 0 \\ -x, & x < 0 \end{cases}, \quad x \in R$$

$\therefore\ f(-2) = |-2| = -(-2) = 2$ and $f(2) = |2| = 2$

So, $f(-2) = f(2)$ but $-2 \neq 2 \Rightarrow$ it is not a one-one function.

Also, $-2 \in \mathbb{R}$ (codomain) but there is no $x \in \mathbb{R}$ (domain) such that

$$f(x) = -2 \quad \begin{bmatrix} \because \text{ Output of modulus function can} \\ \text{never be a negative number} \end{bmatrix}$$

$\Rightarrow$ it is not an onto function.

Hence, the given function f is neither 1-1 nor onto.

(b) We have to show that the function $f : \mathbb{R}\backslash\{5\} \to \mathbb{R}\backslash\{2\}$ defined by

$$f(x) = \frac{2x-3}{x-5}, \quad x \in \mathbb{R}$$

is both 1-1 and onto.

Consider $x_1, x_2 \in \mathbb{R}$ (domain) be such that

$$f(x_1) = f(x_2) \Rightarrow \frac{2x_1 - 3}{x_1 - 5} = \frac{2x_2 - 3}{x_2 - 5}$$

$$\Rightarrow 2x_1x_2 - 10x_1 - 3x_2 + 15 = 2x_1x_2 - 3x_1 - 10x_2 + 15$$

$$\Rightarrow -10x_1 - 3x_2 = -3x_1 - 10x_2 \quad \Rightarrow -7x_1 = -7x_2 \quad \Rightarrow x_1 = x_2$$

$\Rightarrow$ it is a one-one function.

Also, for each $y \in \mathbb{R}$ (codomain) there exists $\frac{5y-3}{y-2} \in \mathbb{R}$ (domain) s.t.

$$f\left(\frac{5y-3}{y-2}\right) = \frac{2\left(\frac{5y-3}{y-2}\right) - 3}{\frac{5y-3}{y-2} - 5} = \frac{10y - 6 - 3y + 6}{5y - 3 - 5y + 10} = \frac{7y}{7} = y$$

$\Rightarrow$ it is an onto function.

Hence, the given function f is both 1-1 and onto.

(c) Here n(A) = 5 = m and n(B) = 7 = n, so total number of functions from A to B are given by $n^m = 7^5 = 16807$

(d) Here n(A) = 5 = m and n(B) = 3 = n, so total number of onto functions from A to B are given by

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

$$= 243 - 96 + 3 = 150$$

(e) Here n(A) = 5 = m and n(B) = 5 = n, so total number of bijective functions from A to B are given by n! = 5! = 120.

2. If the outcomes head and tail are represented by H, T respectively and S denotes the set of all possible outcomes of this experiment then

S = {H, TH, TTH, TTTH, TTTTH, …}

Let X counts number of failures preceding the first success then X takes values 0, 1, 2, 3, 4, … corresponding to the elements of the set S and if

we write values of X in a set then

X = {0, 1, 2, 3, 4, …}

Note that corresponding to each element of S there is exactly one element in X and vice versa so n(S) = n(X). So, in order to prove that S is countable it is enough to prove that X is countable. Define a function $f:\mathbb{N}\to X$ by

$$f(n)=n-1,\quad n\in\mathbb{N}$$

We claim that f is 1-1 and onto. Let us first prove that f is 1-1. Let $n_1, n_2\in\mathbb{N}$ be such that

$$f(n_1)=f(n_2)\ \Rightarrow n_1-1=n_2-1\Rightarrow n_1=n_2$$

$\Rightarrow$ f is a one-one function.

Also, for each $n\in X$ (codomain) there exists $n+1\in\mathbb{N}$ (domain) such that

$$f(n+1)=n+1-1=n$$

$\Rightarrow$ f is an onto function.

Hence, the given function f is both 1-1 and onto. Thus, $\mathbb{N}\sim X\ \Rightarrow$ X is enumerable and so X is countable. Countability of X implies S is countable because n(X) = n(S) $\Rightarrow X\sim S$.

3. We claim that the greatest integer function at x = n, where n is an integer is not continuous. Let us calculate LHL and RHL at x = n.

$$\text{LHL}=\lim_{x\to n^-}f(x)=\lim_{x\to n^-}[x]=n-1\quad\begin{bmatrix}\because x\to n^- \text{ means}\\ n-1<x<n,\text{ so }[x]=n-1\end{bmatrix}$$

$$\text{Similarly, RHL}=\lim_{x\to n^+}f(x)=\lim_{x\to n^+}[x]=n\quad\begin{bmatrix}\because x\to n^+ \text{ means}\\ n<x<n+1,\text{ so }[x]=n\end{bmatrix}$$

$\therefore$ LHL ≠ RHL $\Rightarrow$ f is not continuous at x = n, where n is an integer.

Left and Right continuity of f at n: Obviously, f(n) = [n] = n.

Since $\text{LHL}_{\text{at } x=n}=n-1\neq n=f(n)\Rightarrow$ f is not left continuous at x = n.

However, $\text{RHL}_{\text{at } x=n}=f(n)=n\Rightarrow$ f is right continuous at x = n.

4. We claim that the greatest integer function is continuous at x = m, where m is not an integer. Let us calculate LHL and R HL at x = m. Suppose k is an integer such that k < m < k + 1.

$$\text{LHL}=\lim_{x\to m^-}f(x)=\lim_{x\to m^-}[x]=k\quad\begin{bmatrix}\because x\to m^- \text{ and } k<m<k+1 \text{ means}\\ k<x<k+1,\text{ so }[x]=k\end{bmatrix}$$

$$\text{Similarly, RHL}=\lim_{x\to m^+}f(x)=\lim_{x\to m^+}[x]=k\quad\begin{bmatrix}\because x\to m^+ \text{ and } k<m<k+1\\ \Rightarrow k<x<k+1,\text{ so }[x]=k\end{bmatrix}$$

Also, f(m) = [m] = k $\left[\because k<m<k+1,\text{ where k is an integer}\Rightarrow[m]=k\right]$

$\therefore$ LHL = RHL = f(m) $\Rightarrow$ f is continuous at x = m, where m is not an integer.

UNIT 3

SET FUNCTION AND DISTANCE FUNCTION

Structure

3.1 Introduction

Expected Learning Outcomes

3.2 Extended Real Numbers

3.3 Binary Operations

3.4 Set Function and Measure

3.5 Characteristic, Simple and Step Functions

3.6 Distance Function and Metric

3.7 Summary

3.8 Terminal Questions

3.9 Solutions/Answers

3.1 INTRODUCTION

In Units 1 and 2 you have studied the functions of a single variable. You know that function is a rule which associates each element of a set with a unique element of another set. In this unit, you will study set function in Sec. 3.4 which is also a function from the set X to Y but here members of the set X are set itself. So, the domain of the set function is a collection of subsets of a set and the range may be any subset of the set of all extended real numbers. What we mean by extended real numbers and how we can perform operations in an extended real number system is explained in Sec. 3.2. You know that sum of two natural numbers is again a natural number. In such a case we say that the operation of addition is a binary operation on the set of all natural numbers. In Sec. 3.3 we have introduced the idea of binary operations on the family of numbers. Three special functions namely: characteristic, simple and step functions are defined in Sec. 3.5. In the course MST-026 you have to find out the distance between two points, so some distance functions are discussed in Sec. 3.6.

What we have discussed in this unit is summarised in Sec. 3.7. Self-Assessment Questions (SAQs) have been given in some sections which are generally based on the content discussed in that section. But to give you a good practice of what we have discussed in this unit, some more questions based on the entire unit are given in Sec. 3.8 under the heading Terminal Questions. Due to the reason mentioned in Sec. 1.1 of Unit 1, solutions of all

the SAQs and Terminal Questions are given in Sec. 3.9.

In the next unit, you will study the convergence and divergence of sequences and series.

Expected Learning Outcomes

After completing this unit, you should be able to:

- ❖ explain the rules for addition and multiplication in the world of extended real numbers;
- ❖ define the set function and give some examples of the set function; and
- ❖ define some functions: characteristic function, simple function, step function and metric function.

3.2 EXTENDED REAL NUMBERS

You know when we say that x is a real number it means $-\infty < x < \infty$. In other words, $-\infty$ and $+\infty$ are not included in the set of real numbers. If we include these two symbols in the set of all real numbers then this extended set is called the **extended set of real numbers** and is denoted by $\mathbb{R}^*$. So, $\mathbb{R}^* = (-\infty, +\infty) \cup \{-\infty, +\infty\} = [-\infty, +\infty]$. After including these two symbols we have to perform addition and multiplication with these two new symbols. So, we need to define what are the rules for addition and multiplication with these two new symbols. These rules are explained as follows:

Addition: If x is a real number, i.e., $-\infty < x < +\infty$, then

- $x + (-\infty) = -\infty$ … (3.1)
- $x + (+\infty) = +\infty$ … (3.2)
- $(+\infty) + (+\infty) = +\infty$ ... (3.3)
- $(-\infty) + (-\infty) = -\infty$ ... (3.4)

Multiplication: If x is a real number, i.e., $-\infty < x < +\infty$, then

- If x > 0 then $x(-\infty) = (-\infty)x = -\infty$, $x(+\infty) = (+\infty)x = +\infty$ … (3.5)
- If x < 0 then $x(-\infty) = (-\infty)x = +\infty$, $x(+\infty) = (+\infty)x = -\infty$ … (3.6)
- If x = 0 then $x(-\infty) = (-\infty)x = 0$, $x(+\infty) = (+\infty)x = 0$ … (3.7)
- $(+\infty)(-\infty) = (-\infty)(+\infty) = -\infty$, $(-\infty)(-\infty) = +\infty$, $(+\infty)(+\infty) = +\infty$ … (3.8)

But remember, following are not defined even in the world of extended real numbers.

$(+\infty) + (-\infty)$ and $(-\infty) + (+\infty)$ … (3.9)

Now, you can try the following Self-Assessment Question.

SAQ 1

Suppose you are working in the world of extended real numbers then replace the ? sign by appropriate real numbers or the symbol $-\infty$ or $+\infty$ as per the rules of this world.

(i) $(-3)(-\infty) = ?$ (ii) $(5)(-\infty) = ?$ (iii) $(-10)(+\infty) = ?$

(iv) $-5 + (-\infty) = ?$ (v) $15 + (+\infty) = ?$ (vi) $-7 + (-\infty) = ?$

(vii) $-9+(+\infty)=?$ (viii) $(0)(-\infty)=?$ (ix) $(0)(+\infty)=?$
(x) $(+\infty)+(-\infty)=?$

3.3 BINARY OPERATIONS

Consider two numbers 2 and 5 then you know how to apply four fundamental operations addition, subtraction, multiplication and division on two numbers. So, we have $2+5=7,\ 2-5=-3,\ 2\times 5=10,\ 2\div 5=0.4.$ In mathematics we say that an operation $*$ (say) on a non-empty set X is said to be binary if $a*b\in X,$ for all $a,\ b\in X.$ If $*$ is a binary operation on a non-empty set X then we also say that X is **closed with respect to the operation** $*.$ If so then we also say that **closure property** holds in X with respect to the operation $*.$ For example, set $\mathbb{N}$ of all natural numbers is closed with respect to the operations addition (+) and multiplication $(\times)$ because sum and product of two natural numbers is always a natural number, i.e.,

$$a+b,\ a\times b\in\mathbb{N},\ \text{for all } a, b\in\mathbb{N}. \qquad \ldots (3.10)$$

But the set of all natural numbers is not closed with respect to the subtraction $(-)$ and division $(\div)$ operations because $2, 5\in\mathbb{N}$ but $2-5=-3\notin\mathbb{N}$ and $2\div 5=0.4\notin\mathbb{N}.$

However, addition, subtraction, multiplication and division (except zero) all are binary operations in the set $\mathbb{R}$ of all real numbers.

You have gotten the idea of binary operation. Before defining it in general let us first define cartesian product of two sets.

Cartesian Product: Cartesian product of two sets A and B is denoted by $A\times B$ and define as follows. $A\times B=\{(x, y): x\in A,\ y\in B\}$ … (3.11)

For example, if A = {2, 5} and B = {4, 6, 7} then

$A\times B=\{(2, 4), (2, 6), (2, 7), (5, 4), (5, 6), (5, 7)\}$

Now we can define binary operation in general as follows.

A binary operation $*$ is a function $*: X\times X\to X$ defined by

$$*(a, b)=a*b,\ (a, b)\in X\times X \qquad \ldots (3.12)$$

There are some other terms related to binary operation which are explained as follows.

Commutative Binary Operation: A binary operation $*: X\times X\to X$ is said to be commutative binary operation if it satisfies the following condition

$$a*b=b*a,\ \text{for all } a, b\in X \qquad \ldots (3.13)$$

For example, binary operations addition (+) and multiplication $(\times)$ on the set $\mathbb{N}$ of all natural numbers are commutative since

$$a+b=b+a \text{ and } a\times b=b\times a \text{ for all } a, b\in\mathbb{N} \qquad \ldots (3.14)$$

Operation subtraction $(-)$ is a binary operation on the set of all real numbers $\mathbb{R}.$ However, it is not commutative since there exists $2, 5\in\mathbb{R}$ such that $2-5=-3\in\mathbb{R}$ and $5-2=3\in\mathbb{R},$ but $2-5\neq 5-2,$ hence, subtraction $(-)$ is not commutative on the set of all real numbers $\mathbb{R}.$

Associative: A binary operation $*: X \times X \to X$ is said to be associative if it satisfies the following condition

$$(a * b) * c = a * (b * c), \text{ for all } a, b, c \in X \qquad \ldots (3.15)$$

For example, binary operations addition (+) and multiplication $(\times)$ on the set $\mathbb{N}$ of all natural numbers are associative since

$$(a + b) + c = a + (b + c) \text{ and } (a \times b) \times c = a \times (b \times c) \text{ for all } a, b, c \in \mathbb{N} \quad (3.16)$$

But operation subtraction $(-)$ is not associative on $\mathbb{R}$, the set of all real numbers because there exists $2, 5, 7 \in \mathbb{R}$ such that $(2-5)-7 \neq 2-(5-7)$. Hence, subtraction $(-)$ is not associative on the set of all real numbers $\mathbb{R}$.

Similarly, operation subtraction $(-)$ is not associative on any of the sets $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$.

Existence of Identity: If $*: X \times X \to X$ is a binary operation then we say that an element $e \in X$ if it exists is called identity element with respect to the operation $*$ if

$$e * a = a = a * e, \ \forall\ a \in X \qquad \ldots (3.17)$$

For example, 0 is the identity element with respect to the binary operation addition (+) in the sets $\mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$. Also, 1 is the identity element with respect to the binary operation multiplication $(\times)$ in the sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$.

Subtraction $(-)$ is binary operations in $\mathbb{R}$ but there is no identity element in $\mathbb{R}$ with respect to subtraction operation.

Existence of Inverse: If $*: X \times X \to X$ is a binary operation and e be the identity element in X with respect to the operation $*$ then we say that an element $a \in X$ has its inverse in X if there exists an element $b \in X$ such that

$$b * a = e = a * b \qquad \ldots (3.18)$$

b is known as inverse of a and is generally denoted by $b = a^{-1}$. If each non-zero element of X has its inverse in X then we say that property 'existence of inverse' with respect to $*$ holds in X.

For example, property 'existence of inverse' with respect to addition (+) holds in $\mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$. But property 'existence of inverse' with respect to multiplication $(\times)$ holds in $\mathbb{Q}$ and $\mathbb{R}$ but not in $\mathbb{Z}$.

In the next section we will discuss set function and measure. So, before studying them it will be better if you have understanding of what we mean by a σ-field which is defined as follows.

Let Ω be a non-empty set and $\mathcal{F}$ be a collection of subsets of Ω then we say that the class $\mathcal{F}$ forms a **sigma algebra** (σ-algebra) or **sigma field** (**σ-field**) if it satisfies following three conditions:

(i) $\phi, \ \Omega \in \mathcal{F}$, i.e., empty set and the full space Ω should be members of $\mathcal{F}$

(ii) $A \in \mathcal{F} \Rightarrow A^c (= \Omega \setminus A) \in \mathcal{F}$, i.e., $\mathcal{F}$ is closed with respect to complement

(iii) $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ for any countable collection $\{A_n\}_{n=1}^{\infty}$ in $\mathcal{F}$, i.e., $\mathcal{F}$ is closed with respect to countable union $\qquad \ldots (3.19)$

Let us consider two examples.

Example 1: If Ω = {a, b, c, d} then what is the

(i) smallest σ-field on Ω

(ii) smallest σ-field containing the set A = {a} on Ω

(iii) largest σ-field on Ω.

Solution: We know that a collection $\mathcal{F}$ of subsets of Ω is called a σ-field if it satisfies the following three conditions.

(a) $\phi,\ \Omega \in \mathcal{F}$ (b) $A \in \mathcal{F} \Rightarrow A^c (= \Omega \setminus A) \in \mathcal{F}$ (c) $\{A_n\}_{n=1}^{\infty}$ is in $\mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

(i) So, the smallest σ-field on Ω is $\mathcal{F} = \{\phi,\ \Omega\}$. It satisfies all the three requirements for a collection to be a σ-field. This is known as **trivial** σ-field on Ω.

(ii) Keeping in view the three requirements for a collection to be a σ-field the smallest σ-field containing a set A on Ω is given by $\mathcal{F} = \{\phi,\ A,\ A^c,\ \Omega\}$.

In our case A = {a} and Ω = {a, b, c, d}, therefore, required σ-field is given by

$$\mathcal{F} = \{\phi, \{a\}, \{b, c, d\}, \{a, b, c, d\} = \Omega\}$$

(iii) By definition of σ-field it is a collection of subsets of Ω which satisfies three conditions. Obviously, from school mathematics you know that the largest collection of subsets of any set is its power set. Further, power set is closed with respect to complement and union. So, the largest σ-field on Ω is the power set of Ω So, required σ-field on Ω is given by

$$\mathcal{F} = \left\{\begin{array}{l}\phi, \{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \\ \{c, d\}, \{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\}, \Omega\end{array}\right\} \quad \ldots (3.20)$$

Note 1: We have seen that both $\{\phi,\ \Omega\}$ and $P(\Omega) =$ power set, always form σ-field on Ω so you can say that both are trivial σ-fields on Ω. There exists lots of σ-fields on Ω other than these two. Example 2 given as follows and generalisation given after Example 2 explains a way of creating more σ-fields on Ω.

Example 2: If Ω = {a, b, c, d} then what is the smallest σ-field containing the sets A = {b} and B = {c, d} on Ω.

Solution: We know that a collection of subsets of a set Ω form a σ-field if it contains empty set, full set, closed under complement and closed under countable union. Therefore, a collection containing the sets A and B will form a σ-field on Ω if it contains following subsets of Ω

$$\mathcal{F} = \left\{\begin{array}{l}\phi,\ A,\ B,\ A^c,\ B^c,\ A \cap B,\ A \cap B^c,\ A^c \cap B,\ A^c \cap B^c,\ A \cup B, \\ A \cup B^c,\ A^c \cup B,\ A^c \cup B^c,\ A \Delta B,\ (A \Delta B)^c,\ \Omega\end{array}\right\}$$

In our case A = {b}, B = {c, d} and $\Omega = \{a, b, c, d\}$ so it becomes

$$\mathcal{F} = \{\phi, \{b\}, \{c, d\}, \{a, c, d\}, \{a, b\}, \{a\}, \{b, c, d\}, \{a, b, c, d\}\}$$

Generalisation: Every collection of subsets of a non-empty set Ω do not form a σ-field. But we can add more subsets of Ω in the given collection to form a

σ-field. How many subsets at the most we need in total to form a σ-field can be understood as follows.

From part (ii) of Example 1 note that if a collection has one set other than ϕ and Ω then we need at most $4\left(=2^{2^1}\right)$ subsets of Ω to form a σ-field. From Example 2 note that if a collection has two sets other than ϕ and Ω, then we need at most $16\left(=2^{2^2}\right)$ subsets of Ω to form a σ-field. In general, if a collection has n subsets of Ω other than ϕ and Ω then we need at the most 2^{2^n} subsets of Ω to form a σ-field with these given n subsets of Ω. But in practice we have less number of subsets compare to general case. For example, in Example 2 we have only 8 subsets instead of 16. … (3.21)

Now, you can try the following Self-Assessment Question.

SAQ 2

If Ω = {a, b, c, d} then what is the smallest σ-field containing the set $A = \{a\}$, B = {a, c] on Ω?

Now, we define set function and measure in the next section.

3.4 SET FUNCTION AND MEASURE

In Sec. 1.4 of Unit 1 of this course you have understood the definition of a function. You have seen that a function associates each element of a **set** to a unique element of another set or the same set. Set function is also a function but here domain is not a set like in the case of a function but it is a **collection of subsets of a non-empty set**. Because of the reason that elements of the domain of the set function are itself sets so it is given the name set function and is defined as follows.

Set Function: Let Ω be a non-empty set and $\mathcal{F}$ be any collection of subsets of Ω then a function $\mu : \mathcal{F} \to [0, +\infty]$ is known as a **set function**. … (3.22)

As mentioned earlier note that in a set function inputs are sets and outputs are real numbers or extended real numbers. Generally, output of a set function is some kind of measure of the input set. For example, it may be cardinality of the set, length or area or volume, etc. So, we can interpret it as a measure of the input set. As per the definition of a set function there is no requirement that members of its domain should form some algebraic structure. But many times, in particular, probability theory and measure theory we are interested to measure all members of a σ-field of subsets of Ω. So, instead of taking arbitrary collection of subsets of a set as a domain of a set function we generally take a σ-field of subsets of Ω as a domain of a set function. You will get more explanation regarding it in Unit 2 of the course MST-012. As mentioned earlier output of a set function is some kind of measure of the input set. So, let us define when a set function is called a measure.

Measure: Let Ω be a non-empty set and $\mathcal{F}$ be a collection of subsets of Ω which forms a σ-field then A set function $\mu : \mathcal{F} \to [0, +\infty]$ is called a measure if it satisfies the following two requirements … (3.23)

(i) $\mu(\phi) = 0$, i.e., measure of an empty set should be zero

(ii) μ is countably additive, i.e., whenever $A_1, A_2, A_3, \ldots \in \mathcal{F}$ with $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ and $A_m \cap A_n = \phi$ for all $m \neq n$ then $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$

Since μ is a function from $\mathcal{F}$ to $[0, +\infty]$ so it will associate a unique measure to each element of the σ-field $\mathcal{F}$ and so members of the σ-field $\mathcal{F}$ are known as **measurable sets**. … (3.24)

Thus, we can say that a measure is a set function which associates a measure to the members of the σ-field $\mathcal{F}$. In measure theory which is generally studied as a course in master degree programme of mathematics the triplet $(\Omega, \mathcal{A}, \mu)$ is known as **measure space**. Also Ω together with the σ-field $\mathcal{F}$ i.e., $(\Omega, \mathcal{F})$ is known as **measurable space**. Measurable space $(\Omega, \mathcal{F})$ means members of $\mathcal{F}$ are just one step away to get their measure, i.e., as soon as you define a measure μ on $\mathcal{F}$ then each member of $\mathcal{F}$ will get their corresponding measure. While measure space means members of $\mathcal{F}$ have been measured as per measure function μ. … (3.25)

Another important set function on extended real line is the length function which measures lengths of the intervals and is defined as follows.

Let $\mathcal{I}$ be the collection of all intervals on the real line. Let $I(a, b) \in \mathcal{I}$ be an interval with end points a and b where a and b are extended real numbers. The interval I(a, b) may be both sides open, both sides closed, one side open and one side closed. Then the set function $\lambda : \mathcal{I} \to [0, +\infty]$ defined by

$$\lambda(I(a, b)) = \begin{cases} |b - a|, & \text{if } a, b \in \mathbb{R} \\ +\infty, & \text{if either } a = -\infty \text{ or } b = +\infty \text{ or both} \end{cases} \qquad I(a, b) \in \mathcal{I} \quad \ldots (3.26)$$

is called the **length function**.

Let us now consider one example of a set function.

Example 3: Let $\Omega = \{a, b, c, d\}$ and $\mathcal{F} = \{\phi, \{a\}, \{b, c, d\}, \Omega\}$ then define a set function $\mu : \mathcal{F} \to [0, +\infty]$ by

$$\mu(A) = \begin{cases} n(A), & \text{if } A \text{ is a finite set} \\ +\infty, & \text{if } A \text{ is not a finite set} \end{cases} \qquad A \in \mathcal{F} \qquad \ldots (3.27)$$

where n(A) represents cardinality of the set A. Obtain measure of all the members of $\mathcal{F}$ associated by the set function μ.

Solution: By definition of μ, measures of members of $\mathcal{F}$ are given as follows.

$$\mu(\phi) = n(\phi) = 0, \quad \mu(\{a\}) = n(\{a\}) = 1,$$

$$\mu(\{b, c, d\}) = n(\{b, c, d\}) = 3, \quad \mu(X) = n(X) = 4$$

Remark 1: The set function defined in Example 3 actually counts the number of elements in the input set. So, it applies counting measure on the members of $\mathcal{F}$ so it is known as **counting measure**. … (3.28)

Now, you can try the following Self-Assessment Question.

SAQ 3

Let $\mathcal{I}$ be the collection of all intervals on the real line and $\lambda : \mathcal{I} \to [0, +\infty]$ be

the length function on $\mathcal{I}$ then write the image of the following members of $\mathcal{I}$ under the function λ.

(a) [4, 9] (b) (4, 9] (c) [4, 9) (d) (4, 9) (e) $[-\infty, 9]$
(f) $[-\infty, 9)$ (g) $[4, +\infty]$ (h) $[4, +\infty)$ (i) $[-\infty, +\infty]$ (j) $[-\infty, +\infty)$
(k) $(-\infty, +\infty]$ (l) $(-\infty, +\infty)$

3.5 CHARACTERISTIC, SIMPLE AND STEP FUNCTIONS

Suppose in a study we have 50 males and 52 female subjects. So, our universal set Ω which contains subjects of the study has 102 members/elements. If M and F denote respectively the sets of males and females subjects of the study then we have

- $M = \{m_1, m_2, m_3, \ldots, m_{50}\}$ is a subset of Ω, i.e., $M \subset \Omega$ and n(M) = 50, $n(\Omega) = 102$.
- $F = \{f_1, f_2, f_3, \ldots, f_{52}\}$ is a subset of Ω, i.e., $F \subset \Omega$ and n(F) = 52, $n(\Omega) = 102$.
- $\Omega = \{m_1, m_2, m_3, \ldots, m_{50}, f_1, f_2, f_3, \ldots, f_{52}\}$, $M^c = F$, $F^c = M$, $M \cup F = \Omega$.

If we define two functions $\chi_M : \Omega \to \{0, 1\}$ and $\chi_F : \Omega \to \{0, 1\}$ by

$$\chi_M(\omega) = \begin{cases} 1, & \omega \in M \\ 0, & \omega \notin M \end{cases} \quad \omega \in \Omega \quad \text{and} \quad \chi_F(\omega) = \begin{cases} 1, & \omega \in F \\ 0, & \omega \notin F \end{cases} \quad \omega \in \Omega$$

then what the functions χ_M and χ_F actually are indicating. The function χ_M is indicating whether the subject of the study is male or not. If outcome is 1 it means subject of the study is a male and if outcome is 0 it means subject is not a male. Similarly, the function χ_F is indicating whether the subject of the study is female or not. If outcome is 1 it means subject of the study is a female and if outcome is 0 it means subject is not a female. Such types of indicator functions have given the name characteristic function in mathematics because their outputs characterise the underlying set. So, in mathematics a characteristic function is defined as follows.

Characteristic Function: Let A be any subset of the universal set Ω then characteristic function of A is denoted by χ_A and defined as follows:

$\chi_A : \Omega \to \{0, 1\}$ define by

$$\chi_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases} \quad \omega \in \Omega \qquad \ldots (3.29)$$

Note 2: The characteristic function χ_A of the set A is also called the **indicator function** of A. … (3.30)

Let us do an example.

Example 4: Suppose a red die is thrown and E, F are the events defined as follows. E: getting an odd number, F: getting a multiple of 3. Write events E, F, $E \cup F$, and $E \cap F$. Visualise the characteristic functions of these four events in pictorial form.

Solution: When a red colour die (note here colour of the die is not any issue it is mentioned with the objective that you get exposure of such small points which are sometimes used in the statements of the question just to test the understanding level of the learner) is thrown then sample space is given by

$\Omega = \{1, 2, 3, 4, 5, 6\}$

Required four events are given by

$E = \{1, 3, 5\}, \quad F = \{3, 6\}, \quad E \cup F = \{1, 3, 5, 6\}, \quad E \cap F = \{3\}$

Now, visualisation of the characteristic functions of these four events in pictorial form are respectively shown in Fig. 3.1 (a) to (d).

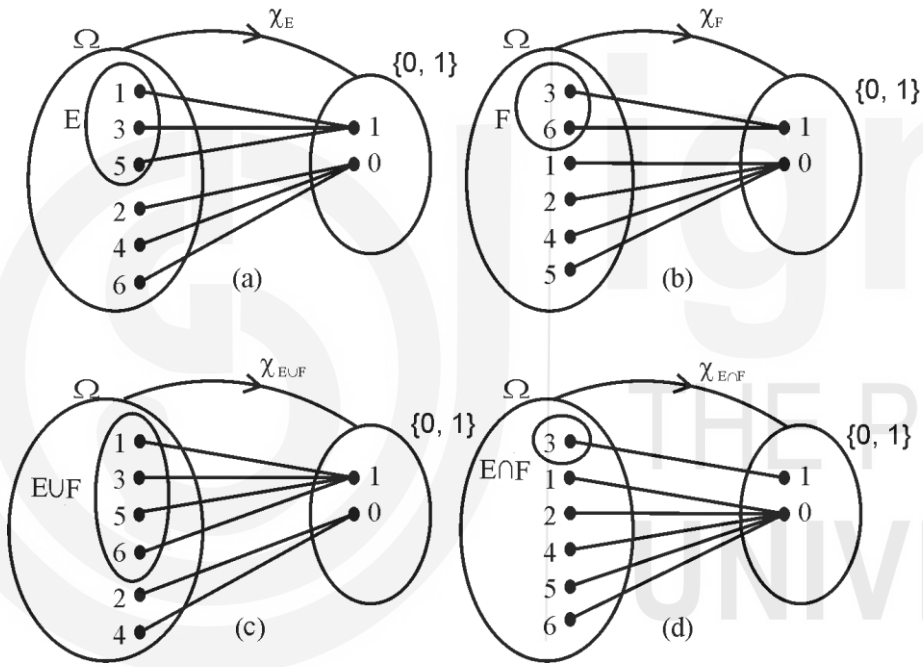


Fig. 3.1: Visualisation of the characteristic functions for the sets (events) (a) E (b) F (c) $E \cup F$ (d) $E \cap F$

Now, we discuss what is a simple function.

We just saw that characteristic function χ_A of a set A assumes only two values 0 and 1, refer (3.29). In Unit 1 of this course, you saw that signum function assumes only three values 0, – 1, 1, refer to (1.51). Similarly, a measurable function which assumes only a finite number of values is called a simple function. So, simple function is defined as follows.

Simple Function: Let $(\Omega, \mathcal{F}, \mu)$ be a measure space then a function $\phi : \Omega \to \mathbb{R}$ is said to be simple function if there is a finite disjoint class $\{E_1, E_2, E_3, \ldots, E_n\}$ of measurable sets (it means $E_i \in \mathcal{F}$) with $\bigcup_{i=1}^{n} E_i = \Omega$ and a finite set $\{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n\}$ of real numbers such that

$$\phi(x) = \alpha_i, \quad x \in E_i, \quad i = 1, 2, 3, 4, \ldots, n \qquad \ldots (3.31)$$

So, we can say that **simple function is a constant function in each** E_i. Characteristic function takes values 1 in E and 0 in E^c. So, it is a simple function. Finally, let us define a step function.

In Unit 1 of this course, we have defined Unit Step Function you may refer (1.49). Here we will define a step function in general. But before going through the definition of step function it is recommended that you should go through the definition of a partition of a closed and finite interval given by (5.20) and (5.21) in Unit 5 of this course.

Step Function: Let [a, b] be a finite and closed interval. A function $\phi : [a, b] \to \mathbb{R}$ is said to be a step function if:

(1) there is a partition P of [a, b], where by partition we mean a finite ordered set $\{x_0, x_1, x_2, x_3, \ldots, x_n\}$ of points of [a, b] such that
$a = x_0 < x_1 < x_2 < x_3 < \ldots < x_{n-1} < x_n = b.$

The n + 1 points $x_0, x_1, x_2, x_3, \ldots, x_n$ are called partition points of P and the n sub-intervals
$I_1 = [x_0, x_1], I_2 = [x_1, x_2], I_3 = [x_2, x_3], \ldots, I_i = [x_{i-1}, x_i], \ldots, I_n = [x_{n-1}, x_n]$
determined by n + 1 points of P are called the segments of the partition P.

(2) function ϕ assumes one and only one value in each **open** sub-intervals I_i, i = 1, 2, 3, …, n. So, there will exits n real numbers α_i such that
$\phi(x) = \alpha_i, \ \forall \ x \in (x_{i-1}, x_i), \ \ i = 1, 2, 3, \ldots, n$

(3) Also, values of the function ϕ at the end points of the sub-intervals I_i is irrelevant. … (3.32)

Remark 2: Like simple function step function is also constant in each sub-interval (x_{i-1}, x_i). Difference between them is in the definition of simple function $E_1, E_2, E_3, \ldots, E_n$ were sets while in the definition of step function we have intervals. We know that an interval is a set but a set may not be an interval. So, there may be simple functions which are not step functions but every step function is a simple function.

3.6 DISTANCE FUNCTION AND METRIC

So, far by distance between two points we mean straight line distance between them. For example:

(1) If you have two points on the real line, i.e., numbers x = 1, y = 4 $\in \mathbb{R}$ then distance between them is 3 units refer Fig. 3.2 (a).

(2) If you have two points $a = A(4, 0), \ b = B(0, \ 3) \in \mathbb{R}^2$ then using Pythagoras theorem in triangle OAB distance between them is 5 units refer to Fig. 3.2 (b).

But in machine learning we also use some other kinds of distances other than straight line distance between two points. The objective of this section is to make you familiar with some of those kinds of distances. First of all, let us explain what we mean by a distance function.

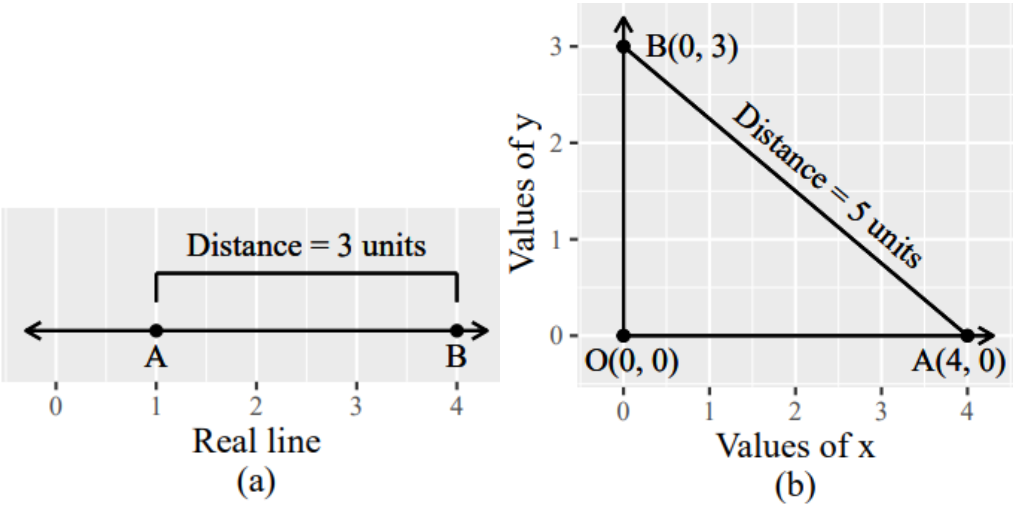


Fig. 3.2: Visualisation of the distance between two points A and B on (a) real line (b) the x-y plane

Distance Function: Let X be a non-empty set then distance function on X is a function which takes a pair of points from X as input and gives a non-negative real number as output. That is distance function looks like $d: X \times X \to [0, \infty]$, so domain of distance function d is $X \times X$ and range may be any subset of $[0, \infty]$. If $(a, b) \in X \times X$ then distance between two points a and b of X are denoted by d(a, b) and $d(a, b) \in [0, \infty]$. … (3.33)

We are interested in a distance function which satisfies following four properties (i) non-negative (ii) coincidence (iii) symmetry (iv) triangle inequality.

A distance function which satisfies these four properties is known as metric. A proper definition of a metric is given as follows.

Metric: Let X be a non-empty set then a distance function $d: X \times X \to [0, \infty]$ is called a metric on X if it satisfies following four properties.

(a) **Non-Negativity Property**: If x and y be any two points of X then we should have

$$d(x, y) \geq 0 \ \forall \ x, y \in X. \qquad \text{… (3.34)}$$

i.e., distance between any two points of X should be non-negative real number or distance between two points should not be negative.

(b) **Coincidence Property**: distance function d should satisfy coincidence property

i.e., $d(x, y) = 0$ if and only if $x = y \ \ \forall \ x, y \in X$ … (3.35)

i.e., distance of each point of X should be zero from itself and vice versa.

(c) **Symmetric Property**: If $x, y \in X$ then distance function d should satisfy

$$d(x, y) = d(y, x) \ \forall \ x, y \in X. \qquad \text{… (3.36)}$$

i.e., if x and y are any two points of X then distance of x from y and distance of y from x both should be equal.

(d) **Triangle Inequality**: If $x, y, z \in X$ then the distance function d should satisfy

$$d(x, y) \le d(x, z) + d(z, y) \quad \forall\ x, y, z \in X \qquad \ldots (3.37)$$

i.e., if x and y are any two points of X then distance between x and y should be less than equal to the sum of distances of x to z and z to y for any point z of X.

In (3.25) we explained the meaning of measure space. Similarly, if d is a metric on X then (X, d) is known as **metric space**. When we say (X, d) is a metric space it means every two points x and y of X have a well-defined distance which is denoted by d(x, y) and given by as per definition of d. … (3.38)

Some particular distance functions that we are going to discuss in this section are given as follows.

(a) Minkowski Distance

(b) Manhattan Distance

(c) Euclidean Distance

(d) Chebyshev Distance

(e) Hamming Distance

Let us discuss them one at a time.

(a) Minkowski Distance: A distance function $d : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty]$ defined by

$$d(x, y) = \sqrt[p]{|x_1 - y_1|^p + |x_2 - y_2|^p + |x_3 - y_3|^p + \ldots + |x_n - y_n|^p}$$

$$= \sqrt[p]{\sum_{i=1}^{n} |x_i - y_i|^p} \qquad \ldots (3.39)$$

where $x = A(x_1, x_2, x_3, \ldots, x_n),\ y = B(y_1, y_2, y_3, \ldots, y_n) \in \mathbb{R}^n,\ p \ge 1$

is called Minkowski distance between two points A and B of $\mathbb{R}^n$. … (3.40)

(b) Manhattan Distance: If p = 1 in Minkowski distance function then corresponding distance function is known as Manhattan distance function. So, a distance function $d : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty]$ defined by

$$d(x, y) = |x_1 - y_1| + |x_2 - y_2| + |x_3 - y_3| + \ldots + |x_n - y_n| = \sum_{i=1}^{n} |x_i - y_i| \qquad \ldots (3.41)$$

where $x = A(x_1, x_2, x_3, \ldots, x_n),\ y = B(y_1, y_2, y_3, \ldots, y_n) \in \mathbb{R}^n$

is called Manhattan distance between two points A and B of $\mathbb{R}^n$. If in a city, roads are like rectangular grids then distance travelled by a taxi to move from

one point to other point is given by Manhattan distance formula. So, it is also known as **taxicab distance** or **city block distance**. In the area of Manhattan in New York roads are like rectangular grids.

For example, the distance shown in Fig. 3.2 (a), gives Manhattan distance between points A and B where $p = 1, n = 1, \ x_1 = 1, y_1 = 4,$ so

$$AB = d(x, y) = \left(\sum_{i=1}^{1} |x_i - y_i|^1 \right)^{1/1} = |x_1 - y_1| = |1 - 4| = |-3| = -(-3) = 3$$

So, Manhattan distance between real numbers 1 and 4 is $d(1, 4) = 3.$ … (3.42)

Let us do an example to explain procedure of calculating Manhattan distance.

Example 5: Find the Manhattan distance between two points (1, 2) and (5, 7). Also, explain it geometrically.

Solution: We are given $(x_1, y_1) = (1, 2)$ and $(x_2, y_2) = (5, 7).$ By definition of Manhattan distance function distance between two points A(1, 2) and B(5. 7) is given by

$$\begin{aligned} AB = d(x, y) &= \sum_{i=1}^{2} |x_i - y_i| = |x_1 - y_1| + |x_2 - y_2| \\ &= |1 - 5| + |2 - 7| \\ &= |-4| + |-5| \\ &= 4 + 5 \\ &= 9 \end{aligned} \qquad \text{... (3.43)}$$

So, Manhattan distance between points A(1, 2) and B(5, 7) is 9 units. … (3.44)

Let us explain it geometrically, refer Fig. 3.3 (a) and (b)

Geometrically Manhattan distance gives distance of two points A and B travel by a taxi under the condition that roads in the city are like rectangular grids. Here straight-line distance will not work because you cannot travel through buildings. So, due to buildings you cannot travel straight and we have assumed that roads are like rectangular grids so you have only two options either you can move horizontally or vertically. Thus, it gives sum of the units of the distance that you have to run horizontally (Run) plus the units of distance you have to go vertically (Rise) to reach at the point B(5, 7) by starting from the point A(1, 2). Note that x-coordinate of the point A is 1 and x-coordinate of the point B is 5 so you have to move 4 (= 5 – 1) units of distance horizontally (Run). Further, y-coordinate of the point A is 2 and y-coordinate of the point B is 7 so you have to move 5 (= 7 – 2) units of distance vertically (Rise). Hence, in total you have to move 9 units of distance including both horizontal and vertical to reach at the point B by starting from the point A. Thus, Manhattan distance between two points A and B is 9 units. Also, note that there are more than one options to reach at the point B by starting from point A and traveling exactly 9 units of distance in total. Four such options are shown in Fig. 3.3 (b) in red, magenta, black and green colours.

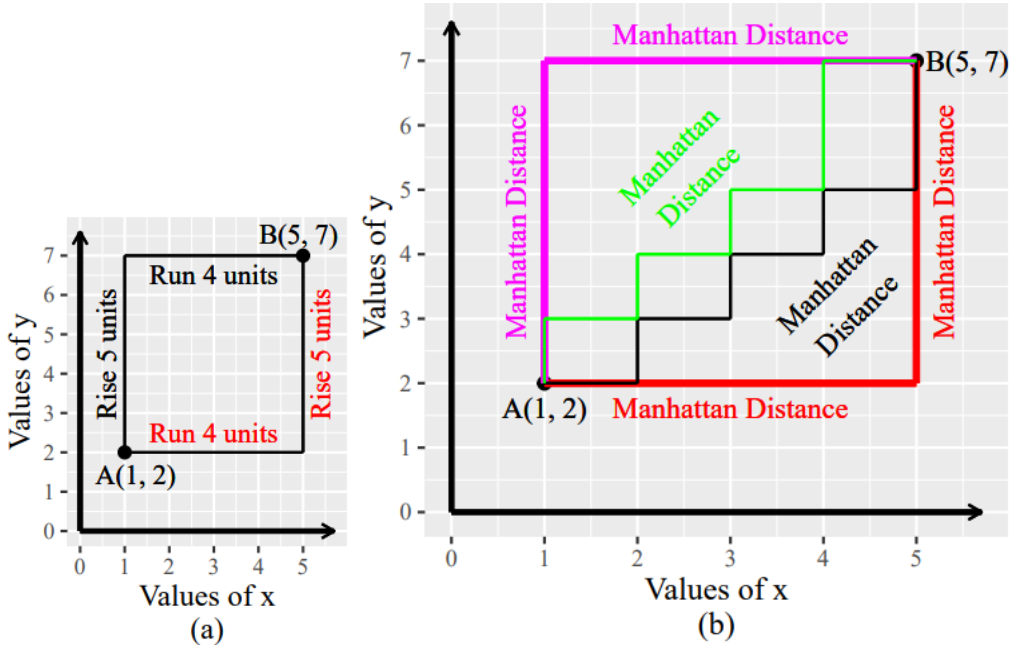


Fig. 3.3: Visualisation of (a) two-point A and B (b) Manhattan distance between two points A and B via four routes

(c) Euclidean Distance: If p = 2 in Minkowski distance function then corresponding distance function is known as Euclidean distance function. So, a distance function $d : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty]$ defined by

$$d(x, y) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + |x_3 - y_3|^2 + \ldots + |x_n - y_n|^2}$$

$$= \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} \qquad \ldots (3.45)$$

where $x = A(x_1, x_2, x_3, \ldots, x_n),\ y = B(y_1, y_2, y_3, \ldots, y_n) \in \mathbb{R}^n$

is called Euclidean distance between two points A and B of $\mathbb{R}^n$. This formula gives straight line distance between two points A and B. So, it is also known as **crow flying distance** because crow can fly over the buildings also. Here we are assuming that walking straight between two points of our interest is feasible.

For example, the distance shown in Fig. 3.2 (b), gives Euclidean distance between points A(4, 0) and B(0, 3) where
$p = 2, n = 2,\ (x_1, x_2) = (4, 0),\ (y_1, y_2) = (0, 3)$. So,

$$AB = d(x, y) = \left(\sum_{i=1}^{2} |x_i - y_i|^2\right)^{1/2} = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2} = \sqrt{|4-0|^2 + |0-3|^2}$$

$$= \sqrt{|4|^2 + |-3|^2} = \sqrt{16+9} = \sqrt{25} = 5$$

So, Euclidean distance between points A(4, 0) and B(0, 3) is 5 units. … (3.46)

Let us do an example to explain procedure of calculating Euclidean distance.

Example 6: Find the Euclidean distance between two points A(1, 2) and B(5, 7). Also, explain it geometrically.

Solution: We are given $(x_1, x_2) = (1, 2)$ and $(y_1, y_2) = (5, 7)$. By definition of

Euclidean distance function distance between two points A(1, 2) and B(5. 7) is given by

$$AB = d(x, y) = \sqrt{\sum_{i=1}^{2} |x_i - y_i|^2} = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2} = \sqrt{|1-5|^2 + |2-7|^2}$$

$$= \sqrt{|-4|^2 + |-5|^2} = \sqrt{16+25} = \sqrt{41} \simeq 6.403124 \text{ units} \quad \ldots (3.47)$$

So, Euclidean distance between points A(1, 2) and B(5, 7) is $\sqrt{41}$ units

… (3.48)

Let us explain it geometrically, refer to Fig. 3.4 (a) and (b).

Geometrically Euclidean distance is the shortest distance between the two points A and B. So, it can be obtained using Pythagoras theorem in triangle ABC as follows.

$$AB = \sqrt{AC^2 + CB^2} = \sqrt{4^2 + 5^2} = \sqrt{16+25} = \sqrt{41} \simeq 6.403124 \text{ units}$$

Thus, Euclidean distance between two points A and B is $\sqrt{41}$ units. If you are working in more than 2-dimension than you have to apply Pythagoras theorem more than once. Also, note that unlike Manhattan distance here you have only one option to reach at the point B by starting from point A and you have to travel a distance of $\sqrt{41}$ units. This unique path is shown in Fig. 3.4 (b).

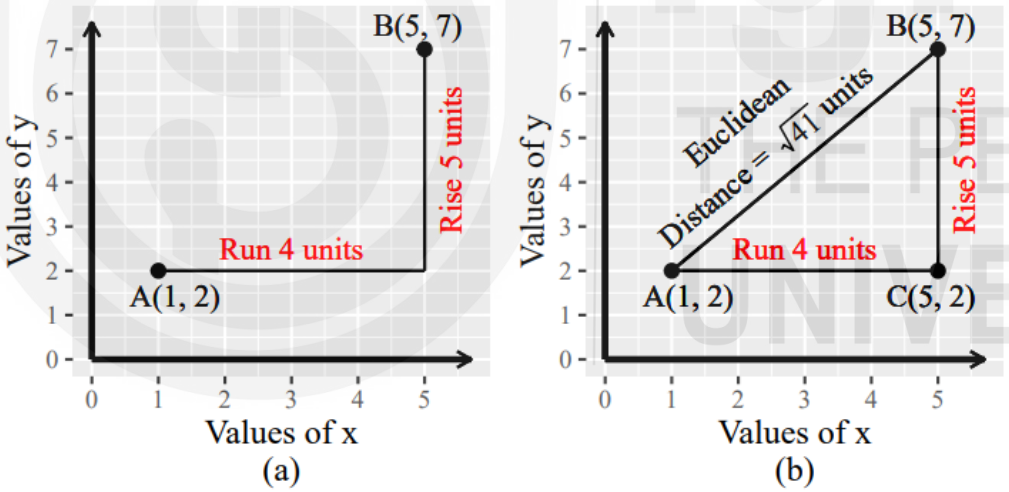


Fig. 3.4: Visualisation of (a) two points A and B (b) Euclidean distance between two points A and B via unique route

On comparing (3.44) and (3.48), we see that Euclidean distance between two points A and B is less than Manhattan distance between the same points A and B.

(d) Chebyshev Distance: If $p \to \infty$ in Minkowski distance function then corresponding distance function is known as Chebyshev distance function. So, a distance function $d : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty]$ defined by

$$d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|, |x_3 - y_3|, \ldots, |x_n - y_n|\} \quad \ldots (3.49)$$

where $x = A(x_1, x_2, x_3, \ldots, x_n)$, $y = B(y_1, y_2, y_3, \ldots, y_n) \in \mathbb{R}^n$

is called Chebyshev distance between two points A and B of $\mathbb{R}^n$. This formula looks at absolute distance between each corresponding coordinates of two points and select maximum among them.

For example, Chebyshev distance between points A(4, 0) and B(0, 3) is shown in Fig. 3.4 (a), and given as follows.

$$AB_{Chebyshev} = d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|, |x_3 - y_3|, \ldots, |x_n - y_n|\}$$

$$= \max\{|x_1 - y_1|, |x_2 - y_2|\}, \text{ where } x_1 = 4, x_2 = 0, y_1 = 0, y_1 = 3.$$

$$= \max\{|4 - 0|, |0 - 3|\} = \max\{|4|, |-3|\} = \max\{4, 3\} = 4.$$

So, Chebyshev distance between points A(4, 0) and B(0, 3) is 4 units. …(3.50)

Let us do an example to explain procedure of calculating Chebyshev distance.

Example 7: Find the Chebyshev distance between two points A(1, 2) and B(5, 7). Also, explain it geometrically.

Solution: We are given $(x_1, x_2) = (1, 2)$ and $(y_1, y_2) = (5, 7)$. By definition of Chebyshev distance function distance between two points A(1, 2) and B(5. 7) is given by

$$AB_{Chebyshev} = d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\},$$
$$\text{where } x_1 = 1, x_2 = 2, y_1 = 5, y_2 = 7.$$

$$= \max\{|1 - 5|, |2 - 7|\}$$

$$= \max\{|-4|, |-5|\} = \max\{4, 5\} = 5 \text{ units.} \quad \ldots (3.51)$$

So, Chebyshev distance between points A(1, 2) and B(5, 7) is 5 units… (3.52)

Let us explain it geometrically, refer to Fig. 3.4 (a).

Geometrically in 2-dimension Chebyshev distance is the maximum among Run and Rise refer Fig. 3.4 (a) where Run = 4 units and Rise = 5 units. So, Chebyshev distance between the two points A and B is max{Run, Rise}. So, $AB_{Chebyshev} = d(x, y) = \max\{\text{Run}, \text{Rise}\} = \max\{4, 5\} = 5$.

Remark 3: If points A and B both lie on a horizontal line or a vertical line then $AB_{Chebyshev} = AB_{Euclidean}$, otherwise $AB_{Chebyshev} \le AB_{Euclidean}$.

Let us do one example which will compare the shapes of a circle using different distance formulae with centre at origin and radius is of 1 unit.

Example 8: Draw a circle with centre at origin and radius of 1 unit when distance between two points is measured using (a) Manhattan distance formula (b) Euclidean distance formula (c) Minkowski distance formula for p > 2 (d) Chebyshev distance formula.

Solution: Centre of the circle is O(0, 0) and radius of the circle = r = 1 unit. We know that distance between two points $A(x_1, x_2)$ and $B(y_1, y_2)$ using different distance formulae are given by

$$AB_{Manhattan} = d(x, y) = |x_1 - y_1| + |x_2 - y_2| \quad \ldots (3.53)$$

$$AB_{Euclidean} = d(x, y) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2} \quad \ldots (3.54)$$

$$AB_{Minkowski} = d(x, y) = \sqrt[p]{|x_1 - y_1|^p + |x_2 - y_2|^p} \quad \ldots (3.55)$$

$$AB_{Chebyshev} = d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\} \quad \ldots (3.56)$$

We know that locus of a point which moves in a plane in such a way that its distance from a fixed point always remains constant is called a circle. Using definitions of distance between two points given by (3.53) to (3.56) shapes of the circles are given in Fig. 3.5 (a) to (d) respectively.

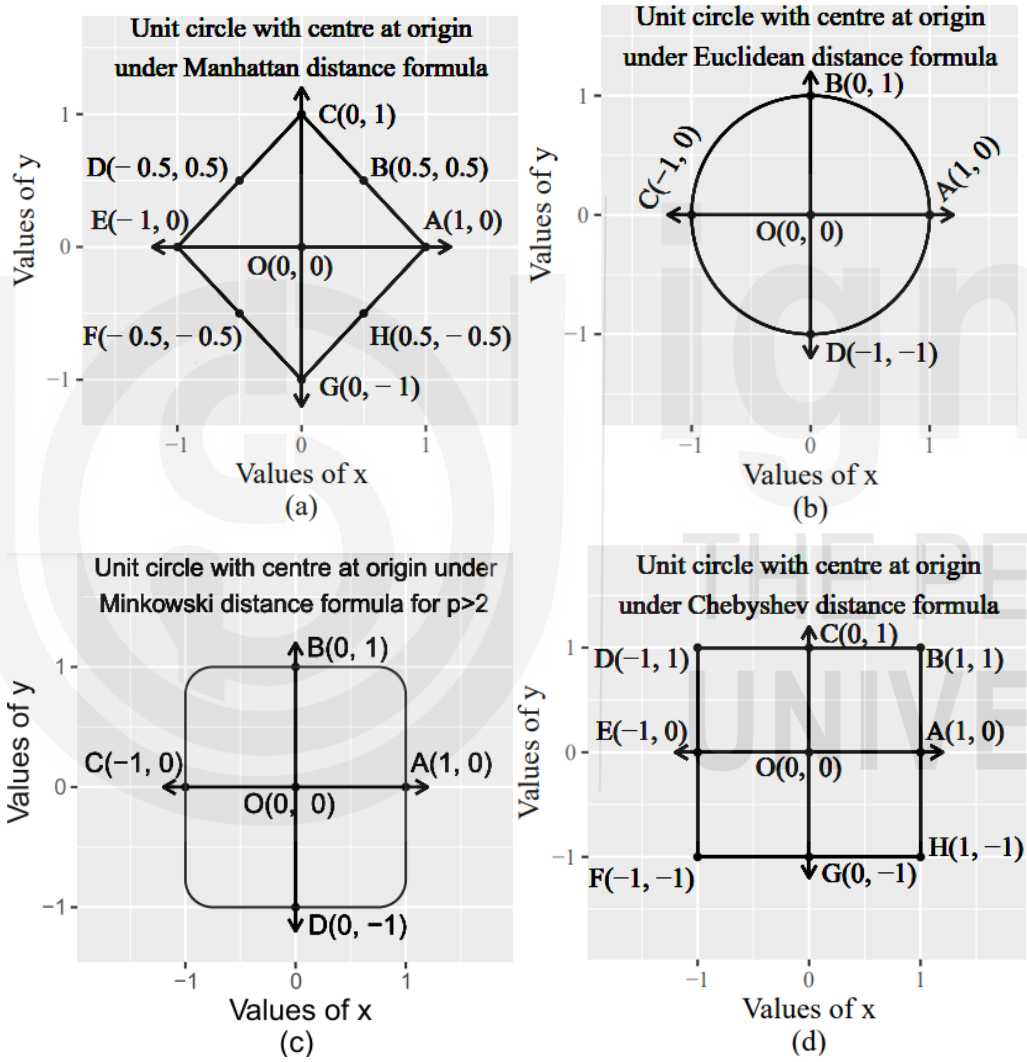


Fig. 3.5: Visualisation of unit circle with centre at the origin and formula for distance is (a) Manhattan (b) Euclidean (c) Minkowski for p > 2 (d) Chebyshev

(e) Hamming Distance: This method is used to obtain distance between two strings of the same length. Under this method distance between two strings is the number of counts of the positions at which the corresponding character in the two strings are different.

For example, Hamming distance between two words “PLANT” and “POINT” is given by counting the number of the positions at which the corresponding

character in the two words “PLANT” and “POINT” are different.

Letter no in the two words	Letter in the word PLANT	Corresponding letter in the word POINT	Match or do not match	Count of distance
1	P	P	Match	0
2	L	O	Do not match	1
3	A	I	Do not match	1
4	N	N	Match	0
5	T	T	Match	0
Hamming Distance between two words “PLANT” and “POINT” =				2

Now, you can try the following Self-Assessment Question.

SAQ 4

In R we have a built-in data set “iris”. Screenshot of the first five rows of this data set is given as follows.

```
> head(iris,5)
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1         3.5          1.4         0.2  setosa
2          4.9         3.0          1.4         0.2  setosa
3          4.7         3.2          1.3         0.2  setosa
4          4.6         3.1          1.5         0.2  setosa
5          5.0         3.6          1.4         0.2  setosa
```

Note that the first four variables of this data set are numeric. So, assuming each row of this data set as a point in 4-dimension. For example, point corresponding to the first row is (5.1, 3.5, 1.4, 0.2). Find the distance between the points corresponding to the first and third rows using the distance formula given by (a) Manhattan (b) Euclidean (c) Chebyshev (d) Minkowski for p = 4.

3.7 SUMMARY

A brief summary of what we have covered in this unit is given as follows:

- If we include two symbols $-\infty$ and $+\infty$ in the set of all real numbers then this extended set is called the **extended set of real numbers** and is denoted by $\mathbb{R}^*$. i.e., $\mathbb{R}^* = (-\infty, +\infty) \cup \{-\infty, +\infty\} = [-\infty, +\infty]$.
- An operation $*$ (say) on a non-empty set X is said to be binary if $a * b \in X$, for all $a, b \in X$. If $*$ is a binary operation on a non-empty set X then we also say that X is **closed with respect to the operation** $*$. If so then we also say that **closure property** holds in X with respect to the operation $*$.
- Let Ω be a non-empty set and $\mathcal{F}$ be a collection of subsets of Ω then we say that the class $\mathcal{F}$ forms a **sigma algebra** (σ-algebra) or **sigma field** (σ**-field**) if (i) $\phi, \Omega \in \mathcal{F}$, (ii) $A \in \mathcal{F} \Rightarrow A^c (= \Omega \setminus A) \in \mathcal{F}$, (iii) $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$
- **Set Function**: Let Ω be a non-empty set and $\mathcal{F}$ be any collection of subsets of Ω then a function $\mu : \mathcal{F} \to [0, +\infty]$ is known as a **set function**.
- **Measure**: Let Ω be a non-empty set and $\mathcal{F}$ be a collection of subsets of Ω which forms a σ-field then a set function $\mu : \mathcal{F} \to [0, +\infty]$ is called a

measure if it satisfies the following two requirements (i) $\mu(\phi) = 0$, (ii) μ is countably additive.

- The set function $\lambda : \mathcal{I} \to [0, +\infty]$ defined by

$$\lambda\big(I(a, b)\big) = \begin{cases} |b - a|, & \text{if } a, b \in \mathbb{R} \\ +\infty, & \text{if either } a = -\infty \text{ or } b = +\infty \text{ or both} \end{cases} \qquad I(a, b) \in \mathcal{I}$$

is called the **length function**.

- **Characteristic Function**: Let A be any subset of the universal set Ω then characteristic function of A is denoted by χ_A and defined as follows:

$\chi_A : \Omega \to \{0, 1\}$ define by

$$\chi_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases} \qquad \omega \in \Omega$$

- **Simple Function**: Let $(\Omega, \mathcal{F}, \mu)$ be a measure space then a function $\phi : \Omega \to \mathbb{R}$ is said to be a simple function if there is a finite disjoint class $\{E_1, E_2, E_3, \ldots, E_n\}$ of measurable sets (it means $E_i \in \mathcal{F}$) with $\bigcup_{i=1}^{n} E_i = \Omega$ and a finite set $\{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n\}$ of real numbers such that

$\phi(x) = \alpha_i, \quad x \in E_i, \; i = 1, 2, 3, 4, \ldots, n$

- **Step Function**: Let [a, b] be a finite and closed interval. A function $\phi : [a, b] \to \mathbb{R}$ is said to be a step function if:

$\phi(x) = \alpha_i, \; \forall \; x \in (x_{i-1}, x_i), \; i = 1, 2, 3, \ldots, n$ where the ordered set $\{x_0, x_1, x_2, x_3, \ldots, x_n\}$ forms a partition of [a, b].

- **Distance Function**: Let X be a non-empty set then distance function on X is a function which takes a pair of points from X as input and gives a non-negative real number as output. That is distance function looks like $d : X \times X \to [0, \infty]$.

- A distance function which satisfies four properties (i) non-negative (ii) coincidence (iii) symmetry, and (iv) triangle inequality is called a metric.

- Here we have discussed five types of distance functions namely:

  (a) Minkowski Distance (b) Manhattan Distance (c) Euclidean Distance

  (d) Chebyshev Distance, and (e) Hamming Distance

3.8 TERMINAL QUESTIONS

1. If Ω = {a, b, c, d} then what is the smallest σ-algebra containing the sets A = {a, b} and B = {b, c} on Ω.

2. Consider the experiment of throwing a fair die. If Ω be the sample space of this experiment and suppose $\mathcal{F}$ is the smallest σ-algebra on Ω containing the event E = {2, 5} then answer the following questions.

(a) Write Ω as a set (b) Write $\mathcal{F}$ (c) Define appropriate probability measure (d) Specify probability space (e) What is measure of the event {2, 5}.

3. Prove that every non-empty set can be treated as a metric space.

3.9 SOLUTIONS/ANSWERS

Self-Assessment Questions (SAQs)

1. As per the rules of the world of extended real numbers appropriate number or symbol $-\infty$ or $+\infty$ at the place of ? sign are given as follows.

   (i) $(-3)(-\infty)=+\infty$ (ii) $(5)(-\infty)=-\infty$ (iii) $(-10)(+\infty)=-\infty$
   (iv) $-5+(-\infty)=-\infty$ (v) $15+(+\infty)=+\infty$ (vi) $-7+(-\infty)=-\infty$
   (vii) $-9+(+\infty)=+\infty$ (viii) $(0)(-\infty)=0$ (ix) $(0)(+\infty)=0$
   (x) $(+\infty)+(-\infty)=$ Not defined

2. We know that the smallest σ-field containing the sets A and B on Ω is given by

$$\mathcal{F}=\left\{\begin{matrix}\phi,\ A,\ B,\ A^c,\ B^c,\ A\cap B,\ A\cap B^c,\ A^c\cap B,\ A^c\cap B^c,\ A\cup B,\\ A\cup B^c,\ A^c\cup B,\ A^c\cup B^c,\ A\Delta B,\ (A\Delta B)^c,\ \Omega\end{matrix}\right\}$$

   In our case A = {a}, B = {a, c} and $\Omega=\{a, b, c, d\}$ so it becomes

$$\mathcal{F}=\{\phi,\ \{a\},\ \{a, c\},\ \{b, c, d\},\ \{b, d\},\ \{c\},\ \{a, b, d\},\ \{a, b, c, d\}\}$$

3. We know that length function $\lambda:\mathcal{I}\to[0, +\infty]$ is defined by

$$\lambda(I(a, b))=\begin{cases}|b-a|, & \text{if } a, b\in\mathbb{R}\\ +\infty, & \text{if either } a=-\infty \text{ or } b=+\infty \text{ or both}\end{cases}\qquad I(a, b)\in\mathcal{I}$$

   Therefore, images of the given members of $\mathcal{I}$ under the length function λ are given by

   (a) $\lambda([4, 9])=|9-4|=|5|=5$

   (b) $\lambda((4, 9])=|9-4|=|5|=5$

   (c) $\lambda([4, 9))=|9-4|=|5|=5$

   (d) $\lambda((4, 9))=|9-4|=|5|=5$

   (e) $\lambda([-\infty, 9])=+\infty$

   (f) $\lambda([-\infty, 9))=+\infty$

   (g) $\lambda([4, +\infty])=+\infty$

   (h) $\lambda([4, +\infty))=+\infty$

   (i) $\lambda([-\infty, +\infty])=+\infty$

   (j) $\lambda([-\infty, +\infty))=+\infty$

(k) $\lambda\left((-\infty, +\infty]\right) = +\infty$

(l) $\lambda\left((-\infty, +\infty)\right) = +\infty$

4. Let us denote points corresponding to the first and the third rows by A and B respectively. So, we have

$A\left(5.1,\ 3.5,\ 1.4,\ 0.2\right) = A(x_1, x_2, x_3, x_4)$ and

$B\left(4.7,\ 3.2,\ 1.3,\ 0.2\right) = B(y_1, y_2, y_3, y_4)$

Now, required distances are given by

(a) $AB_{Manhattan} = d(x, y) = |x_1 - y_1| + |x_2 - y_2| + |x_3 - y_3| + |x_4 - y_4|$

$= |5.1 - 4.7| + |3.5 - 3.2| + |1.4 - 1.3| + |0.2 - 0.2|$

$= |0.4| + |0.3| + |0.1| + |0| = 0.4 + 0.3 + 0.1 + 0 = 0.8$

(b) $AB_{Euclidean} = d(x, y) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + |x_3 - y_3|^2 + |x_4 - y_4|^2}$

$= \sqrt{|5.1 - 4.7|^2 + |3.5 - 3.2|^2 + |1.4 - 1.3|^2 + |0.2 - 0.2|^2}$

$= \sqrt{|0.4|^2 + |0.3|^2 + |0.1|^2 + |0|^2} = \sqrt{0.16 + 0.09 + 0.01 + 0}$

$= \sqrt{0.26} \simeq 0.509902$

(c) $AB_{Chebyshev} = d(x, y) = \max\left\{|x_1 - y_1|,\ |x_2 - y_2|,\ |x_3 - y_3|,\ |x_4 - y_4|\right\}$

$= \max\left\{|5.1 - 4.7|,\ |3.5 - 3.2|,\ |1.4 - 1.3|,\ |0.2 - 0.2|\right\}$

$= \max\left\{|0.4|,\ |0.3|,\ |0.1|,\ |0|\right\} = \max\left\{0.4,\ 0.3,\ 0.1,\ 0\right\} = 0.4$

(d) $AB_{\substack{Minkowski \\ (p=4)}} = d(x, y) = \sqrt[4]{|x_1 - y_1|^4 + |x_2 - y_2|^4 + |x_3 - y_3|^4 + |x_4 - y_4|^4}$

$= \sqrt[4]{|5.1 - 4.7|^4 + |3.5 - 3.2|^4 + |1.4 - 1.3|^4 + |0.2 - 0.2|^4}$

$= \sqrt[4]{|0.4|^4 + |0.3|^4 + |0.1|^4 + |0|^4} = \sqrt[4]{0.0256 + 0.0081 + 0.0001 + 0}$

$= \sqrt[4]{0.0338} \simeq 0.428775$

Terminal Questions

1. We know that the smallest σ-algebra containing the sets A and B on Ω is given by

$$\mathcal{F} = \left\{\begin{array}{l} \phi,\ A,\ B,\ A^c,\ B^c,\ A \cap B,\ A \cap B^c,\ A^c \cap B,\ A^c \cap B^c,\ A \cup B, \\ A \cup B^c,\ A^c \cup B,\ A^c \cup B^c,\ A\Delta B,\ \left(A\Delta B\right)^c,\ \Omega \end{array}\right\}$$

In our case A = {a, b}, B = {b, c} and $\Omega = \{a, b, c, d\}$ so it becomes

$$\mathcal{F} = \left\{\begin{array}{l} \phi,\ \{a, b\},\ \{b, c\},\ \{c, d\},\ \{a, d\},\ \{b\},\ \{a\},\ \{c\},\ \{d\},\ \{a, b, c\}, \\ \{a, b, d\},\ \{b, c, d\},\ \{a, c, d\},\ \{a, c\},\ \{b, d\},\ \{a, b, c, d\} \end{array}\right\}$$

i.e., $\mathcal{F} = P\left(\Omega\right)$ = the largest σ-algebra on Ω

2 (a) $\Omega = \left\{1, 2, 3, 4, 5, 6\right\}$

(b) The smallest σ-algebra on Ω containing the event E = {2, 5} is given by $\mathcal{F} = \{\phi, E, E^c, \Omega\} = \{\phi, \{2, 5\}, \{1, 3, 4, 6\}, \Omega\}$

(c) Appropriate probability measure is defined by $\mu(E) = \dfrac{n(E)}{n(\Omega)}, \quad E \in \mathcal{F}$

(d) Probability space in this case is

$$(\Omega, \mathcal{F}, \mu) = \left(\{1, 2, 3, 4, 5, 6\},\ \{\phi, \{2, 5\}, \{1, 3, 4, 6\}, \Omega\},\ \mu(E) = \frac{n(E)}{n(\Omega)}\right)$$

(e) By definition of probability measure as define in part (c), probability measure of the event E = {2, 5} is given by $\mu(\{2, 5\}) = \dfrac{n(\{2, 5\})}{n(\Omega)} = \dfrac{2}{6} = \dfrac{1}{3}$

3. Let X be a non-empty set. Define a function $f : X \times X \to \mathbb{R}$ by

$$d(x, y) = \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases} \qquad x, y \in X \qquad \ldots (3.57)$$

We claim that d is a metric on X.

(i) **Non-negativity**: By definition of d it takes only two values 0 and 1. So, d(x, y) ≥ 0

(ii) **Coincidence**: d(x, y) = 0 iff x = y [Using (3.57)]

(iii) **Symmetry**: **Case I**: If x = y, then y = x. So, d(x, y) = 0 = d(y, x)

Case II: If x ≠ y, then y ≠ x. So, d(x, y) = 1 = d(y, x)

(iv) **Triangle Inequality**: Required to prove

d(x, y) ≤ d(x, z) + d(z, y) … (3.58)

Case I: x = y, but y ≠ z, then x ≠ z. So, (3.58) reduces to 0 ≤ 1 + 1 which is true

Case II: x = y, and y = z, then x = z. So, (3.58) reduces to 0 ≤ 0 + 0 which is true

Case III: x ≠ y, but y = z, then x ≠ z. So, (3.58) reduces to 1 ≤ 1 + 0 which is true

Case IV: x ≠ y, and y ≠ z, then

Sub case I: x = z. So, (3.58) reduces to 1 ≤ 0 + 1 which is true

Sub case II: x ≠ z. So, (3.58) reduces to 1 ≤ 1 + 1 which is true

Thus, (3.58) holds for all $x, y, z \in X$.

Hence, d is a metric on X and so (X, d) is a metric space.

Remark 4: This metric d is called discrete metric on X. … (3.59)

UNIT 4

SEQUENCES AND SERIES

Structure

4.1 Introduction

Expected Learning Outcomes

4.2 Sequence in the Family of Real Numbers

4.3 Some Results Related to Convergence and Divergence of Sequence and Their Use

4.4 Infinite Series

4.5 Absolute and Conditional Convergence

4.6 Summary

4.7 Terminal Questions

4.8 Solutions/Answers

4.1 INTRODUCTION

You have heard or used the word sequence in some context like putting things in a sequence. In this unit, you will study the sequence of real numbers. In Sec. 4.2 we will define what is the meaning of sequence in mathematics in particular real sequence. In the same section, we will also visualise the sequence geometrically. If after a certain value of n terms of a sequence come closer and closer to a particular real number as n increases then we say that sequence converges to that real number. In Sec. 4.3 first we will state some results related to the convergence of the sequence and then we will apply them to test the convergence of the sequence. In Sec. 4.4 we will do similar things for the series of positive terms. If a series has infinite terms of both signs positive and negative then the underlying series is known as a series of arbitrary terms. Some terms related to a series of arbitrary terms are discussed in Sec. 4.5 in brief.

What we have discussed in this unit is summarised in Sec. 4.6. Self-Assessment Questions (SAQs) have been given in some sections which are generally based on the content discussed in that section. But to give you a good practice of what we have discussed in this unit, some more questions based on the entire unit are given in Sec. 4.7 under the heading Terminal Questions. Due to the reason mentioned in Sec. 1.1 of Unit 1 of this course, solutions of all the SAQs and Terminal Questions are given in Sec. 4.8.

In the next unit, you will study Riemann integration.

Expected Learning Outcomes

After completing this unit, you should be able to:

- define what is a sequence and what is meaning of convergence and divergence of a sequence and how to test the convergence of a sequence;
- define what is a series and what is meaning of convergence and divergence of a series and how to test the convergence of a series; and
- explain the difference between absolute and conditional convergence of a series.

4.2 SEQUENCE IN THE FAMILY OF REAL NUMBERS

In Units 1 and 2 of this course, you have studied function. You know that every function has a domain and range. **Sequence** is also a function having domain as the family of natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$ and range may be any set but, in this course, we will discuss only sequences of real numbers. Hence, the range of every sequence that we are going to consider in this course will be a subset of $\mathbb{R}$, the set of all real numbers. So, throughout the course sequence will mean a sequence of real numbers known as a real sequence. So, let us formally define a real sequence.

Real Sequence: A real sequence is a function whose domain is the set of all natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$ and range may be any subset of $\mathbb{R}$, the set of all real numbers. In general, a sequence is written as

$$a_1, a_2, a_3, a_4, \ldots, a_n, \ldots \qquad \ldots (4.1)$$

In short, we write it as $\{a_n\}$ or (a_n) or $< a_n >$. $\qquad \ldots (4.2)$

In (4.1) or (4.2), a_1 is known as the first term of the sequence $< a_n >$. Similarly, $a_2, a_3, \ldots$ are known as the second, and third terms respectively and so on. Therefore, a_n is known as n^{th} the term of the sequence $< a_n >$. A sequence may be represented in any of the following four ways.

First Way of Representing a Sequence: A sequence can be represented by writing the first few terms of the sequence till a definite rule for writing down other terms becomes clear. For example, 1, 4, 9, 16, 25, … is a sequence.

Second Way of Representing a Sequence: A sequence can also be represented by giving a formula for its nth term. For example, $< \frac{n+1}{n} >$.

Third Way of Representing a Sequence (Recursive Relation): A sequence can also be represented by specifying its first few terms and a formula to write the remaining terms in terms of preceding terms. For example,
$a_1 = 1, a_2 = 1, a_{n+1} = a_n + a_{n-1} \;\; \forall \;\; n \geq 2, n \in \mathbb{N}$

Fourth Way of Representing a Sequence: Sometimes sequence is such that it cannot be represented by giving a single formula for its n term. But its terms are generated by two or more relations. For example,

$$a_n = \begin{cases} \frac{n}{2}, & \text{if } n \text{ is even} \\ -\frac{n-1}{2}, & \text{if } n \text{ is odd} \end{cases}$$

If we write this sequence using the first way of representing a sequence then it can be written as $0, 1, -1, 2, -2, 3, -3, 4, -4, \ldots$

Let us consider six examples of sequences, i.e., Sequence 1 to Sequence 6 given as follows.

Sequence 1: Consider the function $f : \mathbb{N} \to \mathbb{R}$ defined by $f(n) = \frac{1}{n}, \ n \in \mathbb{N}$.

The sequence defined by this function can be written as $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ … (4.3)

or simply by $< \frac{1}{n} >$. If we take $f(n) = a_n, \ n \in \mathbb{N}$, then this sequence can be written as $a_n = \frac{1}{n}$. … (4.4)

So, the first term = $a_1 = 1$, second term = $a_2 = \frac{1}{2}$, third term = $a_3 = \frac{1}{3}$, and so on.

Sequence 2: Consider the function $f : \mathbb{N} \to \mathbb{R}$ defined by $f(n) = \frac{n+1}{n}, \ n \in \mathbb{N}$.

The sequence defined by this function can be written as $2, \frac{3}{2}, \frac{4}{3}, \frac{5}{4}, \ldots$ (4.5)

or simply by $< \frac{n+1}{n} >$. In this sequence, we have $a_n = \frac{n+1}{n}$. … (4.6)

So, the first term = $a_1 = 2$, second term = $a_2 = \frac{3}{2}$, third term = $a_3 = \frac{4}{3}$, and so on.

Sequence 3: Consider the function $f : \mathbb{N} \to \mathbb{R}$ defined by $f(n) = 2n - 1, \ n \in \mathbb{N}$.

The sequence defined by this function can be written as $1, 3, 5, 7, \ldots$ … (4.7)
or simply by $< 2n - 1 >$. In this sequence, we have $a_n = 2n - 1$. … (4.8)
So, the first term = $a_1 = 1$, second term = $a_2 = 3$, third term = $a_3 = 5$, and so on.

Sequence 4: Consider the function $f : \mathbb{N} \to \mathbb{R}$ defined by $f(n) = -2n + 1, \ n \in \mathbb{N}$.

The sequence defined by this function can be written as $-1, -3, -5, \ldots$ (4.9)
or simply by $< -2n + 1 >$. In this sequence, we have $a_n = -2n + 1$. … (4.10)
So, the first term = $a_1 = -1$, second term = $a_2 = -3$, third term = $a_3 = -5$, and so on.

Sequence 5: Consider the function $f : \mathbb{N} \to \mathbb{R}$ defined by $f(n) = (-1)^n, \ n \in \mathbb{N}$.

The sequence defined by this function can be written as $-1, 1, -1, 1, \ldots$ (4.11)
or simply by $< (-1)^n >$. In this sequence, we have $a_n = (-1)^n$. … (4.12)

Clearly the first term = $a_1 = -1$, second term = $a_2 = 1$, third term = $a_3 = -1$, and so on.

Sequence 6: Consider the function $f : \mathbb{N} \to \mathbb{R}$ defined by $f(n) = (-1)^n n,\ n \in \mathbb{N}$.

The sequence defined by this function can be written as $-1, 1, -2, 2, \ldots$ (4.13) or simply by $<(-1)^n n>$. In this sequence, we have $a_n = (-1)^n n$. … (4.14) So, the first term = $a_1 = -1$, second term = $a_2 = 1$, third term = $a_3 = -2$, and so on.

Since sequence is a function, so you can make its graph by taking values of n on horizontal axis and values of the terms of the sequence on vertical axis. Following this rule graphs of the sequences given by (4.3), (4.5), (4.7), (4.9), (4.11) and (4.13) are shown in Fig. 4.1 (a) to (f) respectively. Here values of terms of the sequence are plotted by solid points and by a vertical line from horizontal axis to the corresponding solid point. The vertical lines from horizontal line up to the solid point will help you to understand where the values are going as n increases.

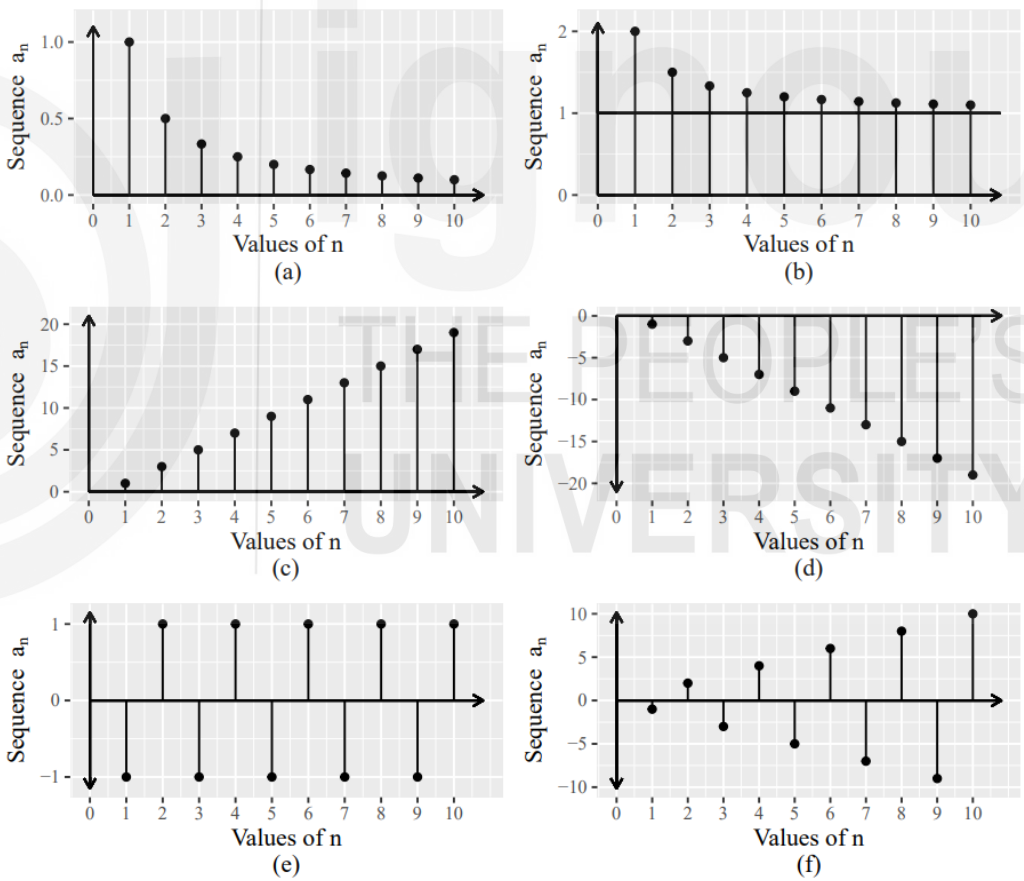


Fig. 4.1: Visualisation of Sequences 1 to 6 (a) 1 (b) 2 (c) 3 (d) 4 (e) 5 (f) 6

Let us comment on the behaviours of the sequences shown in Fig. 4.1 (a) to (f) one at a time.

Comment on the Behaviour of Sequence 1: From Fig. 4.1 (a) you see that terms of the sequence $<\frac{1}{n}>$ are coming closer and closer to horizontal axis as n increases. In fact, terms of this sequence are coming closer and closer to

the value 0 $\left[\because \lim_{n\to\infty} a_n = 0\right]$. If $\lim_{n\to\infty} a_n$ exists and equal to a finite number k (say) then we say that sequence $< a_n >$ **converges to finite number k**. Here value of k is 0.

Comment on the Behaviour of Sequence 2: Following similar approach, the sequence $< \frac{n+1}{n} >$ shown in Fig. 4.1 (b) converges to the number 1 $\left[\because \lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{n+1}{n} = \lim_{n\to\infty}\left(1+\frac{1}{n}\right) = 1+0 = 1\right]$. You can see it in Fig. 4.2 (b) as n increases terms of the sequence are coming closer and closer to the horizontal line which is 1 unit above the horizontal axis.

Comment on the Behaviour of Sequence 3: In Fig. 4.1 (c) you see that terms of the sequence $< 2n-1 >$ are increasing as n increases. In fact, given any real number no matter how large it is there exist terms in this sequence which are even greater than that real number. This is an important point, so let us consider two examples to explain this point.

(i) If you consider the real number 10000 then there exists $m = 5001$ such that $a_n > 10000 \;\; \forall\; n \geq m = 5001$. $\left[\because 2n-1 = 10000 \Rightarrow 2n = 10001 \Rightarrow n = 5000.5\right]$

(ii) If you consider the real number 2^{1000} then there exists $m = 2^{999}+1$ such that

$$a_n > 2^{1000} \;\; \forall\; n \geq m = 2^{999}+1. \quad \left[\begin{array}{l} \because 2n-1 = 2^{1000} \Rightarrow 2n = 2^{1000}+1 \\ \Rightarrow n = \frac{2^{1000}+1}{2} \Rightarrow n = \frac{2^{1000}}{2} + \frac{1}{2} = 2^{999}+0.5 \end{array}\right]$$

If a sequence behaves like this then we say that sequence diverges to infinity. So, sequence $< 2n-1 >$ diverges to ∞. Also, you can note that $\lim_{n\to\infty} a_n = \lim_{n\to\infty}(2n-1) = \infty$.

Comment on the Behaviour of Sequence 4: In Fig. 4.1 (d) you see that terms of the sequence $< -2n+1 >$ are decreasing as n increases. In fact, given any real number no matter how small it is there exist terms in this sequence which are even less than that real number. Like sequence 3, let us consider two examples to explain this point.

(i) If you consider the real number -10000 then there exists $m = 5001$ such that $a_n < -10000 \;\; \forall\; n \geq m = 5001$. $\left[\begin{array}{l} \because -2n+1 = -10000 \\ \Rightarrow -2n = -10001 \Rightarrow n = 5000.5 \end{array}\right]$

(ii) If you consider the real number -2^{1000} then there exists $m = 2^{999}+1$ such that

$$a_n < -2^{1000} \;\; \forall\; n \geq m = 2^{999}+1. \quad \left[\begin{array}{l} \because -2n+1 = -2^{1000} \Rightarrow -2n = -2^{1000}-1 \\ \Rightarrow n = \frac{-2^{1000}-1}{-2} \Rightarrow n = \frac{2^{1000}}{2} + \frac{1}{2} = 2^{999}+0.5 \end{array}\right]$$

If a sequence behaves like this then we say that sequence diverges to minus infinity. So, sequence $< -2n+1 >$ diverges to $-\infty$. Also, you can note that

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty}(-2n+1) = -\infty.$$

Comment on the Behaviour of Sequence 5: In Fig. 4.1 (e) you see that even number terms $a_2, a_4, a_6, \ldots$ of the sequence $<(-1)^n>$ converges to 1 while odd number terms $a_1, a_3, a_5, \ldots$ of the sequence $<(-1)^n>$ converges to -1. So, as n increases its behaviour depends on whether n is even or odd. That is

$$a_n = \begin{cases} 1, & \text{if } n \text{ is even} \\ -1, & \text{if } n \text{ is odd} \end{cases}$$

So,

$$\lim_{n\to\infty} a_n = \begin{cases} \lim\limits_{n\to\infty} 1, & \text{if } n \text{ is even} \\ \lim\limits_{n\to\infty}(-1), & \text{if } n \text{ is odd} \end{cases}$$

$$= \begin{cases} 1, & \text{if } n \text{ is even} \\ -1, & \text{if } n \text{ is odd} \end{cases}$$

If a sequence like this converges to more than one finite number then we say that sequence oscillates finitely. So, sequence $<(-1)^n>$ oscillates finitely.

Comment on the Behaviour of Sequence 6: In Fig. 4.1 (f) you see that even number terms $a_2, a_4, a_6, \ldots$ of the sequence $<(-1)^n n>$ are increasing as n increases. In fact, given any real number no matter how large it is there exist even number terms in this sequence which are greater than that real number. For example,

(i) If you consider the real number 10000 then there exists even natural number $m = 2k = 10002$ such that for all even natural numbers n

$$a_n > 10000 \;\; \forall \; n \geq m = 2k = 10002. \quad \begin{bmatrix} \because \text{ If } n \text{ is even then } (-1)^n n = 10000 \\ \Rightarrow n = 10000 \end{bmatrix}$$

(ii) If you consider the real number -2^{1000} then there exists odd natural number $m = 2k+1 = 2^{1000}+1$ such that for all odd natural numbers n

$$a_n < -2^{1000} \;\; \forall \;\; n \geq m = 2k+1 = 2^{1000}+1. \quad \begin{bmatrix} \because \text{ If } n \text{ is odd then } (-1)^n n = -2^{1000} \\ \Rightarrow -n = -2^{1000} \Rightarrow n = 2^{1000} \end{bmatrix}$$

If a sequence behaves like this then we say that sequence oscillates infinitely.

Also, note that $a_n = \begin{cases} n, & \text{if } n \text{ is even} \\ -n, & \text{if } n \text{ is odd} \end{cases}$

So,

$$\lim_{n\to\infty} a_n = \begin{cases} \lim\limits_{n\to\infty} n, & \text{if } n \text{ is even} \\ \lim\limits_{n\to\infty}(-n), & \text{if } n \text{ is odd} \end{cases}$$

$$= \begin{cases} \infty, & \text{if } n \text{ is even} \\ -\infty, & \text{if } n \text{ is odd} \end{cases}$$

With the help of graphical presentation and discussion of convergence, divergence and oscillation of these six sequences it will be easy for you to understand definition of these ideas which are given as follows.

Convergence of a Sequence: A real sequence $< a_n >$ is said to converges a real number 'a' if for a given $\varepsilon > 0$ however small it may be there exists a natural number m depending on ε such that $|a_n - a| < \varepsilon \ \forall \ n \geq m$ … (4.15)

The real number a is called the limit of the sequence $< a_n >$, and we write it as $a_n \to a$ as $n \to \infty$ or $\lim\limits_{n\to\infty} a_n = a$. … (4.16)

Condition (4.15) can be written as:

$$-\varepsilon < a_n - a < \varepsilon \ \forall \ n \geq m$$
$$\Rightarrow a - \varepsilon < a_n < a + \varepsilon \ \forall \ n \geq m$$

This means all the terms from m onward lie inside a strip of width ε about point a refer Fig. 4.2.

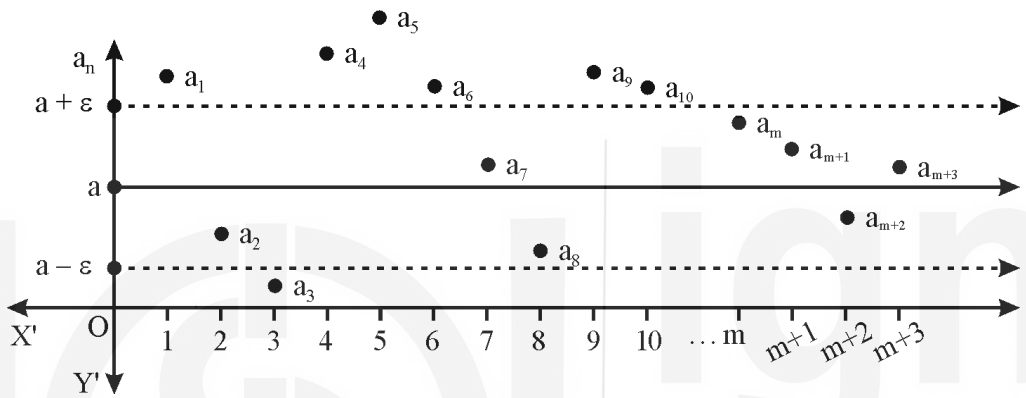


Fig. 4.2: Visualisation of the definition of convergence of a sequence

For example, sequence $< \frac{1}{n} >$ shown in Fig. 4.1 (a) converges to 0. Let us consider two values of $\varepsilon > 0$ so that you can understand well what this definition says.

- Suppose $\varepsilon = 0.1 > 0$ then any positive integer ≥ 11 can be taken as value of m. So,

  $$|a_n - 0| < 0.1 \ \forall \ n \geq 11 \text{ or } \left|\frac{1}{n} - 0\right| < 0.1 \ \forall \ n \geq 11 \ [\because \text{ here } a = 0]$$

  $$\text{or } \left|\frac{1}{n}\right| < 0.1 \ \forall \ n \geq 11 \left[\begin{array}{l} \because 11 > 10 \Rightarrow \frac{1}{11} < \frac{1}{10} \text{ refer (6.89)} \\ \Rightarrow \frac{1}{11} < 0.1 \\ \text{So, } \frac{1}{n} < 0.1 \ \forall \ n \geq 11 \Rightarrow \left|\frac{1}{n}\right| < 0.1 \ \forall \ n \geq 11 \end{array}\right]$$

- Suppose $\varepsilon = 0.01 > 0$ then any positive integer ≥ 101 can be taken as value of m. So, $|a_n - 0| < 0.01 \ \forall \ n \geq 101$ or $\left|\frac{1}{n} - 0\right| < 0.01 \ \forall \ n \geq 101 \ [\because \text{ here } a = 0]$

  $$\text{or } \left|\frac{1}{n}\right| < 0.01 \ \forall \ n \geq 101 \left[\begin{array}{l} \because 101 > 100 \Rightarrow \frac{1}{101} < \frac{1}{100} \Rightarrow \frac{1}{101} < 0.01 \\ \text{So, } \frac{1}{n} < 0.01 \ \forall \ n \geq 101 \Rightarrow \left|\frac{1}{n}\right| < 0.01 \ \forall \ n \geq 101 \end{array}\right]$$

Divergence of a Sequence: A real sequence $< a_n >$ is said to diverges to ∞ if given any positive real number K no matter how large it is there exists a positive integer m such that $a_n > K \ \forall \ n \geq m,$ and we write it as

$$a_n \to \infty \text{ as } n \to \infty \text{ or } \lim_{n\to\infty} a_n = \infty. \qquad \ldots (4.17)$$

A real sequence $< a_n >$ is said to diverges to $-\infty$ if given any positive real number K no matter how large it is there exists a positive integer m such that $a_n < -K \ \forall \ n \geq m,$ and we write it as

$$a_n \to -\infty \text{ as } n \to \infty \text{ or } \lim_{n\to\infty} a_n = -\infty. \qquad \ldots (4.18)$$

For example, sequence $< 2n-1 >$ shown in Fig. 4.1 (c) diverges to ∞. Let us consider two values of the positive real number K so that you can understand well what this definition says.

- Suppose $K = 10000$ then any positive integer ≥ 5001 can be taken as value of m. So,

$$a_n > 10000 \ \forall \ n \geq m = 5001 \text{ or } 2n-1 > 10000 \ \forall \ n \geq m = 5001$$
$$\left[\begin{array}{l} \because n > 5001 \Rightarrow 2n > 10002 \\ \Rightarrow 2n-1 > 10001 \Rightarrow 2n-1 > 10000 \Rightarrow a_n > 10000 \end{array}\right]$$

- Suppose $K = 2^{100000}$ then any positive integer ≥ $2^{99999} + 1$ can be taken as value of m. So,

$$a_n > 2^{100000} \ \forall \ n \geq m = 2^{99999} + 1 \text{ or } 2n-1 > 2^{100000} \ \forall \ n \geq m = 2^{99999} + 1$$
$$\left[\begin{array}{l} \because n > 2^{99999} + 1 \Rightarrow 2n > 2 \times 2^{99999} + 2 \Rightarrow 2n-1 > 2^{100000} + 1 \\ \Rightarrow 2n-1 > 2^{100000} \Rightarrow a_n > 2^{100000} \end{array}\right]$$

Following similar steps as in the case of divergence to ∞. You can consider some values of K and obtain corresponding value of m in the case $a_n \to -\infty$.

Oscillatory Sequence: If a real sequence $< a_n >$ neither converges to a finite real number and nor diverges to ∞ or $-\infty$, then we say that sequence oscillates. As discussed earlier oscillatory sequence may oscillate finitely or infinitely. … (4.19)

For example, sequence $< (-1)^n >$ shown in Fig. 4.1 (e) oscillates finitely because it converges to two finite real numbers 1 and – 1. The sequence $< (-1)^n n >$ shown in Fig. 4.1 (f) oscillates infinitely because it diverges to ∞ and $-\infty$.

Let us now do some examples to explain the idea of convergence and divergence of a sequence.

Example 1: Discuss the convergence of the sequence $< \frac{n+1}{n} >$ using definition.

Solution: Here $a_n = \frac{n+1}{n}$ and

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{n+1}{n} = \lim_{n\to\infty}\left(\frac{n}{n} + \frac{1}{n}\right) = \lim_{n\to\infty}\left(1 + \frac{1}{n}\right) = 1 + \lim_{n\to\infty}\left(\frac{1}{n}\right) = 1 + 0 = 1$$

Now, using definition we prove that it converges to 1. Let $\varepsilon > 0$ be given then

$$|a_n - 1| < \varepsilon \Rightarrow \left|\frac{n+1}{n} - 1\right| < \varepsilon \Rightarrow \left|1 + \frac{1}{n} - 1\right| < \varepsilon \Rightarrow \left|\frac{1}{n}\right| < \varepsilon \Rightarrow \frac{1}{n} < \varepsilon \left[\because n > 0, \text{ so } \left|\frac{1}{n}\right| = \frac{1}{n}\right]$$

$$\Rightarrow n > \frac{1}{\varepsilon} \left[\because a > 0,\ b > 0 \text{ and } a < b \Rightarrow \frac{1}{a} > \frac{1}{b} \text{ refer (6.89)}\right]$$

If we take positive integer $m > \frac{1}{\varepsilon}$, then we have $|a_n - 1| < \varepsilon \ \forall\ n \geq m$

Hence, by definition sequence $< \frac{n+1}{n} >$ converges to 1.

Example 2: Discuss the convergence of the sequence $< \frac{5n+4}{2n+3} >$ using definition.

Solution: Here $a_n = \frac{5n+4}{2n+3}$ and $\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{5n+4}{2n+3} = \lim_{n\to\infty} \left(\frac{5 + \frac{4}{n}}{2 + \frac{3}{n}}\right) = \left(\frac{5+0}{2+0}\right) = \frac{5}{2}$

Now, using definition we prove that it converges to 5/2. Let $\varepsilon > 0$ be given then

$$\left|a_n - \frac{5}{2}\right| < \varepsilon \Rightarrow \left|\frac{5n+4}{2n+3} - \frac{5}{2}\right| < \varepsilon \Rightarrow \left|\frac{10n+8-10n-15}{2(2n+3)}\right| < \varepsilon \Rightarrow \left|\frac{-7}{2(2n+3)}\right| < \varepsilon$$

$$\Rightarrow \left|\frac{7}{2(2n+3)}\right| < \varepsilon \Rightarrow \frac{7}{2(2n+3)} < \varepsilon \left[\begin{array}{l} \because n > 0, \text{ so } \frac{7}{2(2n+3)} > 0 \\ \Rightarrow \left|\frac{7}{2(2n+3)}\right| = \frac{7}{2(2n+3)} \end{array}\right]$$

Now, $\frac{7}{2(2n+3)} < \varepsilon \Rightarrow \frac{7}{2n+3} < 2\varepsilon \Rightarrow 7 < 4\varepsilon n + 6\varepsilon$

$$\Rightarrow 7 - 6\varepsilon < 4\varepsilon n \Rightarrow \frac{7-6\varepsilon}{4\varepsilon} < n \text{ or } n > \frac{7-6\varepsilon}{4\varepsilon}$$

If we take positive integer $m > \frac{7-6\varepsilon}{4\varepsilon}$, then we have $\left|a_n - \frac{5}{2}\right| < \varepsilon \ \forall\ n \geq m$

Hence, by definition sequence $< \frac{5n+4}{2n+3} >$ converges to $\frac{5}{2}$.

Remark 1: For better understanding of the concept, you should try experiments like: In particular, if $\varepsilon = 0.01$, then $m > \frac{7-0.06}{0.04} = \frac{6.94}{0.04} = 173.5$.

So, in this case we may take m = 174 or any value higher than 174 but not less than 173. **In other words, it means only first 173 terms of the given sequence are at a distance of 0.01 or higher from 5/2. All other terms from 174 onward are at a distance less than 0.01 from 5/2**.

Example 3: Show that the sequence $< 3^n >$ diverges using definition.

Solution: Here $a_n = 3^n$. Let K be any positive real number no matter how large it is. We claim that there exists positive integer m such that $a_n > K \ \ \forall\ n \geq m$

Now, $a_n > K \Rightarrow 3^n > K \Rightarrow n\log 3 > \log K \Rightarrow n > \dfrac{\log K}{\log 3}$

If we take positive integer $m > \dfrac{\log K}{\log 3}$, then we have $a_n > K \;\; \forall\, n \geq m$.

Hence, by definition sequence $< 3^n >$ diverges to ∞.

Remark 2: For better understanding of the concept, you should try experiments like: In particular, if $K = 10000$, then $m > \dfrac{\log 10000}{\log 3} \simeq 8.3836$. So, in this case we may take m = 9.

Now, you can try the following Self-Assessment Question.

SAQ 1

Write n^{th} term of the sequence $9, \dfrac{11}{4}, \dfrac{13}{9}, \dfrac{15}{16}, \dfrac{17}{25}, \ldots$

4.3 SOME RESULTS RELATED TO CONVERGENCE AND DIVERGENCE OF SEQUENCE AND THEIR USE

Before stating some results, which help in checking whether the given sequence converges or diverges or oscillates, let us explain some terms which will be required in the results which we are going to discuss here.

Bounded above Sequence: A sequence $< a_n >$ is said to be bounded above if there exists a real number K such that

$$a_n \leq K \;\; \forall\, n \in \mathbb{N} \qquad \ldots (4.20)$$

For example, consider the sequence $< \dfrac{1}{2^n} >$, this sequence is bounded above since $a_n \leq \dfrac{1}{2} \;\; \forall\, n \in \mathbb{N} \left[\because a_1 = \dfrac{1}{2} \leq \dfrac{1}{2},\, a_2 = \dfrac{1}{4} \leq \dfrac{1}{2},\, a_3 = \dfrac{1}{8} \leq \dfrac{1}{2},\text{ and so on}\right]$

Bounded below Sequence: A sequence $< a_n >$ is said to be bounded below if there exists a real number k such that

$$a_n \geq k \;\; \forall\, n \in \mathbb{N} \qquad \ldots (4.21)$$

For example, consider the sequence $< \dfrac{1}{2^n} >$, this sequence is bounded below since $a_n \geq 0 \;\; \forall\, n \in \mathbb{N} \left[\because a_1 = \dfrac{1}{2} \geq 0,\, a_2 = \dfrac{1}{4} \geq 0,\, a_3 = \dfrac{1}{8} \geq 0,\text{ and so on}\right]$

Bounded Sequence: A sequence $< a_n >$ is said to be bounded if it is both bounded above and bounded below. So, if a sequence is bounded then there will exist two real numbers k and K such that

$$k \leq a_n \leq K \;\; \forall\, n \in \mathbb{N} \qquad \ldots (4.22)$$

For example, consider the sequence $< \dfrac{1}{2^n} >$, this sequence is bounded since it is both bounded above and bounded below. Here k = 0 and $K = \dfrac{1}{2}$, i.e.,

$0 \le a_n \le \frac{1}{2} \ \forall \ n \in \mathbb{N} \left[\because a_1 = \frac{1}{2} \text{ and } 0 \le \frac{1}{2} \le \frac{1}{2}, a_2 = \frac{1}{4} \text{ and } 0 \le \frac{1}{4} \le \frac{1}{2}, \text{ and so on} \right]$

Least Upper Bound (lub) of a Sequence: Least upper bound of a sequence can be defined in two ways. Let us call them definition 1 and definition 2 given as follows.

Definition 1: Suppose sequence $< a_n >$ is bounded above then there will exists a real number u such that

$a_n \le u \ \forall \ n \in \mathbb{N}$

The real number u is known as **upper bound** of the sequence $< a_n >$. The upper bound u is said to be the least upper bound of the sequence $< a_n >$ if whenever we have any other upper bound u_1 then $u \le u_1$. That is no real number less than u can be an upper bound of the sequence $< a_n >$. … (4.23)

For example, $u = \frac{1}{2}$ is the least upper bound of the sequence $< \frac{1}{2^n} >$, since if we have any other upper bound u_1 then $\frac{1}{2} \le u_1$. e.g., $u_1 = 0.6$ is an upper bound of the sequence $< \frac{1}{2^n} >$, clearly $u = 0.5 < 0.6 = u_1$.

Definition 2: Let $< a_n >$ be a real and bounded above sequence then a real number u is said to be least upper bound (lub) of the sequence $< a_n >$ if

(i) $a_n \le u \ \forall \ n \in \mathbb{N}$ and

(ii) given any $\varepsilon > 0$, no matter how small it is, there exists at least one $n \in \mathbb{N}$ such that $a_n > u - \varepsilon$. … (4.24)

For example, $u = \frac{1}{2}$ is the least upper bound of the sequence $< \frac{1}{2^n} >$, since if we take any value of $\varepsilon > 0$, then there exists at least one value 1 of n such that

$a_1 = \frac{1}{2} > \frac{1}{2} - \varepsilon$

Remark 3: (i) In particular, if $\varepsilon = 0.4 > 0$, then there are only three values 1, 2 and 3 of n such that

$a_1 = \frac{1}{2} = 0.5 > 0.1 = \frac{1}{2} - 0.4 = \frac{1}{2} - \varepsilon$, $a_2 = \frac{1}{4} = 0.25 > 0.1 = \frac{1}{2} - 0.4 = \frac{1}{2} - \varepsilon$,

$a_3 = \frac{1}{8} = 0.125 > 0.1 = \frac{1}{2} - 0.4 = \frac{1}{2} - \varepsilon$.

But $a_4 = \frac{1}{16} = 0.0625 < 0.1 = \frac{1}{2} - 0.4 = \frac{1}{2} - \varepsilon$

(ii) If $\varepsilon = 0.3 > 0$, then there are only two values 1 and 2 of n such that

$a_1 = \frac{1}{2} = 0.5 > 0.2 = \frac{1}{2} - 0.3 = \frac{1}{2} - \varepsilon$, $a_2 = \frac{1}{4} = 0.25 > 0.2 = \frac{1}{2} - 0.3 = \frac{1}{2} - \varepsilon$.

But $a_3 = \frac{1}{8} = 0.125 < 0.2 = \frac{1}{2} - 0.3 = \frac{1}{2} - \varepsilon$.

(iii) If $\varepsilon = 0.2 > 0$, then there is only one value 1 of n such that

$$a_1 = \frac{1}{2} = 0.5 > 0.3 = \frac{1}{2} - 0.2 = \frac{1}{2} - \varepsilon.$$

But $a_2 = \frac{1}{4} = 0.25 < 0.3 = \frac{1}{2} - 0.2 = \frac{1}{2} - \varepsilon.$

If you take any value of $\varepsilon > 0,$ no matter how small it is then you will always find that $a_1 = \frac{1}{2} > \frac{1}{2} - \varepsilon.$ Hence, $\frac{1}{2}$ is the lub of the sequence $< \frac{1}{2^n} >$.

Greatest Lower Bound (glb) of a Sequence: Greatest lower bound of a sequence can be defined in two ways. Let us call them definition 1 and definition 2 given as follows.

Definition 1: Suppose sequence $< a_n >$ is bounded below then there will exists a real number g such that

$a_n \geq g \;\; \forall \; n \in \mathbb{N}.$

A real number g which satisfy this condition is known as **lower bound** of the sequence $< a_n >$. The lower bound g is said to be the greatest lower bound of the sequence $< a_n >$ if whenever we have any other lower bound g_1 then $g_1 \leq g.$ That is no real number > g can be a lower bound of the sequence $< a_n >$. … (4.25)

For example, $g = 0$ is the greatest lower bound of the sequence $< \frac{1}{2^n} >,$ since if we have any other lower bound g_1 then $g_1 \leq g.$ e.g., $g_1 = -0.1$ is a lower bound of the sequence $< \frac{1}{2^n} >,$ clearly $g_1 = -0.1 < 0 = g.$

Definition 2: Let $< a_n >$ be a real and bounded below sequence then a real number g is said to be greatest lower bound (glb) of the sequence $< a_n >$ if

(i) $a_n \geq g \;\; \forall \; n \in \mathbb{N}$ and

(ii) given any $\varepsilon > 0,$ no matter how small it is, there exists at least one $n \in \mathbb{N}$ such that $a_n < g + \varepsilon.$ … (4.26)

For example, $g = 0$ is the greatest lower bound of the sequence $< \frac{1}{2^n} >,$ since if we take any value of $\varepsilon > 0,$ then there exists at least one value of n such that $a_n < 0 + \varepsilon.$

Remark 4: (i) In particular, if $\varepsilon = 0.1 > 0,$ then except first three terms all other terms are ≤ 0 + 0.1 = g + $\varepsilon.$ That is

$$a_4 = \frac{1}{16} = 0.0625 < 0.1 = 0 + 0.1 = g + \varepsilon,$$

$$a_5 = \frac{1}{32} = 0.03125 < 0.1 = 0 + 0.1 = g + \varepsilon, \text{ and so on.}$$

But $a_3 = \frac{1}{8} = 0.125 > 0.1 = 0 + 0.1 = g + \varepsilon$

(ii) If $\varepsilon = 0.2 > 0,$ then except first two terms all other terms are ≤ 0 + 0.1 = g + $\varepsilon.$ That is

$a_3 = \frac{1}{8} = 0.125 < 0.2 = 0 + 0.2 = g + \varepsilon,$

$a_4 = \frac{1}{16} = 0.0625 < 0.2 = 0 + 0.2 = g + \varepsilon,$ and so on.

But $a_2 = \frac{1}{4} = 0.25 > 0.2 = 0 + 0.2 = g + \varepsilon.$

(iii) If $\varepsilon = 0.3 > 0,$ then except first term all other terms are ≤ 0 + 0.1 = g + ε. That is

$a_2 = \frac{1}{4} = 0.25 < 0.3 = 0 + 0.3 = g + \varepsilon,$

$a_3 = \frac{1}{8} = 0.125 < 0.3 = 0 + 0.3 = g + \varepsilon,$ and so on.

But $a_1 = \frac{1}{2} = 0.5 > 0.3 = 0 + 0.3 = g + \varepsilon.$

For the sequence $< \frac{1}{2^n} >,$ if you will take any value of $\varepsilon > 0,$ which is greater than 0.5 then all terms of the given sequence will be < $g + \varepsilon.$ But if you take any value of $\varepsilon > 0,$ no matter how small it is then you will always find at least one value of n such that $a_n < g + \varepsilon.$ Hence, 0 is the glb of the sequence $< \frac{1}{2^n} >.$

Monotonic Sequence

Before defining monotonic sequence, we have to define monotonically increasing and decreasing sequences.

Monotonically Increasing Sequence: A real sequence $< a_n >$ is said to be monotonically increasing sequence if $a_n \le a_{n+1} \;\; \forall\, n \in \mathbb{N}$

i.e., $a_1 \le a_2 \le a_3 \le a_4 \le a_5 \ldots$ … (4.27)

For example, consider the sequence $< 2^n >,$ this sequence is monotonically increasing since

$a_n \le a_{n+1} \;\; \forall\, n \in \mathbb{N} \left[\because \forall\, n \in \mathbb{N} \Rightarrow n < n+1 \Rightarrow 2^n < 2^{n+1} \Rightarrow a_n < a_{n+1} \Rightarrow a_n \le a_{n+1}\right]$

Monotonically Decreasing Sequence: A real sequence $< a_n >$ is said to be monotonically decreasing sequence if $a_n \ge a_{n+1} \;\; \forall\, n \in \mathbb{N}$

i.e., $a_1 \ge a_2 \ge a_3 \ge a_4 \ge a_5 \ldots$ … (4.28)

For example, consider the sequence $< \frac{1}{2^n} >,$ this sequence is monotonically decreasing since

$$a_n \ge a_{n+1} \;\; \forall\, n \in \mathbb{N} \left[\begin{array}{l} \because \forall\, n \in \mathbb{N} \Rightarrow n < n+1 \Rightarrow 2^n < 2^{n+1} \Rightarrow \frac{1}{2^n} > \frac{1}{2^{n+1}} \text{ refer (6.89)} \\ \Rightarrow a_n > a_{n+1} \Rightarrow a_n \ge a_{n+1} \end{array}\right]$$

Monotonic Sequence: A real sequence $< a_n >$ is said to be monotonic if either it is monotonically increasing or monotonically decreasing sequence.

… (4.29)

Cauchy Sequence: A sequence $< a_n >$ is said to be a Cauchy sequence if given $\varepsilon > 0$, no matter how small it is there exists a positive integer n_0 such that $|a_n - a_m| < \varepsilon \ \forall \ m, n \geq n_0$ … (4.30)

Now, let us state some results without proof as proofs are beyond the scope of this course.

Result 1: If a sequence $< a_n >$ converges then it will converge to a unique limit. … (4.31)

Result 2: A monotonic sequence $< a_n >$ is convergent if and only if it is bounded. … (4.32)

Result 3: A sequence $< a_n >$ which is both monotonically increasing and bounded above converges to its lub. … (4.33)

Result 4: A sequence $< a_n >$ which is both monotonically decreasing and bounded below converges to its glb. … (4.34)

Result 5: A sequence $< a_n >$ which is monotonically increasing but not bounded above diverges to ∞. … (4.35)

Result 6: A sequence $< a_n >$ which is monotonically decreasing but not bounded below diverges to $-\infty$. … (4.36)

Remark 5: In view of results 2 to 6 you can say that a monotonic sequence either converges or diverges. … (4.37)

Result 7: Every convergent sequence is bounded. … (4.38)

Result 8: If a sequence $< a_n >$ converges to a and $a_n \geq 0 \ \forall \ n \in \mathbb{N}$, then $a \geq 0$. … (4.39)

Result 9: If sequences $< a_n >$ and $< b_n >$ are such that $a_n \leq b_n \ \forall \ n \in \mathbb{N}$ and $a_n \to a$ and $b_n \to b$ then $a \leq b$. … (4.40)

Result 10: **Squeeze Principle**: If sequences $< a_n >$, $< b_n >$ and $< c_n >$ are such that $a_n \leq b_n \leq c_n \ \forall \ n \geq n_0$ for some positive integer n_0 and $a_n \to a$ and $c_n \to a$ then $b_n \to a$. … (4.41)

Result 11: **Cauchy's First Theorem on Limit**: If sequence $< a_n >$ converges to 'a' then the sequence $< b_n >$ also converges to 'a' where

$b_n = \dfrac{a_1 + a_2 + a_3 + \ldots + a_n}{n} \ \forall \ n \in \mathbb{N}$, i.e., b_n is arithmetic mean of the first n terms of the sequence $< a_n >$. … (4.42)

Result 12: **Cauchy's Second Theorem on Limit**: If all the terms of the sequence $< a_n >$ are positive then $\lim\limits_{n\to\infty}(a_n)^{1/n} = \lim\limits_{n\to\infty}\dfrac{a_{n+1}}{a_n}$ provided $\lim\limits_{n\to\infty}\dfrac{a_{n+1}}{a_n}$ exists as finite or infinite. … (4.43)

Result 13: A real sequence $< a_n >$ is convergent if and only if it is a Cauchy sequence. … (4.44)

Result 14: **Algebra of Limits**: If sequences $< a_n >$ and $< b_n >$ converge to a and b respectively. i.e., $\lim\limits_{n\to\infty} a_n = a$ and $\lim\limits_{n\to\infty} b_n = b$, then following hold:

- $\lim_{n\to\infty}(a_n + b_n) = a + b$ … (4.45)
- $\lim_{n\to\infty}(a_n - b_n) = a - b$ … (4.46)
- $\lim_{n\to\infty} a_n b_n = ab$ … (4.47)
- $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b},\ b \neq 0$ … (4.48)
- $\lim_{n\to\infty} \frac{1}{a_n} = \frac{1}{a},\ a \neq 0$ … (4.49)
- $\lim_{n\to\infty} |a_n| = |a|$ … (4.50)
- $\lim_{n\to\infty} \alpha a_n = \alpha a,\ \alpha$ is any real number … (4.51)

Now, let us apply these results in solving problems on convergence of sequence.

Example 4: Discuss the convergence of the sequence $<\frac{1}{n}>$ which is shown in Fig. 4.1 (a) and given by (4.3) or (4.4).

Solution: Here $a_n = \frac{1}{n}$, so $a_{n+1} = \frac{1}{n+1}$.

Now $n < n+1 \Rightarrow \frac{1}{n} > \frac{1}{n+1} \Rightarrow a_n > a_{n+1} \Rightarrow < a_n >$ is decreasing. … (4.52)

Another way of showing the sequence $< a_n >$ as a decreasing sequence is:

$$a_n - a_{n+1} = \frac{1}{n} - \frac{1}{n+1} = \frac{n+1-n}{n(n+1)} = \frac{1}{n(n+1)} > 0 \ \left[\because n > 0 \Rightarrow n(n+1) > 0\right]$$

$\Rightarrow a_n - a_{n+1} > 0 \Rightarrow a_n > a_{n+1}$. Hence, $< a_n >$ is decreasing.

You can prove it decreasing using either way you like.

Also, $a_n = \frac{1}{n} > 0\ \forall\ n \in \mathbb{N} \Rightarrow$ Sequence $< a_n >$ is bounded below. … (4.53)

By (4.52) and (4.53) the given sequence $< a_n >$ is both bounded below as well as decreasing. Hence, using (4.34) given sequence $<\frac{1}{n}>$ converges to its glb where $\text{glb} = \lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{1}{n} = 0$. Hence, $<\frac{1}{n}> \to 0$.

Example 5: Discuss the convergence of the sequence $<\frac{n}{n+1}>$.

Solution: Here $a_n = \frac{n}{n+1}$, so $a_{n+1} = \frac{n+1}{n+2}$.

We claim that given sequence $< a_n >$ is an increasing sequence.

$$a_n - a_{n+1} = \frac{n}{n+1} - \frac{n+1}{n+2} = \frac{n^2 + 2n - (n^2 + 2n + 1)}{(n+1)(n+2)}$$

$$= \frac{-1}{(n+1)(n+2)} < 0 \quad [\because n > 0 \Rightarrow (n+1)(n+2) > 0]$$

$\Rightarrow a_n - a_{n+1} < 0 \Rightarrow a_n < a_{n+1}$. Hence, $< a_n >$ is increasing. … (4.54)

Also, $n < n+1 \Rightarrow \frac{n}{n+1} < 1 \; \forall \; n \in \mathbb{N} \Rightarrow a_n < 1 \;\; \forall \; n \in \mathbb{N}$

$\Rightarrow$ Sequence $< a_n >$ is bounded above. … (4.55)

By (4.54) and (4.55) the given sequence $< a_n >$ is both bounded above as well as increasing. Hence, using (4.33) given sequence $< \frac{n}{n+1} >$ converges to its lub where $\text{lub} = \lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{n}{n+1} = \lim_{n\to\infty} \frac{1}{1+\frac{1}{n}} = \frac{1}{1+0} = 1$. Hence, $< \frac{n}{n+1} > \to 1$.

Example 6: Discuss the convergence of the sequence $< n^{1/n} >$.

Solution: Let $a_n = n$, so $a_{n+1} = n+1$

Now, $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{n+1}{n} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = 1 + 0 = 1$ … (4.56)

By Cauchy's second theorem on limit, we have

$\lim_{n\to\infty} (a_n)^{1/n} = \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = 1 \qquad [\text{Using } (4.56)]$

Hence, the sequence $< n^{1/n} >$ converges to 1. … (4.57)

First 35 terms of this sequence are shown in Table 4.1 and have been visualised in Fig. 4.3 (a) by 35 vertical lines. From Table 4.1 note that first three terms are increasing as n increases but after that continuously terms are decreasing as n increases. So, this trend verifies the same thing that we have shown in Example 6 that $a_n \to 1$ as $n \to \infty$.

Example 7: Discuss the convergence of the sequence

$< \frac{1 + 2^{1/2} + 3^{1/3} + \ldots + n^{1/n}}{n} >$.

Solution: Let $a_n = n^{1/n}$, then using (4.57) it converges to 1.

By Cauchy's first theorem on limit, we have

$$\lim_{n\to\infty} \frac{a_1 + a_2 + a_3 + \ldots + a_n}{n} = \lim_{n\to\infty} a_n = 1$$

$$\Rightarrow \lim_{n\to\infty} \frac{1 + 2^{1/2} + 3^{1/3} + \ldots + n^{1/n}}{n} = 1.$$

Hence, the given sequence $< \frac{1 + 2^{1/2} + 3^{1/3} + \ldots + n^{1/n}}{n} >$ converges to 1.

First 35 terms of this sequence are shown in Table 4.2 and have been visualised in Fig. 4.3 (b) by 35 vertical lines. From Table 4.2 note that first six terms are increasing as n increases but after that continuously terms are decreasing as n increases. So, this trend verifies the same thing that we have shown in Example 7 that $a_n \to 1$ as $n \to \infty$.

Table 4.1: First 35 terms of the sequence discussed in Example 6. Values of these terms are obtained using R software

Value of n	n^{th} term of the sequence	Value of n	n^{th} term of the sequence	Value of n	n^{th} term of the sequence
1	1.000000	13	1.218114	25	1.137411
2	1.414214	14	1.207442	26	1.133501
3	1.442250	15	1.197860	27	1.129831
4	1.414214	16	1.189207	28	1.126378
5	1.379730	17	1.181352	29	1.123124
6	1.348006	18	1.174187	30	1.120050
7	1.320469	19	1.167623	31	1.117142
8	1.296840	20	1.161586	32	1.114387
9	1.276518	21	1.156013	33	1.111772
10	1.258925	22	1.150851	34	1.109286
11	1.243575	23	1.146055	35	1.106920
12	1.230076	24	1.141586		

Table 4.2: First 35 terms of the sequence discussed in Example 7. Values of these terms are obtained using R software

Value of n	n^{th} term of the sequence	Value of n	n^{th} term of the sequence	Value of n	n^{th} term of the sequence
1	1.000000	13	1.295610	25	1.234164
2	1.207107	14	1.289312	26	1.230293
3	1.285488	15	1.283215	27	1.226572
4	1.317669	16	1.277340	28	1.222993
5	1.330081	17	1.271694	29	1.219550
6	1.333069	18	1.266277	30	1.216233
7	1.331269	19	1.261084	31	1.213037
8	1.326965	20	1.256109	32	1.209954
9	1.321360	21	1.251343	33	1.206978
10	1.315116	22	1.246775	34	1.204105
11	1.308613	23	1.242396	35	1.201328
12	1.302068	24	1.238196		

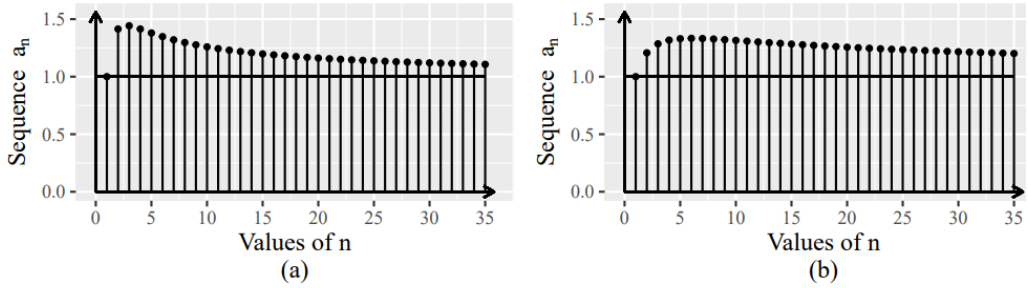


Fig. 4.3: Visualisation of successive terms of the sequence discussed in Examples (a) 6 (b) 7

Example 8: Prove that the sequence $<\frac{1}{n}>$ is a Cauchy sequence.

Solution: Here $a_n = \frac{1}{n}$. Let $\varepsilon > 0$ be given. Let n > m then

$$|a_n - a_m| = \left|\frac{1}{n} - \frac{1}{m}\right| = \left|-\left(\frac{1}{m} - \frac{1}{n}\right)\right|$$

$$= \left|\frac{1}{m} - \frac{1}{n}\right| = \frac{1}{m} - \frac{1}{n} \quad \left[\begin{array}{l} \because n > m \Rightarrow \frac{1}{n} < \frac{1}{m} \Rightarrow \frac{1}{m} - \frac{1}{n} > 0 \\ \Rightarrow \left|\frac{1}{m} - \frac{1}{n}\right| = \frac{1}{m} - \frac{1}{n} \text{ as if } x \geq 0 \text{ then } |x| = x \end{array}\right]$$

$$\Rightarrow |a_n - a_m| = \frac{1}{m} - \frac{1}{n} < \frac{1}{m} \quad \left[\because n > 0 \Rightarrow \frac{1}{n} > 0\right]$$

So, $|a_n - a_m| < \varepsilon \Rightarrow \frac{1}{m} < \varepsilon \Rightarrow m > \frac{1}{\varepsilon}$

If we take a positive integer $n_0 > \frac{1}{\varepsilon}$ then

$\Rightarrow |a_n - a_m| < \varepsilon \quad \forall\ m, n \geq n_0$

Hence, the given sequence $<\frac{1}{n}>$ is a Cauchy sequence.

Remark 6: In particular, if $\varepsilon = 0.1 > 0$, then $n_0 > \frac{1}{\varepsilon} \Rightarrow n_0 > \frac{1}{0.1} \Rightarrow n_0 > 10$. It means for $\varepsilon = 0.1 > 0$, difference between any two terms after tenth term is less than 0.1. if $\varepsilon = 0.01 > 0$, then $n_0 > \frac{1}{\varepsilon} \Rightarrow n_0 > \frac{1}{0.01} \Rightarrow n_0 > 100$. It means for $\varepsilon = 0.01 > 0$, difference between any two terms after 100th term is less than 0.01.

Example 9: Prove that the sequence $<\frac{1}{n^2} + \frac{1}{(n+1)^2} + \frac{1}{(n+2)^2} + \ldots + \frac{1}{(2n)^2}>$ converges to 0.

Solution: Here $a_n = \frac{1}{n^2} + \frac{1}{(n+1)^2} + \frac{1}{(n+2)^2} + \ldots + \frac{1}{(2n)^2}$, then

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

$$< \frac{1}{n^2} + \frac{1}{n^2} + \frac{1}{n^2} + \ldots + \frac{1}{n^2} \quad \left[\because \frac{1}{(n+1)^2} < \frac{1}{n^2}, \text{ etc.}\right]$$

$$= \frac{n+1}{n^2} \quad \left[\because \underbrace{a + a + a + \ldots + a}_{(n+1) \text{ times}} = (n+1)a, \text{ here } a = \frac{1}{n^2}\right]$$

Also,

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

$$> \frac{1}{(2n)^2} + \frac{1}{(2n)^2} + \frac{1}{(2n)^2} + \ldots + \frac{1}{(2n)^2} \left[\because \frac{1}{n^2} > \frac{1}{(2n)^2}, \text{ etc.}\right]$$

$$= \frac{n+1}{(2n)^2} \qquad \left[\because \underbrace{a + a + a + \ldots + a}_{(n+1)\text{ times}} = (n+1)a, \text{ here } a = \frac{1}{(2n)^2}\right]$$

Thus, we have

$$\frac{n+1}{(2n)^2} < a_n < \frac{n+1}{n^2} \qquad \ldots (4.58)$$

But $\lim\limits_{n\to\infty} \frac{n+1}{(2n)^2} = \lim\limits_{n\to\infty} \frac{n+1}{4n^2} = \lim\limits_{n\to\infty} \frac{1+\frac{1}{n}}{4n} = 0$ and $\lim\limits_{n\to\infty} \frac{n+1}{n^2} = \lim\limits_{n\to\infty} \left(\frac{1}{n} + \frac{1}{n^2}\right) = 0 \ldots (4.59)$

Taking limit in (4.58), we have

$$\lim_{n\to\infty} \frac{n+1}{(2n)^2} < \lim_{n\to\infty} a_n < \lim_{n\to\infty} \frac{n+1}{n^2}$$

$$\Rightarrow 0 < \lim_{n\to\infty} a_n < 0 \qquad [\text{Using } (4.59)]$$

$$\Rightarrow \lim_{n\to\infty} a_n = 0 \qquad [\text{Using Squeeze principle refer } (4.41)]$$

$$\therefore \lim_{n\to\infty} a_n = 0 \Rightarrow \frac{1}{n^2} + \frac{1}{(n+1)^2} + \frac{1}{(n+2)^2} + \ldots + \frac{1}{(2n)^2} \to 0 \text{ as } n \to \infty$$

Hence, given sequence $< \frac{1}{n^2} + \frac{1}{(n+1)^2} + \frac{1}{(n+2)^2} + \ldots + \frac{1}{(2n)^2} >$ converges to 0.

Now, you can try the following Self-Assessment Question.

SAQ 2

If n^{th} term of a sequence $< a_n >$ is $< \frac{1 + 2^{1/2} + 3^{1/3} + \ldots + n^{1/n}}{n} >$, then find a_3 and a_{n+1}.

4.4 INFINITE SERIES

In Sec. 4.3 you have studied about sequence and discussed their converges. In this section you will study about infinite series of **positive terms**.

You know that a sequence $< a_n >$ can also be written as $a_1, a_2, a_3, a_4, \ldots$

If the terms of this sequence are added instead of separated by commas, then resulting expression will be $a_1 + a_2 + a_3 + a_4 + \ldots$ … (4.60)

In mathematics the expression written in (4.60) can be written using summation notation as follows

$$\sum_{n=1}^{\infty} a_n \qquad \ldots (4.61)$$

Both expressions given by (4.60) and (4.61) are known as **infinite series**.

You know that if you add some finite real numbers then result is a finite real number. For example, (i) $5+10+15+20=50$ which is a finite number.

(ii) $1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}=\frac{16+8+4+2+1}{16}=\frac{31}{16}$ which is a finite number. What about if you add infinite real numbers. Answer of this question is: it may be infinite or finite depends on the numbers in our series. Throughout this section we will trying to answer this question through different examples. But before going to discuss examples we have to define some terms which are necessary for further discussion.

Sequence of Partial Sum: If we have an infinite series $\sum_{n=1}^{\infty} a_n$ then the sequence $s_1, s_2, s_3, s_4, \ldots, s_n, \ldots$, where

$s_1 = a_1,\ s_2 = a_1 + a_2,\ s_3 = a_1 + a_2 + a_3,\ s_4 = a_1 + a_2 + a_3 + a_4, \ldots,$

$s_n = a_1 + a_2 + a_3 + \ldots + a_n, \ldots$

is known as sequence of partial sum of the given infinite series $\sum_{n=1}^{\infty} a_n$.

Convergence and Divergence of Infinite Series: If the sum of infinite terms

$$a_1 + a_2 + a_3 + \ldots + a_n + \ldots \qquad \ldots (4.62)$$

is a real number then we say that infinite series $\sum_{n=1}^{\infty} a_n$ converges, otherwise diverges. But we will not obtain sum of all terms to check its convergence. We will test convergence of series using some results. One easy way of testing whether given series converges or not is discussed as follows.

Note that $a_1 + a_2 + a_3 + \ldots + a_n + \ldots = \lim_{n\to\infty}(a_1 + a_2 + a_3 + \ldots + a_n) = \lim_{n\to\infty} s_n \ldots (4.63)$

From (4.63), we note that convergence or divergence of infinite series $\sum_{n=1}^{\infty} a_n$ depends on convergence or divergence of the sequence $< s_n >$ of partial sum. So, one of the testing procedures to test convergence of a given series is via testing convergence of the sequence $< s_n >$ of partial sum.

Now, let us state some results without proof as proofs are beyond the scope of this course which will be used to test whether a given series converges or diverges.

Result 15: If an infinite series $\sum_{n=1}^{\infty} a_n$ converges then $\lim_{n\to\infty} a_n = 0$. $\ldots$ (4.64)

Result 16: **Geometric Series**: Consider the infinite geometric series

$$\sum_{n=1}^{\infty} r^{n-1} = 1 + r + r^2 + r^3 + \ldots$$

It converges if $0 < r < 1$ and diverges if $r \geq 1$. $\ldots$ (4.65)

Result 17: **p-series**: An infinite series of the form

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = \frac{1}{1^p} + \frac{1}{2^p} + \frac{1}{3^p} + \frac{1}{4^p} + \ldots \qquad \ldots (4.66)$$

is known as p-series. It converges if $p > 1$ and diverges if $p \leq 1$. $\ldots$ (4.67)

Result 18: **Comparison Test**: If $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ are two positive terms series

and:

- $a_n \le k\, b_n \;\; \forall\, n \in \mathbb{N}$

  then (i) convergence of $\sum_{n=1}^{\infty} b_n$ implies $\sum_{n=1}^{\infty} a_n$ converges. … (4.68)

  (ii) divergence of $\sum_{n=1}^{\infty} a_n$ implies $\sum_{n=1}^{\infty} b_n$ diverges. … (4.69)

- $\lim_{n\to\infty} \frac{a_n}{b_n} = l,$ where l is a non-zero finite real number … (4.70)

  then $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ converges or diverges together. … (4.71)

  if l is 0 and $\sum_{n=1}^{\infty} b_n$ converges implies $\sum_{n=1}^{\infty} a_n$ converges. … (4.72)

  if l is ∞ and $\sum_{n=1}^{\infty} b_n$ diverges implies $\sum_{n=1}^{\infty} a_n$ diverges. … (4.73)

Result 19: **D' Alembert's Ratio Test**: If $\sum_{n=1}^{\infty} a_n$ is a series of positive terms and

$\lim_{n\to\infty} \frac{a_n}{a_{n+1}} = l,$ then

(i) $\sum_{n=1}^{\infty} a_n$ converges if l > 1 … (4.74)

(ii) $\sum_{n=1}^{\infty} a_n$ diverges if l < 1. … (4.75)

(iii) Test fails if l = 1 … (4.76)

(iv) $\sum_{n=1}^{\infty} a_n$ converges if $l = \infty$. … (4.77)

Result 20: **Cauchy's Root Test**: If $\sum_{n=1}^{\infty} a_n$ is a series of positive terms and

$\lim_{n\to\infty} (a_n)^{1/n} = l,$ then

(i) $\sum_{n=1}^{\infty} a_n$ converges if l < 1 … (4.78)

(ii) $\sum_{n=1}^{\infty} a_n$ diverges if l > 1. … (4.79)

(iii) Test fails if l = 1 … (4.80)

(iv) $\sum_{n=1}^{\infty} a_n$ diverges if $l = \infty$. … (4.81)

Result 21: Convergence or divergence of an infinite series $\sum_{n=1}^{\infty} a_n$ remains unaltered on omission of some finite terms. … (4.82)

Result 22: If a series $\sum_{n=1}^{\infty} a_n$ converges to a then for a real number c series $\sum_{n=1}^{\infty} c\, a_n$ also converges and converges to ca. … (4.83)

Result 23: If a series $\sum_{n=1}^{\infty} a_n$ diverges then for a non-zero real number c series $\sum_{n=1}^{\infty} c\, a_n$ also diverges. … (4.84)

Result 24: If series $\sum_{n=1}^{\infty} a_n$ converges to a and series $\sum_{n=1}^{\infty} b_n$ converges to b then the series $\sum_{n=1}^{\infty} (a_n + b_n)$ also converges and converges to a + b. … (4.85)

Result 25: If series $\sum_{n=1}^{\infty} a_n$ converges to a and series $\sum_{n=1}^{\infty} b_n$ converges to b then the series $\sum_{n=1}^{\infty} (a_n - b_n)$ also converges and converges to a – b. … (4.86)

Now, let us test convergence of some series using these results.

Example 10: Test the convergence of the series $\sum_{n=1}^{\infty} \sqrt{\frac{n+1}{n}}$.

Solution: Here $a_n = \sqrt{\frac{n+1}{n}}$.

Now, $\lim_{n\to\infty} a_n = \lim_{n\to\infty} \sqrt{\frac{n+1}{n}} = \lim_{n\to\infty} \sqrt{1+\frac{1}{n}} = \sqrt{1+0} = 1 \neq 0$

Hence, the given series $\sum_{n=1}^{\infty} \sqrt{\frac{n+1}{n}}$ does not converges. [Using (4.64)]

Example 11: Test the convergence of the series $\sum_{n=1}^{\infty} \frac{1}{n^n}$.

Solution: Here given series is $\sum_{n=1}^{\infty} \frac{1}{n^n}$. After omitting first two terms of the given series then it reduces to $\sum_{n=3}^{\infty} \frac{1}{n^n}$.

So, $\sum_{n=3}^{\infty} \frac{1}{n^n} = \frac{1}{3^3} + \frac{1}{4^4} + \frac{1}{5^5} + \frac{1}{6^6} + \ldots < \frac{1}{3^3} + \frac{1}{3^4} + \frac{1}{3^5} + \frac{1}{3^6} + \ldots \left[\begin{array}{l} \because 3 < 4 \Rightarrow 3^4 < 4^4 \\ \Rightarrow \frac{1}{4^4} < \frac{1}{3^4}, \text{ etc.} \end{array}\right]$

But $\frac{1}{3^3} + \frac{1}{3^4} + \frac{1}{3^5} + \frac{1}{3^6} + \ldots$ is an infinite geometric series with common ratio $\frac{1}{3} < 1$. Therefore it converges. [refer (4.65)]

But then using comparison test given series converges. [refer (4.68)]

Example 12: Test the convergence of the series $\frac{1}{5.9} + \frac{1}{6.11} + \frac{1}{7.13} + \frac{1}{8.15} + \ldots$

Solution: Here

$$a_n = \frac{1}{(n^{th} \text{ term of } 5, 6, 7, 8, \ldots)(n^{th} \text{ term of } 9, 11, 13, 15, \ldots)}$$

$$= \frac{1}{(n+4)(2n+7)} \quad \left[\begin{array}{l} \because n^{th} \text{ term of } 5, 6, 7, 8, \ldots \text{ which is an AP is} \\ a+(n-1)d = 5+(n-1)(1) = n+4 \\ \text{Similarly, } n^{th} \text{ term of } 9, 11, 13, 15, \ldots \\ \text{which is also an AP is } 9+(n-1)(2) = 2n+7 \end{array}\right]$$

$$= \frac{1}{n^2(1+4/n)(2+7/n)}$$

Let $b_n = \frac{1}{n^2}$

Now,

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{\frac{1}{n^2(1+4/n)(2+7/n)}}{1/n^2} = \lim_{n\to\infty} \frac{1}{(1+4/n)(2+7/n)} = \frac{1}{(1+0)(2+0)} = \frac{1}{2}$$

which is non-zero finite number. Thus, by comparison test series $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ converges or diverges together. But series $\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} \frac{1}{n^2}$ converges by p-test as p = 2 > 1. [refer (4.67)]. Hence, given series converges.

Example 13: Test the convergence of the series $\frac{1}{5} + \frac{1}{2.5^2} + \frac{1}{3.5^3} + \frac{1}{4.5^4} + \ldots$

Solution: Here

$$\sum_{n=1}^{\infty} a_n = \frac{1}{5} + \frac{1}{2.5^2} + \frac{1}{3.5^3} + \frac{1}{4.5^4} + \ldots = \frac{1}{1.5} + \frac{1}{2.5^2} + \frac{1}{3.5^3} + \frac{1}{4.5^4} + \ldots$$

$$a_n = \frac{1}{(n^{th} \text{ term of } 1, 2, 3, 4, \ldots)(n^{th} \text{ term of } 5, 5^2, 5^3, 5^4, \ldots)}$$

$$= \frac{1}{n.5^n} \quad \left[\begin{array}{l} \because n^{th} \text{ term of } 1, 2, 3, 4, \ldots \text{ which is an AP is} \\ a+(n-1)d = 1+(n-1)(1) = n \\ \text{Similarly, } n^{th} \text{ term of } 5, 5^2, 5^3, 5^4, \ldots \text{ which is a GP is} \\ ar^{n-1} = 5(5)^{n-1} = 5^n \end{array}\right]$$

$$\therefore a_{n+1} = \frac{1}{(n+1)5^{n+1}}$$

Now,

$$\lim_{n\to\infty} \frac{a_n}{a_{n+1}} = \lim_{n\to\infty} \frac{\frac{1}{n5^n}}{\frac{1}{(n+1)5^{n+1}}} = \lim_{n\to\infty} \frac{5(n+1)}{n} = \lim_{n\to\infty} 5\left(\frac{n+1}{n}\right) = \lim_{n\to\infty} 5\left(1+\frac{1}{n}\right) = 5(1+0) = 5$$

Since 5 > 1, so by D Alembert Ratio test given series converges. [refer (4.74)]

Example 14: Test the convergence of the series $\frac{x}{2.3} + \frac{x^2}{5.8} + \frac{x^3}{8.13} + \frac{x^4}{11.18} + \ldots$ where $x > 0$.

Solution: Here $a_n = \frac{n^{th} \text{ term of } x, x^2, x^3, x^4, \ldots}{(n^{th} \text{ term of } 2, 5, 8, 11, \ldots)(n^{th} \text{ term of } 3, 8, 13, 18, \ldots)}$

$$= \frac{x^n}{(3n-1)(5n-2)} \quad \left[\begin{array}{l} \because n^{th} \text{ term of } x, x^2, x^3, x^4, \ldots \text{ which is a GP is} \\ ar^{n-1} = x(x)^{n-1} = x^n \\ n^{th} \text{ term of } 2, 5, 8, 11, \ldots \text{ which is an AP is} \\ a + (n-1)d = 2 + (n-1)(3) = 3n-1 \\ \text{Similarly, } n^{th} \text{ term of } 3, 8, 13, 18, \ldots \text{ which is also} \\ \text{an AP is } a + (n-1)d = 3 + (n-1)(5) = 5n-2 \end{array}\right]$$

$$\therefore a_{n+1} = \frac{x^{n+1}}{\left(3(n+1)-1\right)\left(5(n+1)-2\right)} = \frac{x^{n+1}}{(3n+2)(5n+3)}$$

Now,

$$\lim_{n\to\infty} \frac{a_n}{a_{n+1}} = \lim_{n\to\infty} \frac{\dfrac{x^n}{(3n-1)(5n-2)}}{\dfrac{x^{n+1}}{(3n+2)(5n+3)}} = \lim_{n\to\infty} \frac{(3n+2)(5n+3)}{(3n-1)(5n-2)} \frac{1}{x}$$

$$= \lim_{n\to\infty} \frac{\left(3+\frac{2}{n}\right)\left(5+\frac{3}{n}\right)}{\left(3-\frac{1}{n}\right)\left(5-\frac{2}{n}\right)} \frac{1}{x} = \frac{(3+0)(5+0)}{(3-0)(5-0)} \frac{1}{x} = \frac{1}{x}$$

So, by D Alembert Ratio test given series converges if $\frac{1}{x} > 1 \Rightarrow x < 1$, diverges if $\frac{1}{x} < 1 \Rightarrow x > 1$ and test fails if $\frac{1}{x} = 1 \Rightarrow x = 1$. [refer (4.74) to (4.76)]

Now, when x = 1, then

$$a_n = \frac{x^n}{(3n-1)(5n-2)} = \frac{1}{(3n-1)(5n-2)} = \frac{1}{n^2\left(3-\frac{1}{n}\right)\left(5-\frac{2}{n}\right)}$$

Let $b_n = \frac{1}{n^2}$

Therefore,

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{\dfrac{1}{n^2\left(3-\frac{1}{n}\right)\left(5-\frac{2}{n}\right)}}{\dfrac{1}{n^2}} = \lim_{n\to\infty} \frac{1}{\left(3-\frac{1}{n}\right)\left(5-\frac{2}{n}\right)} = \frac{1}{(3-0)(5-0)} = \frac{1}{15}$$

which is non-zero finite number. Thus, by comparison test series $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ converges or diverges together [refer (4.71)]. But series $\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} \frac{1}{n^2}$ converges by p-test as p = 2 > 1. [refer (4.67)]. Hence, given series converges.

Now, you can try the following Self-Assessment Question.

SAQ 3

Test the convergence of the series $\sum_{n=1}^{\infty} \left(\frac{n}{n+2}\right)^{n^2}$.

4.5 ABSOLUTE AND CONDITIONAL CONVERGENCE

In Sec. 4.4 you have studied about convergence or divergence of positive term series. But there may be series having infinite terms of both signs positive and negative such series are known as **series of arbitrary terms**. If we have a series $\sum_{n=1}^{\infty} u_n$ of arbitrary terms then we say that it **converges absolutely** if series $\sum_{n=1}^{\infty} |u_n|$ converges. … (4.87)

We say that series $\sum_{n=1}^{\infty} u_n$ of arbitrary terms **converges conditionally** if series $\sum_{n=1}^{\infty} u_n$ converges but $\sum_{n=1}^{\infty} |u_n|$ diverges. … (4.88)

We will not discuss idea of absolute and conditional convergence in detail. We shall discuss only one example which is based on Leibnitz's test. So, first we have to state Leibnitz's test. But Leibnitz's test is used to check the convergence behaviour of alternating term series only, therefore, first we define what we mean by alternating series.

Alternating Series: A series having terms alternatively positive and negative is known as alternating series.

For example, series $\sum_{n=1}^{\infty} (-1)^{n-1} n = 1 - 2 + 3 - 4 + 5 - 6 + 7 - 8 + \ldots$ is an alternating series.

Now, we state Leibnitz's test.

Leibnitz's Test: Alternating series

$\sum_{n=1}^{\infty} (-1)^{n-1} u_n = u_1 - u_2 + u_3 - u_4 + u_5 - u_6 + u_7 - u_8 + \ldots$ converges if

(i) $u_{n+1} \le u_n \;\; \forall\, n \in \mathbb{N}$ and (ii) $\lim_{n\to\infty} u_n = 0$, where $u_i > 0$ for all i. … (4.89)

Example 15: Test the convergence of the series $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$.

Solution: Here given series is $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$. Let $\sum_{n=1}^{\infty} u_n = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \sum_{n=1}^{\infty} (-1)^{n-1} v_n$,

where $v_n = \frac{1}{n}$

$\therefore\; v_{n+1} = \frac{1}{n+1}$

Since $n < n+1 \Rightarrow \frac{1}{n} > \frac{1}{n+1} \Rightarrow v_n > v_{n+1}$. Also, $\lim_{n\to\infty} v_n = \lim_{n\to\infty} \frac{1}{n} = 0$.

Thus, both conditions of Leibnitz's test satisfied and hence given alternating series converges [refer (4.89)]. But series $\sum_{n=1}^{\infty} |u_n| = \sum_{n=1}^{\infty} \left|\frac{(-1)^{n-1}}{n}\right| = \sum_{n=1}^{\infty} \frac{1}{n}$ diverges

by p-test as p = 1. [refer (4.67)]

Hence, given series converges conditionally not absolutely.

4.6 SUMMARY

A brief summary of what we have covered in this unit is given as follows:

- **Sequence** is a function having domain as the family of natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$ and range may be any set but in this course, we have discussed only sequences of real numbers.
- **Real Sequence**: A real sequence is a function whose domain is the set of all natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$ and range may be any subset of $\mathbb{R}$, the set of all real numbers.
- **Ways of Representing a Sequence**: Here we have discussed four ways of representing a sequence.
  - **First way**: A sequence can be represented by writing the first few terms till a definite rule for writing down other term becomes clear. For example, 1, 4, 9, 16, 25, … is a sequence written in this way.
  - **Second Way**: A sequence can also be represented by giving a formula for its nth term. For example, $< \frac{n+1}{n} >$.
  - **Third Way (Recursive Relation)**: A sequence can also be represented by specifying its first few terms and a formula to write remaining terms in terms of the preceding terms. For example, $a_1 = 1, a_2 = 1, a_{n+1} = a_n + a_{n-1} \ \forall \ n \geq 2, n \in \mathbb{N}$
  - **Fourth Way**: Sometimes sequence is such that it cannot be represented by giving a single formula for its n term. But its terms are generated by two or more relations. For example, $a_n = \begin{cases} \frac{n}{2}, & \text{if } n \text{ is even} \\ -\frac{n-1}{2}, & \text{if } n \text{ is odd} \end{cases}$
- **Convergence of a Sequence**: A real sequence $< a_n >$ is said to converges a real number 'a' if for a given $\varepsilon > 0$ however small it may be there exists a natural number m depending on ε such that $|a_n - a| < \varepsilon \ \forall \ n \geq m$.
- **Divergence of a Sequence**: A real sequence $< a_n >$ is said to diverges to ∞ if given any positive real number K no matter how large it is there exists a positive integer m such that $a_n > K \ \forall \ n \geq m$, and we write it as $a_n \to \infty$ as $n \to \infty$ or $\lim_{n\to\infty} a_n = \infty$.
- **Oscillatory Sequence**: If a real sequence $< a_n >$ neither converges to a finite real number and nor diverges to ∞ or $-\infty$, then we say that sequence oscillates. An oscillatory sequence may **oscillate finitely or infinitely**.
- **Bounded above Sequence**: A sequence $< a_n >$ is said to be bounded above if there exists a real number K such that $a_n \leq K \ \forall \ n \in \mathbb{N}$
- **Bounded below Sequence**: A sequence $< a_n >$ is said to be bounded

below if there exists a real number k such that $a_n \geq k \ \forall \ n \in \mathbb{N}$

- **Bounded Sequence**: A sequence $< a_n >$ is said to be bounded if it is both bounded above and bounded below.
- **Least Upper Bound**: Let $< a_n >$ be a real and bounded above sequence then a real number u is said to be least upper bound (lub) of the sequence $< a_n >$ if

  (i) $a_n \leq u \ \forall \ n \in \mathbb{N}$ and

  (ii) given any $\varepsilon > 0$, no matter how small it is, there exists at least one $n \in \mathbb{N}$ such that $a_n > u - \varepsilon$.
- **Greatest Lower Bound**: Let $< a_n >$ be a real and bounded below sequence then a real number g is said to be greatest lower bound (glb) of the sequence $< a_n >$ if

  (i) $a_n \geq g \ \forall \ n \in \mathbb{N}$ and

  (ii) given any $\varepsilon > 0$, no matter how small it is, there exists at least one $n \in \mathbb{N}$ such that $a_n < g + \varepsilon$.
- **Monotonically Increasing Sequence**: A real sequence $< a_n >$ is said to be monotonically increasing sequence if $a_n \leq a_{n+1} \ \forall \ n \in \mathbb{N}$
- **Monotonically Decreasing Sequence**: A real sequence $< a_n >$ is said to be monotonically decreasing sequence if $a_n \geq a_{n+1} \ \forall \ n \in \mathbb{N}$
- **Monotonic Sequence**: A real sequence $< a_n >$ is said to be monotonic if either it is monotonically increasing or monotonically decreasing sequence.
- **Cauchy Sequence**: A sequence $< a_n >$ is said to be a Cauchy sequence if given $\varepsilon > 0$, no matter how small it is there exists a positive integer n_0 such that $|a_n - a_m| < \varepsilon \ \forall \ m, n \geq n_0$
- The expression $a_1 + a_2 + a_3 + a_4 + \ldots$ or $\sum_{n=1}^{\infty} a_n$ is known as **infinite series**.
- **Sequence of Partial Sum**: If we have an infinite series $\sum_{n=1}^{\infty} a_n$ then the sequence $s_1, s_2, s_3, s_4, \ldots, s_n, \ldots,$ where

  $s_1 = a_1, s_2 = a_1 + a_2, s_3 = a_1 + a_2 + a_3, s_4 = a_1 + a_2 + a_3 + a_4, \ldots,$
  $s_n = a_1 + a_2 + a_3 + \ldots + a_n, \ldots$

  is known as sequence of partial sum of the given infinite series $\sum_{n=1}^{\infty} a_n$.
- **Convergence and Divergence of Infinite Series**: If the sum of infinite terms $a_1 + a_2 + a_3 + \ldots + a_n + \ldots$ is a real number, then we say that infinite series $\sum_{n=1}^{\infty} a_n$ converges, otherwise diverges.
- If a series have infinite terms of both signs positive and negative such series are known as **series of arbitrary terms**.
- If we have a series $\sum_{n=1}^{\infty} u_n$ of arbitrary terms then we say that it **converges**

absolutely if series $\sum_{n=1}^{\infty}|u_n|$ converges.

- We say that series $\sum_{n=1}^{\infty}u_n$ of arbitrary terms **converges conditionally** if series $\sum_{n=1}^{\infty}u_n$ converges but $\sum_{n=1}^{\infty}|u_n|$ diverges.

- **Alternating Series**: A series having terms alternatively positive and negative is known as alternating series.

- We have also mentioned some standard results related to test of convergence of sequence and series without proof. We have also learnt to apply these tests with the help of many examples.

4.7 TERMINAL QUESTIONS

1. Test the convergence of the series
$$\frac{3}{4}+\frac{7}{3}+\frac{3}{4^2}+\frac{7}{3^2}+\frac{3}{4^3}+\frac{7}{3^3}+\frac{3}{4^4}+\frac{7}{3^4}+\ldots$$

2. Test the convergence of the series $\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n^2}$.

4.8 SOLUTIONS/ANSWERS

Self-Assessment Questions (SAQs)

1. In numerator, we have an AP, so n^{th} term in numerator is given by
$$a+(n-1)d=9+(n-1)(2)=9+2n-2=2n+7$$
In denominator it is obvious that n^{th} term is n^2. Hence, n^{th} term of the given sequence is $\frac{2n+7}{n^2}$.

2. If $a_n=\frac{1+2^{1/2}+3^{1/3}+\ldots+n^{1/n}}{n}$, then $a_3=\frac{1+2^{1/2}+3^{1/3}}{3}$ and
$$a_{n+1}=\frac{1+2^{1/2}+3^{1/3}+\ldots+n^{1/n}+(n+1)^{1/(n+1)}}{n+1}.$$

3. Here $a_n=\left(\frac{n}{n+2}\right)^{n^2}$, then
$$\lim_{n\to\infty}(a_n)^{1/n}=\lim_{n\to\infty}\left(\frac{n}{n+2}\right)^n=\lim_{n\to\infty}\left(\frac{1}{1+\frac{2}{n}}\right)^n=\lim_{n\to\infty}\frac{(1)^n}{\left(1+\frac{2}{n}\right)^n}=\lim_{n\to\infty}\frac{1}{\left(\left(1+\frac{2}{n}\right)^{\frac{n}{2}}\right)^2}$$
$$=\frac{1}{\left(\lim_{n\to\infty}\left(1+\frac{2}{n}\right)^{\frac{n}{2}}\right)^2}=\frac{1}{(e)^2}\left[\because \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x=e\right]$$

But $\frac{1}{e^2} < 1$ $\left[\begin{array}{l} \because \text{ Refer (1.32), } e = 1 + \frac{1}{\lfloor 1} + \frac{1}{\lfloor 2} + \frac{1}{\lfloor 3} + \ldots \text{ which is } > 1 \\ \Rightarrow e^2 > 1 \Rightarrow \frac{1}{e^2} < 1 \end{array}\right]$

Hence, by Cauchy's root test given series converges. [refer (4.78)]

Terminal Questions

1. Here $\sum_{n=1}^{\infty} a_n = \frac{3}{4} + \frac{7}{3} + \frac{3}{4^2} + \frac{7}{3^2} + \frac{3}{4^3} + \frac{7}{3^3} + \frac{3}{4^4} + \frac{7}{3^4} + \ldots$

$$= \left(\frac{3}{4} + \frac{3}{4^2} + \frac{3}{4^3} + \frac{3}{4^4} + \ldots\right) + \left(\frac{7}{3} + \frac{7}{3^2} + \frac{7}{3^3} + \frac{7}{3^4} + \ldots\right)$$

$$= 3\left(\frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \frac{1}{4^4} + \ldots\right) + 7\left(\frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^3} + \frac{1}{3^4} + \ldots\right)$$

$$= \sum_{n=1}^{\infty} 3b_n + \sum_{n=1}^{\infty} 7c_n$$

Series $\sum_{n=1}^{\infty} b_n = \left(\frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \frac{1}{4^4} + \ldots\right)$, $\sum_{n=1}^{\infty} c_n = \left(\frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^3} + \frac{1}{3^4} + \ldots\right)$ both being geometric series with common ratio $\frac{1}{4} < 1$ and $\frac{1}{3} < 1$ respectively converges. [refer (4.65)]. Further, we know that if series $\sum_{n=1}^{\infty} b_n$ converges then $\sum_{n=1}^{\infty} c\, b_n$ also converges for finite real number c. [refer (4.83)]

Thus, series $\sum_{n=1}^{\infty} 3b_n$ and $\sum_{n=1}^{\infty} 7c_n$ both converges. We also know that if series $\sum_{n=1}^{\infty} u_n$ and $\sum_{n=1}^{\infty} v_n$ converges then their sum $\sum_{n=1}^{\infty} (u_n + v_n)$ converges. [refer (4.85)].

Hence, given series converges.

2. Here given series is $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2}$. Let $\sum_{n=1}^{\infty} u_n = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2} = \sum_{n=1}^{\infty} (-1)^{n-1} v_n$,

where $v_n = \frac{1}{n^2}$

$\therefore\ v_{n+1} = \frac{1}{(n+1)^2}$

Since $n^2 < (n+1)^2 \Rightarrow \frac{1}{n^2} > \frac{1}{(n+1)^2} \Rightarrow v_n > v_{n+1}\ \ \forall\ n \in \mathbb{N}$

Also, $\lim_{n\to\infty} v_n = \lim_{n\to\infty} \frac{1}{n^2} = 0.$

Thus, both conditions of Leibnitz's test satisfied and hence given alternating series converges [refer (4.89)]. Also, series

$$\sum_{n=1}^{\infty}|u_n| = \sum_{n=1}^{\infty}\left|\frac{(-1)^{n-1}}{n^2}\right| = \sum_{n=1}^{\infty}\frac{1}{n^2}$$ converges by p-test as p = 2>1. [refer (4.67)]

Hence, given series is absolutely convergent also.

UNIT 5

RIEMANN INTEGRATION

Structure

5.1 Introduction

Expected Learning Outcomes

5.2 A Review of Integration

How the Job of Summation in the Discrete World is done by Integration in the Continuous World and in What Way

How Integration gives Area Bounded by Four Things: Curve of the Function, Lower and Upper Limits of Integration and the Axis of Integration.

5.3 Two Definitions of Riemann Integration

5.4 Properties of Riemann Integration

5.5 Summary

5.6 Terminal Questions

5.7 Solutions/Answers

5.1 INTRODUCTION

From earlier classes, you are familiar with sigma notation, e.g., $\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \ldots + x_n$, which is used to sum at the most countable observations in a discrete world. But in the field of statistics and data science you not only need to deal with countable values in a discrete world but you also need to deal with uncountable values in a continuous world. How the role of summation in a discrete world is played by integration in a continuous world and in what way is explained in detail in Sec. 5.2. After understanding the concept that how integration plays the role of summation and in what way, the next thing, being a student of statistics or data scientist that you should know are the two definitions of Riemann integration. These two definitions of Riemann integration and some terms related to the two definitions are explained in detail in Sec. 5.3. Some properties of Riemann integration are stated without proof in Sec. 5.4. From earlier classes you are familiar with three important ideas continuity, differentiation and integration. These three concepts have some connections which are required to study some important ideas of probability theory. So, some such connections are also discussed in Sec. 5.4.

What we have discussed in this unit is summarised in Sec. 5.5. Self-Assessment Questions (SAQs) have been given in some sections which are generally based on the content discussed in that section. But to give you a good practice of what we have discussed in this unit some more questions based on the entire unit are given in Sec. 5.6 under the heading Terminal Questions. Due to the reason mentioned in Sec. 1.1 of Unit 1 of this course, solutions of all the SAQs and Terminal Questions are given in Sec. 5.7.

In the next unit, you will study to write the equation of hyperplane.

Expected Learning Outcomes

After completing this unit, you should be able to:

- ❖ explain how and in what way the role of summation in the discrete world is played by integration in the continuous world;
- ❖ explain the fact that integration gives an area bounded by four things: (i) the function which we have to integrate (ii) the lower limit of integration (iii) the upper limit of integration and (iv) the axis of integration;
- ❖ apply both definitions of Riemann integration to solve examples; and
- ❖ state some properties and theorems on Riemann integration. Prove some theorems on Riemann integration which will be used in understanding some concepts of probability theory in the course MST-012.

5.2 A REVIEW OF INTEGRATION

In school mathematics, you have seen integration as an antiderivative of a function. For example, derivative of $\sin x + c$ is $\frac{d}{dx}(\sin x + c) = \cos x$ and integration of $\cos x$ is $\int \cos x\, dx = \sin x + c$. From earlier classes you also know one important application of definite integration is that it gives you an area under the function to which we are integrating and bounded by lower limit, upper limit and axis of integration. For example, the integration

$$\int_1^3 x(5-x)\,dx = \int_1^3 \left(5x - x^2\right)dx = \left[\frac{5x^2}{2} - \frac{x^3}{3}\right]_1^3 = \left(\frac{45}{2} - \frac{27}{3} - \frac{5}{2} + \frac{1}{3}\right)$$

$$= \frac{45}{2} - 9 - \frac{5}{2} + \frac{1}{3} = \frac{135 - 54 - 15 + 2}{6} = \frac{68}{6} = \frac{34}{3}\text{unit}^2$$

gives the area of the shaded region shown in blue colour in Fig. 5.1 (a), which is bounded by one curve and three lines mentioned as follows: … (5.1)

Curve: $y = x(5-x)$ (corresponding to the function we are integrating),

First Line: $x = 1$ (corresponding to equation of lower limit of integration),

Second Line: $x = 3$ (corresponding to equation of upper limit of integration),

Third line: y = 0 (corresponding to equation of axis of variable of integration)

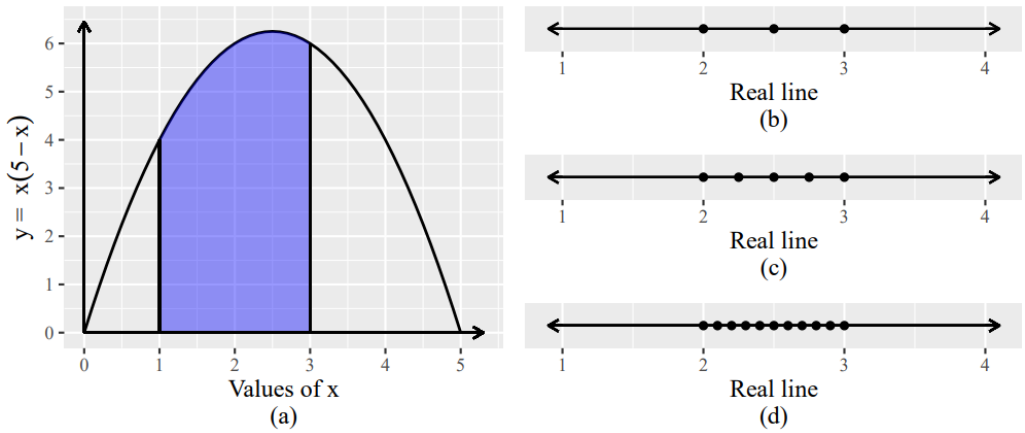


Fig. 5.1: (a) Visualisation of area under the curve of the function x(5 – x) in blue colour and visualisation of points between 2 and 3 including 2 and 3 (b) 3 points (c) 5 points (d) 11 points

In this section we will discuss two things:

(1) How the job of summation in a discrete world is done by integration in a continuous world and in what way? Understanding this idea is very important for a statistician and a data scientist because whenever we want to obtain a statistical measure in a continuous world, we just replace summation by integration in the corresponding formula of the discrete world. For example, the expected value of a discrete random variable X is given by $E(X) = \sum_{x=1}^{\infty} x\,p(x)$ while expected value of a continuous random variable X is given by $E(X) = \int_{-\infty}^{\infty} x\,f(x)\,dx$, where p(x) is the probability mass function of the discrete random variable X and f(x) is the probability density function of continuous random variable X. You will study about this and other statistical measures in detail in Units 6 to 14 of the course MST-012. … (5.2)

(2) How integration gives area bounded by four things (i) under the curve of the function to which we are integrating (ii) equation corresponding to the lower limit (iii) equation corresponding to the upper limit (iv) equation corresponding to the axis of integration. Understanding this idea is also very important for a statistician and a data scientist because the idea of probability in the case of continuous distributions is based on this important mathematical idea. … (5.3)

We will explain both points (1) and (2) mentioned in (5.2) and (5.3) in the following two subsections.

5.2.1 How the Job of Summation in the Discrete World is done by Integration in the Continuous World and in What Way

From the discussion of Sec. 2.3 in Unit 2 of this course you know the meaning of a countable set. If we have at the most countable values then we use summation to deal with their sum. For example, if we have

(a) only 3 numbers 2, 2.5, 3 between 2 and 3 including 2 and 3, then to add all of them we can use summation notation as follows.

$\sum_{x=1}^{3} x_i = x_1 + x_2 + x_3 = 2 + 2.5 + 3 = 7.5,$ where $x_1 = 2,\ x_2 = 2.5,\ x_3 = 3.$ To see it geometrically refer to Fig. 5.1 (b). … (5.4)

(b) only 5 numbers 2, 2.25, 2.5, 2.75, 3 between 2 and 3 including 2 and 3, then to add all of them we can use summation notation as follows.

$\sum_{x=1}^{5} x_i = x_1 + x_2 + x_3 + x_4 + x_5 = 2 + 2.25 + 2.5 + 2.75 + 3 = 12.5,$ where $x_1 = 2,\ x_2 = 2.25,\ x_3 = 2.5,\ x_4 = 2.75,\ x_5 = 3.$ To see it geometrically refer to Fig. 5.1 (c). … (5.5)

(c) only 11 numbers 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3 between 2 and 3 including 2 and 3, then to add all of them we can use summation notation as follows.

$$\sum_{x=1}^{11} x_i = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10} + x_{11}$$

$$= 2 + 2.1 + 2.2 + 2.3 + 2.4 + 2.5 + 2.6 + 2.7 + 2.8 + 2.9 + 3 = 27.5,$$

where $x_1 = 2,\ x_2 = 2.1,\ x_3 = 2.2,\ x_4 = 2.3,\ x_5 = 2.4,\ x_6 = 2.5,\ x_7 = 2.6,$ $x_8 = 2.7,\ x_9 = 2.8,\ x_{10} = 2.9,\ x_{11} = 3.$ To see it geometrically refer to Fig. 5.1 (d). … (5.6)

(d) In this way if we go on increasing numbers between 2 and 3 including 2 and 3 then we can add them using summation if they are at the most countable numbers. But if we want to add all the numbers between 2 and 3 including 2 and 3 then you cannot do so using summation because they are not countable. In fact, you even cannot tell what is the next number after 2. If you say 2.1 is the next real number after 2, no it is not the next real number after 2 because 2.01 is between 2 and 2.1. If you say 2.01 is the next real number after 2, no it is not the next real number after 2 because 2.001 is between 2 and 2.01. Similarly, if you say 2.001 is the next real number after 2, no it is not the next real number after 2 because 2.0001 is between 2 and 2.001 and so on. **So, summation notation fails to work in the case of uncountable numbers**. **So, what is the solution? The solution is integration**. But integration does not add the individual numbers like the way we add in summation. **Integration actually measures the size of the shape constituted by all the points of interest**. In the present case, all the points between 2 and 3 including 2 and 3 constitute line segments joining point 2 and point 3 on the real line. So, the shape constituted by all possible numbers between 2 and 3 including 2 and 3 is a line segment. So, integration should give us the size of this line segment. We know that measure of a line segment is its length. So, the length of the line segment joining points 2 and 3 on the real line is 3 – 2 = 1. So, we expect integration should give us value 1. Let us evaluate the integration and verify it. … (5.7)

$$\int_2^3 dx = [x]_2^3 = 3 - 2 = 1. \qquad \text{… (5.8)}$$

Great integration has done the job that we expected from it.

Now, our next target is to geometrically visualise LHS of (5.8) so that understanding of the idea of integration becomes clear to you. First of all, let us denote points 2 and 3 on real line by points A and B respectively. Look at Fig. 5.2 (a) and consider small portion of the line segment AB which starts from point A and denote it by dx. We are using variable x because generally horizontal axis is represented by variable X. The length dx which is shown in Fig. 5.2 (a) is very large compared to what actually it represents mathematically. Mathematically it is very small which even cannot be seen by eyes. But to understand the idea we are representing it much longer than what it actually means mathematically. Now if we join such very small portions dx end to end (refer Fig. 5.2 (b) and (c)) $1/\,dx$ many times we will finally reach at point B. For example, in particular, (i) if dx = 0.001 then $\frac{1}{dx}=\frac{1}{0.001}=1000$ so we have to join 1000 such dx end to end to reach at the point B. (ii) if dx = 0.0001 then $\frac{1}{dx}=\frac{1}{0.0001}=10000$ so we have to join 10000 such dx end to end to reach at the point B. (iii) if dx = 0.00001 then $\frac{1}{dx}=\frac{1}{0.00001}=100000$ so we have to join 100000 such dx end to end to reach at the point B and so on. So, mathematically it is possible by starting from point A to reach at the point B by taking dx > 0 no matter how small it is. So, you can imagine that these small pieces of length each equal to dx will ultimately constitute the complete line segment AB where you need $1/\,dx$ such small pieces which is mathematically possible. Before the entry of integration in to picture one more clarification is required. We saw that: if we take dx = 0.001 then we need to join 1000 such small pieces, if we take dx = 0.0001 then we need to join 10000 such small pieces, if we take dx = 0.00001 then we need to join 100000 such small pieces. But then steps fall in the category of countable and so we can use summation. Now, here the idea of calculus enters into the picture which forces us to replace summation with integration. This idea of calculus is $dx \to 0^+$ and you know the meaning of something tends to 0 from right from your school mathematics. Conceptually it is so small that we actually cannot specify its value, e.g., you can think of its value like this $dx = 0.\underbrace{00000000\ldots001}_{\text{very very large no of times}}$. So, we have to apply the appropriate tool of calculus to handle this mathematically. The appropriate tool in calculus to obtain a measure of the shape constituted by joining such tends to zero things end to end is integration. That is the explanation of the symbol used in LHS of (5.8).

… (5.9)

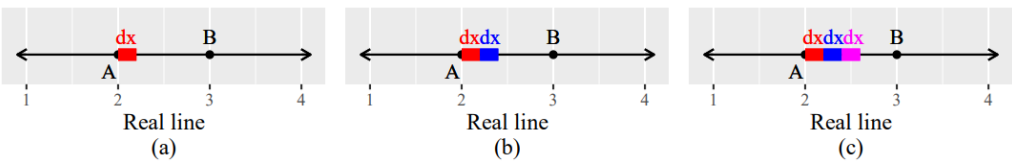


Fig. 5.2: Visualisation of small piece of length dx (a) first piece in red colour (b) one more dx in blue colour is added (c) one more dx in magenta colour is added

This completes the discussion of the point (1) where we explained how the job of summation in discrete word is done by integration in continuous word and in what way.

5.2.2 How Integration gives Area Bounded by Four Things: Curve of the Function, Lower and Upper Limits of Integration and the Axis of Integration

To explain it we need four things. (i) one function which is to be integrated (ii) lower limit of the integration (iii) upper limit of the integration (iv) variable of integration. You know that limits of integration may be $-\infty$ or ∞. But to understand the idea more clearly, we will consider both lower and upper limits of integration some finite real numbers. Let us start the explanation of point (2) mentioned in (5.3) by considering the required four things as follows.

Function: y = f(x) = x + 1 … (5.10)

Lower limit: x = 1 … (5.11)

Upper limit: x = 3 … (5.12)

Variable of integration: x … (5.13)

First of all, let us plot these four things mentioned in (5.10) to (5.13) and shad the region in blue colour bounded by the function given by (5.10) and three lines given by (5.11) to (5.13) in Fig. 5.3 (a).

We will explain why and how the integration $\int_1^3 (x+1)dx$ represents/gives the area shown in Fig. 5.3 (a) in blue colour. Let us first evaluate the area of the shaded region in Fig. 5.3 (a) without using integration. Note that:

$$\text{Area of the shaded region in Fig. 5.3 (a)} = \text{Area of the trapezium ABCD}$$

$$= \frac{1}{2}(AD+BC)AB \quad \left[\because \text{Area of the trapezium} = \frac{1}{2}(\text{Sum of parallel sides}) \times \text{height}\right]$$

$$= \frac{1}{2}(2+4)(2) = \frac{1}{2}(6)(2) = 6 \text{ unit}^2 \qquad \text{... (5.14)}$$

Now, we explain how this area can be obtained using integration. Consider a vertical strip of small width dx starting from point A refer to Fig. 5.3 (b). The shaded area in blue colour in Fig. 5.3 (a) is made up by joining such vertical strips end to end. In Fig. 5.3 (c) we have added one more vertical strip of blue colour adjacent to the first red colour strip. In Fig. 5.3 (d) we have added one more vertical strip of magenta colour adjacent to the second strip of blue colour. If we keep on adding strips adjacent to the previous one, one at a time, ultimately, we will reach at point B. So, the required area is equal to the sum of the areas of all these vertical strips. But before obtaining the areas of all such vertical strips we have to consider small pieces of vertical height like we have dx along the horizontal direction. As usual along the vertical axis, we have variable y or values of the function f(x) = x + 1. So, a small portion of the height along the vertical direction is denoted by dy like dx along the horizontal direction. So, every vertical strip is made up of joining such small pieces dy. In Fig. 5.3 (b) to (d) we have shown one small piece of height dy in each vertical strip by dy while in Fig. 5.3 (e) we have shown 3 such pieces within a single

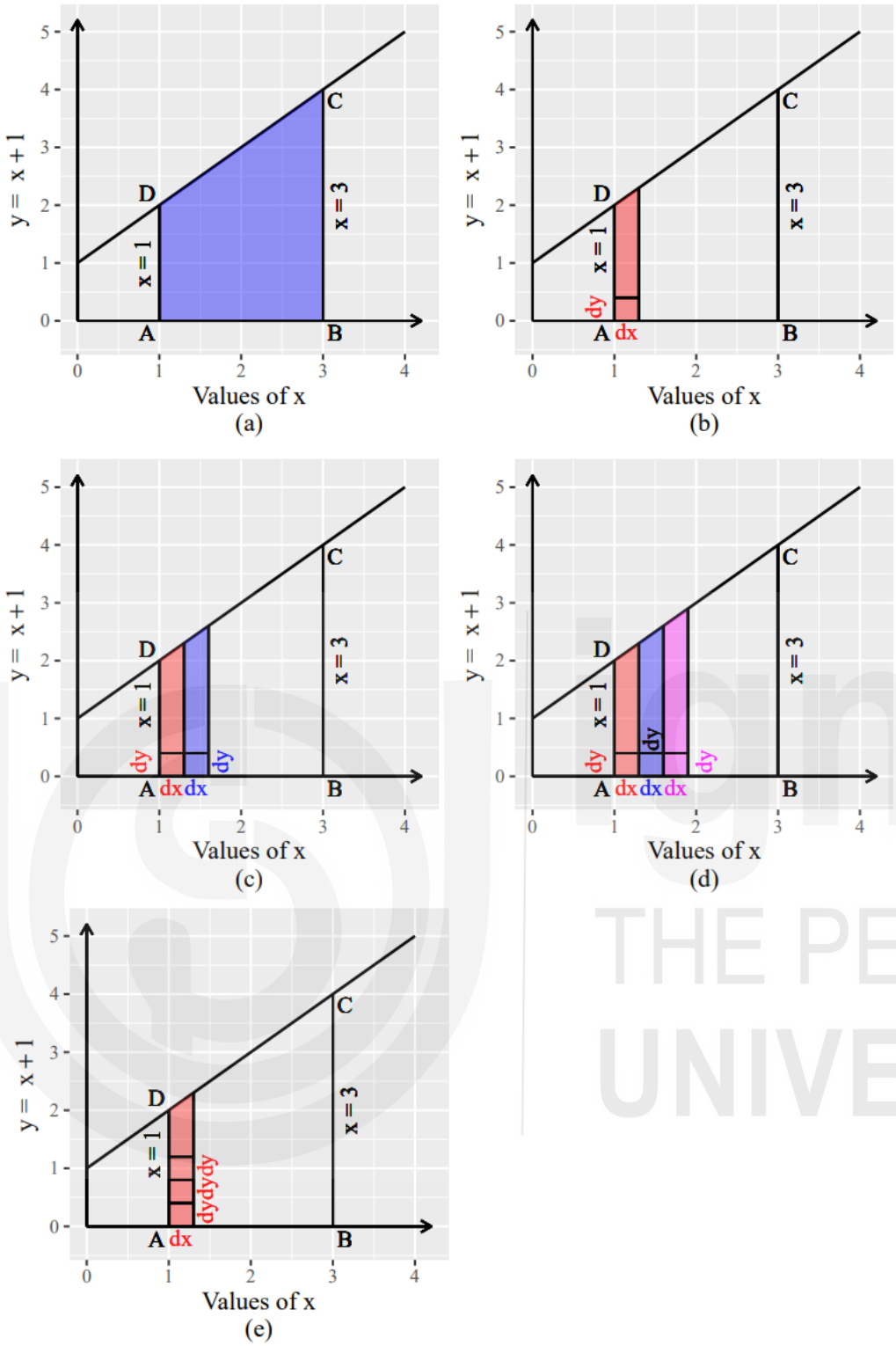


Fig. 5.3: Visualisation of (a) the region bounded by (5.10) to (5.13) (b) the first vertical strips of width dx in red colour (c) one more strip of width dx in blue colour is added (d) one more strip of width dx in magenta colour is added (e) 3 small pieces of height dy within a single strip of width dx

vertical strip each of height dy. So, to obtain area of the shaded region in Fig. 5.3 (a) we have to do two things: (i) first we have to calculate area of one vertical strip and secondly (ii) combine areas of all vertical strips.

To obtain area of one vertical strip is exactly similar to what we have done in explanation of point (1) refer (5.9). Now, as we know dy and dx both are very small, in fact $dy \to 0$ and $dx \to 0$ so the shape constituted by $dy \to 0$ and

$dx \to 0$ is the vertical line segment which starts from the horizontal axis where y = 0 and ends at the curve of the function where y = f(x). So, due to the discussion of point (1) in (5.9) we know that the appropriate tool of calculus to measure the shape constituted by joining such tends to zero things end to end is integration and so:

$$\text{Length of vertical strip} = \int_{0}^{f(x)} dy \qquad \ldots (5.15)$$

Equation (5.15) gives measure of one vertical strip at a general point from x = 1 to x = 3. Thus, the required shaded area is the measure of the shape obtained by joining all such possible vertical line segments (whose measure is given by (5.15)) from x = a = 1 to x = b = 3. But these vertical strips are each of width dx where $dx \to 0$. Again due to the discussion of point (1) in (5.9) we know that the appropriate tool of calculus to measure the shape constituted by joining such tends to zero things end to end is integration and so required shaded area shown in Fig. 5.3 (a) is given by:

$$\text{Area of the shaded region shown in Fig. 5.3 (a)} = \int_{x=a}^{x=b}\left(\int_{0}^{f(x)} dy\right)dx \qquad \ldots (5.16)$$

Note that in equation (5.16) we have two signs of integration known as double integration which is discussed in Unit 9 of this course. But after solving inside integral the expression given in (5.16) can simply be written as follows.

$$\int_{x=a}^{x=b}\left(\int_{0}^{f(x)} dy\right)dx = \int_{x=a}^{x=b}\left([y]_{y=0}^{y=f(x)}\right)dx = \int_{x=a}^{x=b}\left(f(x)-0\right)dx = \int_{x=a}^{x=b} f(x)\,dx \qquad \ldots (5.17)$$

The equation (5.17) gives area of the shaded region shown in Fig. 5.3 (a) which is bounded by one curve and three lines mentioned in (5.10) to (5.13) where a = 1 and b = 3. … (5.18)

But in Fig. 5.3 (a), we have particular values of f(x), a and b where f(x) = x + 1, a = 1 and b = 3. So, finally, area of the shaded region shown in Fig. 5.3 (a) is given by just replacing f(x) by x + 1, a by 1 and b by 3 in (5.17). So, we have

$$\text{Area of the shaded region shown in Fig. 5.3 (a)} = \int_{x=a}^{x=b} f(x)\,dx = \int_{x=1}^{x=3}(x+1)\,dx$$

$$= \left[\frac{x^2}{2}+x\right]_1^3 = \frac{9}{2}+3-\frac{1}{2}-1 = 6\,\text{unit}^2 \qquad \ldots(5.19)$$

Great! This matches with the value of the area which we obtained without using integration refer (5.14).

This completes the discussion of the point (2) mentioned in equation (5.3) where we explained how integration gives area bounded by four things (i) function or curve which we are integrating (ii) equation corresponding to the lower limit of integration (iii) equation corresponding to the upper limit of integration (iv) equation corresponding to the axis of integration.

5.3 TWO DEFINITIONS OF RIEMANN INTEGRATION

To understand definition of Riemann Integration you should have understanding of following terms.

- Closed and Finite Interval
- Partition of a Finite and Closed Interval
- Norm of a Partition
- Refinement of a Partition
- Least Upper Bound (lub) and Greatest Lower Bound of a Bounded Function over an Interval
- Lower and Upper Sums of a Function over a Partition (Darboux Sums)
- Oscillatory Sum
- Upper and Lower Riemann Integrals

So, let us explain what we mean by these terms one at a time.

Closed and Finite Interval

In equations (1.11) and (1.15) in Unit 1 of this course we have discussed both closed and finite interval respectively. So, before moving ahead just revise what is discussed there.

Partition of a Finite and Closed Interval

If I = [a, b] be a finite and closed interval then by a partition P of I, we mean a finite ordered set $\{x_0, x_1, x_2, x_3, \ldots, x_n\}$ of points of [a, b] such that

$$a = x_0 < x_1 < x_2 < x_3 < \ldots < x_{n-1} < x_n = b. \qquad \ldots (5.20)$$

The n + 1 points $x_0, x_1, x_2, x_3, \ldots, x_n$ are called partition points of P and the n sub-intervals

$$I_1 = [x_0, x_1], I_2 = [x_1, x_2], I_3 = [x_2, x_3], \ldots, I_i = [x_{i-1}, x_i], \ldots, I_n = [x_{n-1}, x_n] \quad \ldots (5.21)$$

determined by n + 1 points of P are called the segments of the partition P. Fig. 5.4 visualise the positions of n + 1 points $x_0, x_1, x_2, x_3, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n$ and n sub-intervals $I_1, I_2, I_3, \ldots, I_i, I_{i+1}, \ldots, I_n$.

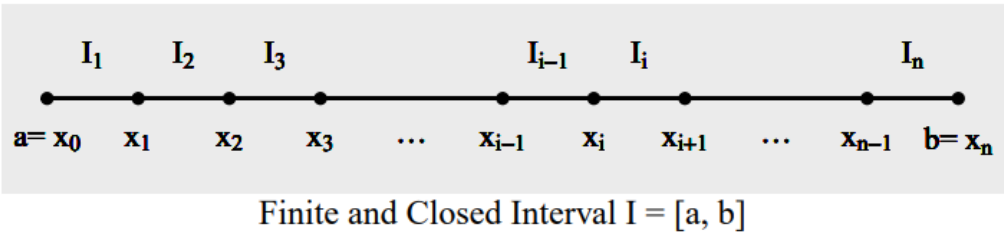


Fig. 5.4: Visualisation of the points of partition of the finite and closed interval I = [a, b] and n sub-intervals $I_1, I_2, I_3, \ldots, I_n$

Notation for Length of Sub-Intervals Formed by Points of a Partition

If I = [a, b] be a finite and closed interval and $P = \{x_0, x_1, x_2, x_3, \ldots, x_n\}$ be a partition of [a, b] such that $a = x_0 < x_1 < x_2 < x_3 < \ldots < x_{n-1} < x_n = b$, then n + 1 points $x_0, x_1, x_2, \ldots, x_n$ forms n sub-intervals given by $I_1 = [x_0, x_1], I_2 = [x_1, x_2]$,

$I_3 = [x_2, x_3], \ldots, I_i = [x_{i-1}, x_i], \ldots, I_n = [x_{n-1}, x_n]$. The lengths of these n sub-intervals are respectively denoted by δ_i, $i = 1, 2, 3, \ldots, n$. That is,

$$\delta_i = x_i - x_{i-1}, \; i = 1, 2, 3, \ldots, n. \qquad \ldots (5.22)$$

For example, if I = [2, 5] and P = {2, 2.5, 4, 5} then $I_1 = [2, 2.5], I_2 = [2.5, 4]$, $I_3 = [4, 5]$ and so $\delta_1 = 2.5 - 2 = 0.5, \delta_2 = 4 - 2.5 = 1.5, \delta_3 = 5 - 4 = 1$.

Norm of a Partition

If I = [a, b] be a finite and closed interval and $P = \{x_0, x_1, x_2, x_3, \ldots, x_n\}$ be a partition of [a, b], then norm of the partition P is denoted by $\|P\|$ and defined as follows.

$$\|P\| = \max\{x_i - x_{i-1} : i = 1, 2, 3, 4, \ldots, n\} = \max\{\delta_i : i = 1, 2, 3, 4, \ldots, n\} \qquad \ldots (5.23)$$

For example, if I = [2, 5] and P = {2, 2.5, 4, 5} then $I_1 = [2, 2.5], I_2 = [2.5, 4]$, $I_3 = [4, 5]$ and so $\delta_1 = 2.5 - 2 = 0.5, \delta_2 = 4 - 2.5 = 1.5, \delta_3 = 5 - 4 = 1$.

Hence, $\|P\| = \max\{\delta_i : i = 1, 2, 3\} = \max\{0.5, 1.5, 1\} = 1.5$

Refinement of a Partition

If I = [a, b] be a finite and closed interval and P_1 be a partition of [a, b], then another partition P_2 of [a, b] is called refinement of the partition P_1 if $P_1 \subset P_2$. That is, every member of P_1 is also a member of P_2. If the partition P_2 is a refinement of partition P_1 then we also say that P_2 is finer than P_1. … (5.24)

For example, if I = [2, 5] and $P_1 = \{2, 2.5, 4, 5\}$, $P_2 = \{2, 2.5, 3, 4, 5\}$ are two partitions of I = [2, 5] then P_2 is a refinement of P_1 since all elements of P_1 are also elements of P_2 also partition P_2 has one additional element 3.

Least Upper Bound (lub) and Greatest Lower Bound of a Bounded Function over an Interval

In Unit 4 of this course, we have defined both lub and glb refer (4.23) to (4.26) just replace sequence by function, a_n by f(x) and $\mathbb{N}$ by domain of the function. So, before moving ahead just revise what is discussed there.

Lower and Upper Sums of a Function over a Partition (Darboux Sums)

Lower and Upper sums of a bounded function over a finite and closed interval [a, b] are defined as follows.

Let I = [a, b] be a finite and closed interval and $P = \{x_0, x_1, x_2, x_3, \ldots, x_n\}$ be a partition of [a, b]. So, this partition divides the interval [a, b] into n sub-intervals $I_i = [x_{i-1}, x_i]$, $i = 1, 2, 3, \ldots, n$ and lengths of these sub-intervals are $\delta_i = x_i - x_{i-1}$, $i = 1, 2, 3, \ldots, n$.

Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Since f is a bounded function on [a, b] so it will be bounded on each of the sub-intervals $I_i = [x_{i-1}, x_i]$, $i = 1, 2, 3, \ldots, n$.

So, lub and glb of f on each $I_i = [x_{i-1}, x_i],\ i = 1, 2, 3, \ldots, n$ will exist. Suppose

$$m_i = \text{glb}\{f(x) : x \in [x_{i-1}, x_i]\},\ M_i = \text{lub}\{f(x) : x \in [x_{i-1}, x_i]\},\ i = 1, 2, 3, 4, \ldots, n \quad (5.25)$$

then the sum

$$m_1\delta_1 + m_2\delta_2 + m_3\delta_3 + m_4\delta_4 + \ldots + m_n\delta_n = \sum_{i=1}^{n} m_i\delta_i \quad \ldots (5.26)$$

is called the **lower Darboux sum** of f over the partition P and is denoted by

$$L(f, P).\ \text{So},\ L(f, P) = \sum_{i=1}^{n} m_i\delta_i \quad \ldots (5.27)$$

and the sum

$$M_1\delta_1 + M_2\delta_2 + M_3\delta_3 + M_4\delta_4 + \ldots + M_n\delta_n = \sum_{i=1}^{n} M_i\delta_i$$

is called the **upper Darboux sum** of f over the partition P and is denoted by

$$U(f, P).\ \text{So},\ U(f, P) = \sum_{i=1}^{n} M_i\delta_i \quad \ldots (5.28)$$

Oscillatory Sum

If we consider the function f and all other terms same as defined in lower and upper Darboux sums then oscillatory sum of function f over a partition P is denoted by $\omega(f, P)$ and defined as follows

$$\omega(f, P) = U(f, P) - L(f, P) = \sum_{i=1}^{n} M_i\delta_i - \sum_{i=1}^{n} m_i\delta_i = \sum_{i=1}^{n} (M_i - m_i)\delta_i = \sum_{i=1}^{n} o_i\delta_i, \quad \ldots (5.29)$$

where $o_i = M_i - m_i$

Experiment: Let us do an experiment where we will obtain lower and upper Darboux sums of the function y = f(x) = x + 1 over three partitions $P_1 = \{1, 2, 4\}$, $P_2 = \{1, 2, 3, 4\}$, and $P_3 = \{1, 2, 3, 3.5, 4\}$ all are partitions of the finite and closed interval [1, 4]. Note that partition P_2 is a refinement of the partition P_1 and partition P_3 is refinement of P_2 and hence of P_1 also. After completing this experiment, we will discuss some interesting observations of the experiment.

Let us divide this experiment into three parts as follows.

Part I: Calculation of L(f, P_1) and U(f, P_1)

Here, y = f(x) = x + 1, $P_1 = \{1, 2, 4\}$, so

$I_1 = [1, 2],\ I_2 = [2, 4],\ \delta_1 = 2 - 1 = 1,\ \delta_2 = 4 - 2 = 2,\ m_1 = f(1) = 2,\ m_2 = f(2) = 3,$

$M_1 = f(2) = 3,\ M_2 = f(4) = 5.$

$$\text{So},\ L(f, P_1) = \sum_{i=1}^{2} m_i\delta_i = (2)(1) + (3)(2) = 8 \quad \ldots (5.30)$$

$$U(f, P_1) = \sum_{i=1}^{2} M_i\delta_i = (3)(1) + (5)(2) = 13 \quad \ldots (5.31)$$

Values of (5.30) and (5.31) are visualised in Fig. 5.5 (a) and (b) respectively. Areas of two solid rectangles (L1 and L2) in blue colour in Fig. 5.5 (a) represent two terms (2 and 6) of RHS of (5.30) respectively. While areas of two solid rectangles (U1 and U2) in red colour in Fig. 5.5 (b) represent two terms (3 and 10) of RHS of (5.31) respectively.

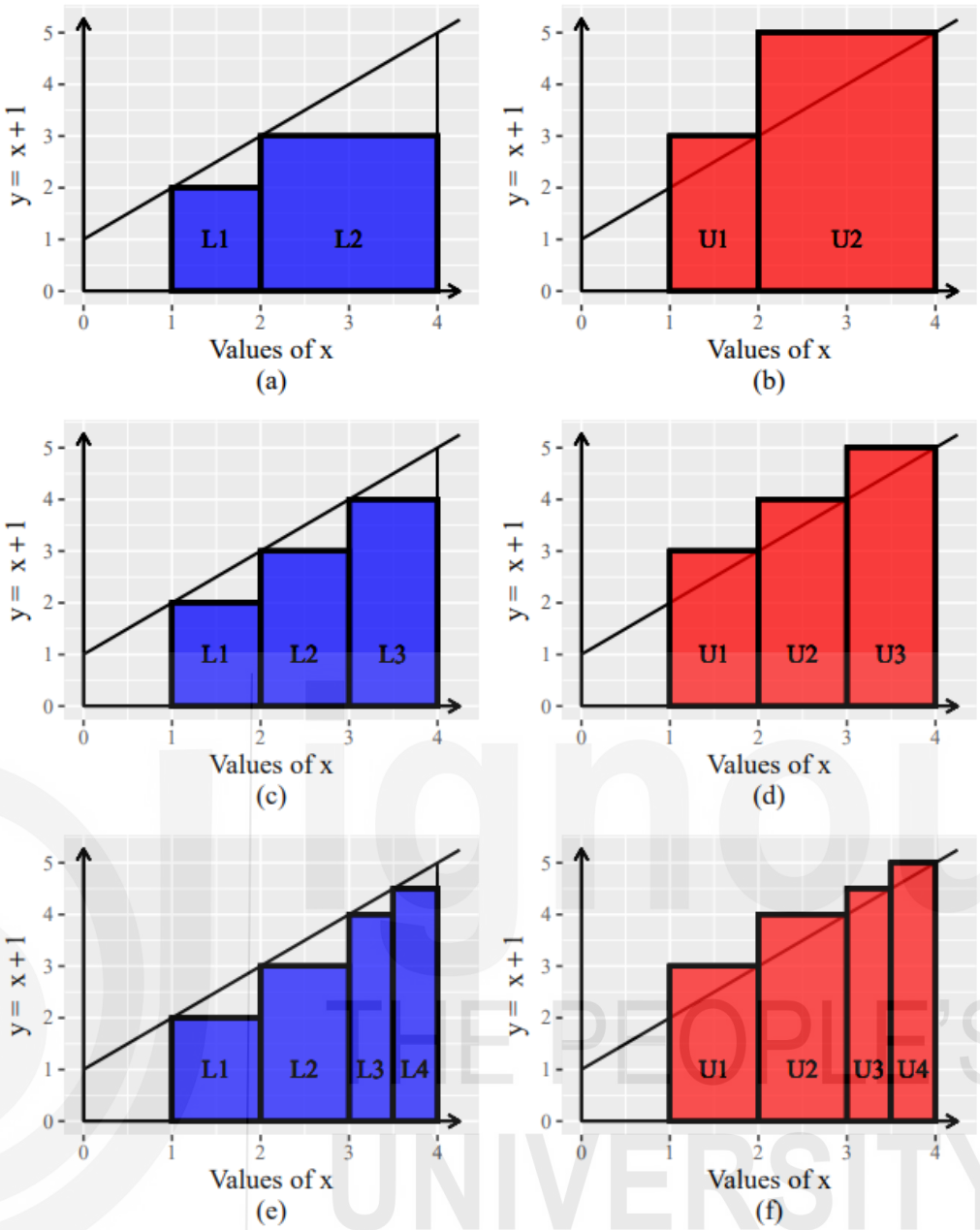


Fig. 5.5: Visualisation of Darboux sums over partition (a) P_1 lower sum (b) P_1 upper sum (c) P_2 lower sum (d) P_2 upper sum (e) P_3 lower sum (f) P_3 upper sum

Part II: Calculation of L(f, P_2) and U(f, P_2)

Here, y = f(x) = x + 1, $P_2 = \{1, 2, 3, 4\}$, so

$$I_1 = [1, 2], I_2 = [2, 3], I_3 = [3, 4], \delta_1 = 2 - 1 = 1, \delta_2 = 3 - 2 = 1, \delta_3 = 4 - 3 = 1,$$

$$m_1 = f(1) = 2, m_2 = f(2) = 3, m_3 = f(3) = 4, M_1 = f(2) = 3, M_2 = f(3) = 4, M_3 = f(4) = 5.$$

So, $$L(f, P_2) = \sum_{i=1}^{3} m_i\delta_i = (2)(1) + (3)(1) + (4)(1) = 9, \qquad \ldots (5.32)$$

$$U(f, P_2) = \sum_{i=1}^{3} M_i\delta_i = (3)(1) + (4)(1) + (5)(1) = 12 \qquad \ldots (5.33)$$

Values of (5.32) and (5.33) are visualised in Fig. 5.5 (c) and (d) respectively. Areas of three solid rectangles (L1, L2 and L3) in blue colour in Fig. 5.5 (c) represent three terms (2, 3 and 4) of RHS of (5.32) respectively. While areas of three solid rectangles (U1, U2 and U3) in red colour in Fig. 5.5 (d) represent

three terms (3, 4 and 5) of RHS of (5.33) respectively.

Part III: Calculation of L(f, P_3) and U(f, P_3)

Here, y = f(x) = x + 1, $P_3 = \{1, 2, 3, 3.5, 4\}$, so

$I_1 = [1, 2], I_2 = [2, 3], I_3 = [3, 3.5], I_4 = [3.5, 4], \delta_1 = 2 - 1 = 1, \delta_2 = 3 - 2 = 1,$
$\delta_3 = 3.5 - 3 = 0.5, \delta_4 = 4 - 3.5 = 0.5, m_1 = f(1) = 2, m_2 = f(2) = 3, m_3 = f(3) = 4,$
$m_4 = f(3.5) = 4.5, M_1 = f(2) = 3, M_2 = f(3) = 4, M_3 = f(3.5) = 4.5, M_4 = f(4) = 5.$

So, $L(f, P_3) = \sum_{i=1}^{4} m_i\delta_i = (2)(1) + (3)(1) + (4)(0.5) + (4.5)(0.5) = 9.25,$ … (5.34)

$U(f, P_3) = \sum_{i=1}^{4} M_i\delta_i = (3)(1) + (4)(1) + (4.5)(0.5) + (5)(0.5) = 11.75$ … (5.35)

Values of (5.34) and (5.35) are visualised in Fig. 5.5 (e) and (f) respectively. Areas of four solid rectangles (L1, L2, L3 and L4) in blue colour in Fig. 5.5 (e) represent four terms (2, 3, 2 and 2.25) of RHS of (5.34) respectively. While areas of four solid rectangles (U1, U2, U3 and U4) in red colour in Fig. 5.5 (f) represent four terms (3, 4, 2.25 and 2.5) of RHS of (5.35) respectively.

Now, as promised let us discuss some interesting observations of the experiment as follows.

- First of all let us calculate the exact area under the curve y = (x) = x + 1 and between the lines x = 1, x = 4 and y = 0 using formula for area of trapezium.
  Required area $= \frac{1}{2}(2+5)(3) = \frac{1}{2}(21) = 10.5 \text{ unit}^2$ ... (5.36)

- **Observation 1**: Note that lower sum under estimate the exact area. It can be seen from Fig. 5.5 (a) and equations (5.30), (5.32) and (5.34) where values of lower sum are $L(f, P_1) = 8, L(f, P_2) = 9, L(f, P_3) = 9.25$ which are < 10.5 (exact value). But important and interesting observation is lower sum is increasing and coming closer and closer to exact value as we add more points in the partition. On comparing Fig. 5.5 (a), (c) and (e) you can observe that as number of points in the partition of finite and closed interval [1, 4] increases then lower sum will also increase and each new point in the partition contributes some new area in the old lower sum. No doubt as we add a new point in the partition then lower sum increase but the amount of new area contributed by new point decrease as we go on increasing points in the partition. This can be seen on comparing values of lower sum given by equations (5.30). (5.32) and (5.34). Now, let us first compare values of lower sum given by equations (5.30) and (5.32). The value of lower sum in (5.32) is 1 unit^2 (8 vs 9) more that the value in (5.30) where partition used in (5.32) has one additional point compare to the partition used to evaluate (5.30). On the other hand, the value of lower sum in (5.34) is 0.25 unit^2 (9 vs 9.25) more that the value in (5.32) where partition used in (5.34) has one additional point compare to the partition used to evaluate (5.32). So, moral of the story is addition of a new point in the partition contributes some additional area in the previous lower sum but in most of the cases this contribution decreases except some cases. The reason for such a behaviour is: initially difference between value of lower sum and value of exact area is more compare to later stages. For example, consider four

more partitions of finite and closed interval [1, 4] given as follows.

$P_4 = \{1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, 2.75, 3, 3.25, 3.5, 3.75, 4\}$,

$P_5 = \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, \ldots, 4\}$,

$P_6 = \{1, 1.01, 1.02, 1.03, 1.04, 1.05, 1.06, 1.07, 1.08, 1.09, 1.10, \ldots, 4\}$,

$P_7 = \{1, 1.001, 1.002, 1.003, 1.004, 1.005, 1.006, \ldots, 4\}$

Now, values of lower sums obtained over the partitions P_1 to P_7 are compared with exact value of the area of 10.5 unit^2 look at the Table 5.1 given as follows. Also, lower and upper sums corresponding to partitions P_4 and P_5 are shown in Fig. 5.6 (a), (c) and (b), (d) respectively.

Table 5.1 Comparison of lower sums in different stages with exact area

Stage	Name of the Partition	Exact Area	Lower Sum	Underestimate in Area	Underestimate in Area in %
1	P_1	10.5	8	2.5	$\frac{2.5}{10.5} \simeq 23.81\%$
2	P_2	10.5	9	1.5	$\frac{1.5}{10.5} \simeq 14.29\%$
3	P_3	10.5	9.25	1.25	$\frac{1.25}{10.5} \simeq 11.90\%$
4	P_4	10.5	10.125	0.375	$\frac{0.375}{10.5} \simeq 3.57\%$
5	P_5	10.5	10.35	0.15	$\frac{1.25}{10.5} \simeq 1.43\%$
6	P_6	10.5	10.485	0.015	$\frac{0.015}{10.5} \simeq 0.14\%$
7	P_7	10.5	10.4985	0.0015	$\frac{0.015}{10.5} \simeq 0.01\%$

Note that in the first stage percentage of the difference between values of exact area and lower sum with exact area was 23.81% (approx.) while in second stage this difference reduces to 14.29% (approx.) and in third stage this difference reduces to 11.90% (approx.) and so on in seventh stage when we have 3001 points in the partition P_7 then value of lower sum is just 0.0015 unit^2 less than the exact area and in percentage it is 0.01% (approx.). Thus, we can say that if given function is Riemann integrable then as number of points in the partition will tend to infinity then lower sum will also tend to exact value of the integral. This limiting value if exists is called **lower integral** and is denoted by

$$\int_1^4 f(x)\,dx = \lim_{n\to\infty} L(f, P) = \lim_{\|P\|\to 0} L(f, P) \qquad \ldots (5.37)$$

whereas n, the number of points in the partition P tends to ∞ then norm of the partition $\|P\| \to 0$.

Observation 2: Note that upper sum overestimates the exact area. It can be seen from Fig. 5.5 (b) and equations (5.31), (5.33) and (5.35) where values of upper sum are $U(f, P_1) = 13$, $U(f, P_2) = 12$, $U(f, P_3) = 11.75$ which are > 10.5 (exact value). But important and interesting observation is upper sum is decreasing and coming closer and closer to exact value as we add

more points in the partition. On comparing Fig. 5.5 (b), (d) and (f) you can observe that as number of points in the partition of finite and closed interval [1, 4] increases then upper sum decreases and each new point in the partition remove some area from the old upper sum. No doubt as we add a new point in the partition then upper sum decreases but the amount of area removed by new point decreases as we go on increasing points in the partition. This can be seen on comparing values of upper sum given by equations (5.31). (5.33) and (5.35). Now, let us first compare values of upper sum given by equations (5.31) and (5.33). The value of upper sum in (5.33) is 1 unit^2 (12 vs 13) less that the value in (5.31) where partition used in (5.33) has one additional point compare to the partition used to evaluate (5.31). On the other hand, the value of upper sum in (5.35) is 0.25 unit^2 (11.75 vs 12) less than that of the value in (5.33) where partition used in (5.35) has one additional point compare to the partition used to evaluate (5.33). So, moral of the story is addition of a new point in the partition remove some area from the previous upper sum and reduces it but in most of the cases this deletion of area decreases except some cases. The reason for such a behaviour is: initially difference between value of upper sum and the value of the exact area is more compare to later stages. For example, values of upper sums obtained over the partitions P_1 to P_7 are compared with exact value of the area of 10.5 unit^2 look at the Table 5.2 given as follows.

Table 5.2 Comparison of upper sums in different stages with exact area

Stage	Name of the Partition	Exact Area	Upper Sum	Overestimate in Area	Overestimate in Area in %
1	P_1	10.5	13	2.5	$\frac{2.5}{10.5} \simeq 23.81\%$
2	P_2	10.5	12	1.5	$\frac{1.5}{10.5} \simeq 14.29\%$
3	P_3	10.5	11.75	1.25	$\frac{1.25}{10.5} \simeq 11.90\%$
4	P_4	10.5	10.875	0.375	$\frac{0.375}{10.5} \simeq 3.57\%$
5	P_5	10.5	10.65	0.15	$\frac{1.25}{10.5} \simeq 1.43\%$
6	P_6	10.5	10.515	0.015	$\frac{0.015}{10.5} \simeq 0.14\%$
7	P_7	10.5	10.5015	0.0015	$\frac{0.015}{10.5} \simeq 0.01\%$

Note that in the first stage percentage of the difference between values of exact area and upper sum was 23.81% (approx.) while in second stage this difference reduces to 14.29% (approx.) and in third stage this difference reduces to 11.90% (approx.) and so on in seventh stage when we have 3001 points in the partition P_7 then value of upper sum is just 0.0015 unit^2 more than the exact area and in percentage it is 0.01% (approx.). Thus, we can say that if given function is Riemann integrable then as number of points in the partition will tend to infinity then upper sum will also tend to exact value of the integral. This limiting value if exists is called **upper**

integral and is denoted by

$$\overline{\int_1^4} f(x)\,dx = \lim_{n\to\infty} U(f, P) = \lim_{\|P\|\to 0} U(f, P) \qquad \ldots (5.38)$$

whereas n, the number of points in the partition P tends to ∞ then norm of the partition $\|P\| \to 0$.

- **Observation 3**: $n \to \infty$ and $\|P\| \to 0$ are equivalent … (5.39)

  because as the number of points in the partition increases then norm of the partition either remain the same if additional point(s) does(do) not fall in the sub-interval which decides norm of the partition and it will decrease if additional point(s) falls(fall) in the sub-interval which decides norm of the partition. But as $n \to \infty$ then length of every sub-interval will approach to zero and so $\|P\| \to 0$. This gives us the way to do examples. You will experience it in Example 1.

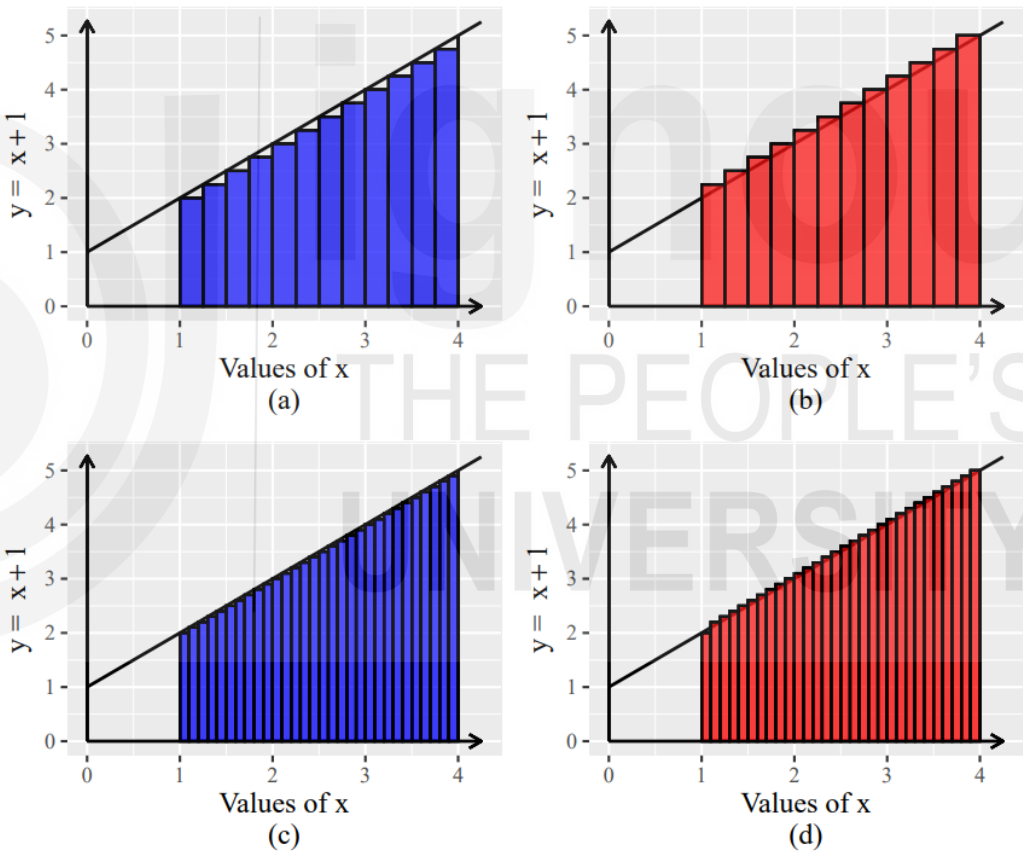


Fig. 5.6: Visualisation of Darboux sums over partition (a) P_4 lower sum (b) P_4 upper sum (c) P_5 lower sum (d) P_5 upper sum

Upper and Lower Riemann Integrals

Before defining lower and upper integral we have to prove a result given as follows.

Result 1: Let [a, b] be a finite and closed interval and P[a, b] denote the set of all partitions over [a, b]. If $f : [a, b] \to \mathbb{R}$ be a bounded function then

$$m(b-a) \le L(f, P) \le U(f, P) \le M(b-a) \quad \forall\, P \in P[a, b] \qquad \ldots (5.40)$$

where m = glb of f and M = lub of f on [a, b].

Proof: Consider the partition $P = \{a = x_0, x_1, x_2, x_3, \ldots, x_n = b\}$ of [a, b], then in usual notations, we have

$I_i = [x_{i-1}, x_i],\ \delta_i = x_i - x_{i-1},\ i = 1, 2, 3, \ldots, n$

Since f is bounded on [a, b] $\Rightarrow$ f is bounded on each sub-interval I_i.

Let $m_i = \text{glb of } f \text{ on } I_i$ and $M_i = \text{lub of } f \text{ on } I_i,\ i = 1, 2, 3, \ldots, n.$

Therefore, m_i, M_i, m and M will always satisfy the following relation

$$m \le m_i \le M_i \le M \ \ \forall\ i,\ i = 1, 2, 3, \ldots, n$$

$$\Rightarrow m\delta_i \le m_i\delta_i \le M_i\delta_i \le M\delta_i \ \ \forall\ i,\ i = 1, 2, 3, \ldots, n\ [\because \delta_i = \text{length of an interval} > 0]$$

$$\Rightarrow \sum_{i=1}^{n} m\delta_i \le \sum_{i=1}^{n} m_i\delta_i \le \sum_{i=1}^{n} M_i\delta_i \le \sum_{i=1}^{n} M\delta_i$$

$$\Rightarrow m\sum_{i=1}^{n} \delta_i \le \sum_{i=1}^{n} m_i\delta_i \le \sum_{i=1}^{n} M_i\delta_i \le M\sum_{i=1}^{n} \delta_i$$

$$\Rightarrow m(b-a) \le L(f, P) \le U(f, P) \le M(b-a) \ \left[\because \sum_{i=1}^{n} \delta_i = b - a\right]$$

Hence, we have proved that

$$m(b-a) \le L(f, P) \le U(f, P) \le M(b-a) \ \ \forall\ P \in P[a, b]$$

Now, we can define lower and upper Riemann integrals as follows.

Let [a, b] be a finite and closed interval and $f : [a, b] \to \mathbb{R}$ be a bounded function then we know that

$$m(b-a) \le L(f, P) \le U(f, P) \le M(b-a) \ \ \forall\ P \in P[a, b] \qquad \ldots (5.41)$$

From (5.41), we have

$$L(f, P) \le M(b-a) \ \ \forall\ P \in P[a, b] \ \text{ and } \ U(f, P) \ge m(b-a) \ \ \forall\ P \in P[a, b]$$

$\Rightarrow$ the set $\{L(f, P)\}_{P \in P[a, b]}$ of all lower sums is bounded above by M(b – a) and therefore has the lub. Similarly, the set $\{U(f, P)\}_{P \in P[a, b]}$ of all upper sums is bounded below by m(b – a) and therefore has the glb.

where m = glb of f and M = lub of f.

Now, **lower Riemann integral** of f on [a, b] is denoted by $\underline{\int_a^b} f(x)\,dx$ and is

defined by $\underline{\int_a^b} f(x)\,dx = \text{lub}\{L(f, P)\}_{P \in P[a, b]}$ … (5.42)

and **upper Riemann integral** of f on [a, b] is denoted by $\overline{\int_a^b} f(x)\,dx$ and is

defined by $\overline{\int_a^b} f(x)\,dx = \text{glb}\{U(f, P)\}_{P \in P[a, b]}$ … (5.43)

So, far we have discussed all the terms required to understand the two definitions of Riemann integration given as follows.

Definition 1: Let [a, b] be a finite and closed interval and $f:[a, b]\to\mathbb{R}$ be a bounded function, then we say that f is **Riemann integrable** if lower Riemann integral and upper Riemann integral are equal. i.e., if

$$\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx \qquad \ldots (5.44)$$

The common value of these integrals is called the Riemann integral of f on [a, b] and is denoted by $\int_a^b f(x)\,dx$.

$$\text{i.e., } \underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx = \int_a^b f(x)\,dx. \qquad \ldots (5.44\ D1)$$

Definition 2: Let [a, b] be a finite and closed interval and $f:[a, b]\to\mathbb{R}$ be a bounded function, then we say that f is **Riemann integrable** if

$$\lim_{\|P\|\to 0}\sum_{i=1}^{n} f(\xi_i)\delta_i \qquad \ldots (5.45)$$

exists and is independent of the choice of sub-interval $[x_{i-1}, x_i]$ and of the point $\xi_i \in [x_{i-1}, x_i]$ … (5.46)

If limit given by (5.45) exists then value of this limit is called the Riemann integral of f on [a, b] and is denoted by $\int_a^b f(x)\,dx$.

$$\text{i.e., } \lim_{\|P\|\to 0}\sum_{i=1}^{n} f(\xi_i)\delta_i = \int_a^b f(x)\,dx. \qquad \ldots (5.46\ D2)$$

Now, let us do an example using both the definitions so that you become familiar with the procedure of evaluating integral using definition of Riemann integration.

Example 1: If $f:[2, 10]\to\mathbb{R}$ be a function defined by $f(x) = 4x + 3,\ x\in[2, 10]$. Show that f is Riemann integrable using definition 1 and definition 2 and $\int_2^{10} f(x)\,dx = 216$.

Solution: In usual notations we are given

a = 2, b = 10, f(x) = 4x + 3, length of the interval [2, 10] is 10 – 2 = 8 = b – a.

For the finite and closed interval [a, b] we form the partition

$$P = \{a = a + 0h,\ a + h,\ a + 2h,\ a + 3h,\ \ldots,\ a + nh = b\},\ h = \frac{b-a}{n} \qquad \ldots (5.47)$$

So, in our case partition will be

$$P = \{2 = 2 + 0h,\ 2 + h,\ 2 + 2h,\ 2 + 3h,\ \ldots,\ 2 + nh = 10\},\ \text{where } h = \frac{10-2}{n} = \frac{8}{n}$$

Therefore, i^{th} sub interval $= [2 + (i-1)h,\ 2 + ih]$, so

$$x_{i-1} = 2 + (i-1)h = 2 + \frac{8(i-1)}{n},\ x_i = 2 + ih = 2 + \frac{8i}{n},\ \text{as } h = \frac{8}{n} \text{ and } \delta_i = h = \frac{8}{n}$$

Now, we proceed to apply definition 1 and definition 2 of Riemann integration as follows.

Using Definition 1

Let $m_i = \text{glb of } f \text{ on } I_i = \left[2+\frac{8(i-1)}{n},\ 2+\frac{8i}{n}\right],\ \ i=1, 2, 3, \ldots, n$

$M_i = \text{lub of } f \text{ on } I_i = \left[2+\frac{8(i-1)}{n},\ 2+\frac{8i}{n}\right],\ \ i=1, 2, 3, \ldots, n$

Since f is an increasing function over the interval [2, 10] and hence over each sub-interval $I_i = \left[2+\frac{8(i-1)}{n},\ 2+\frac{8i}{n}\right],\ \ i=1, 2, 3, \ldots, n,$ therefore

$m_i = f\left(2+\frac{8(i-1)}{n}\right),\ \ M_i = f\left(2+\frac{8i}{n}\right),\ \ i=1, 2, 3, \ldots, n,$

So,

$$L(f, P) = \sum_{i=1}^{n} m_i\,\delta_i = \sum_{i=1}^{n} f\left(2+\frac{8(i-1)}{n}\right)\frac{8}{n} = 8\sum_{i=1}^{n}\left[4\left(2+\frac{8(i-1)}{n}\right)+3\right]\frac{1}{n}$$

$$= 8\sum_{i=1}^{n}\left[\frac{32}{n^2}(i-1)+\frac{11}{n}\right] = \frac{256}{n^2}\sum_{i=1}^{n}(i-1)+\frac{8}{n}\sum_{i=1}^{n}11$$

$$= \frac{256}{n^2}(0+1+2+3+\ldots+n-1)+\frac{8}{n}(11n) = \frac{256}{n^2}\frac{n(n-1)}{2}+88 = 128\left(1-\frac{1}{n}\right)+88$$

$$\left[\because 1+2+3+\ldots+n = \frac{n(n+1)}{2},\ \text{so}\ 1+2+3+\ldots+n-1 = \frac{n(n-1)}{2}\right]$$

$$U(f, P) = \sum_{i=1}^{n} M_i\,\delta_i = \sum_{i=1}^{n} f\left(2+\frac{8i}{n}\right)\frac{8}{n} = 8\sum_{i=1}^{n}\left[4\left(2+\frac{8i}{n}\right)+3\right]\frac{1}{n}$$

$$= 8\sum_{i=1}^{n}\left[\frac{32}{n^2}i+\frac{11}{n}\right] = \frac{256}{n^2}\sum_{i=1}^{n}i+\frac{8}{n}\sum_{i=1}^{n}11$$

$$= \frac{256}{n^2}(1+2+3+\ldots+n)+\frac{8}{n}(11n) = \frac{256}{n^2}\frac{n(n+1)}{2}+88 = 128\left(1+\frac{1}{n}\right)+88$$

By definition of lower Riemann integration refer (5.42), we have

$$\underline{\int_2^{10}} f(x)\,dx = \text{lub}\{L(f, P)\}_{P\in P[2,10]} = \lim_{\|P\|\to 0}\left(128\left(1-\frac{1}{n}\right)+88\right)$$

$$= \lim_{n\to\infty}\left(128\left(1-\frac{1}{n}\right)+88\right) \qquad [\text{Using } (5.39)]$$

$$= 128(1-0)+88 = 216$$

By definition of upper Riemann integration refer (5.43), we have

$$\overline{\int_2^{10}} f(x)\,dx = \text{glb}\{U(f, P)\}_{P\in P[2,10]} = \lim_{\|P\|\to 0}\left(128\left(1+\frac{1}{n}\right)+88\right) = \lim_{n\to\infty}\left(128\left(1+\frac{1}{n}\right)+88\right)$$

$$= 128(1+0)+88 = 216$$

Since $\underline{\int_2^{10}} f(x)\,dx = 216 = \overline{\int_2^{10}} f(x)\,dx,$ so by definition 1 of Riemann integration f is Riemann integrable and $\int_2^{10} f(x)\,dx = 216.$

Using Definition 2:

Let $\xi_i = 2 + \frac{8i}{n}$, $i = 1, 2, 3, \ldots, n$ [Refer (5.46)]

Now,

$$\lim_{\|P\|\to 0}\sum_{i=1}^{n} f(\xi_i)\delta_i = \lim_{\|P\|\to 0}\sum_{i=1}^{n} f\left(2+\frac{8i}{n}\right)\frac{8}{n} \quad \left[\because \xi_i = 2+\frac{8i}{n} \text{ and } \delta_i = \frac{8}{n}\right]$$

$$= 8\lim_{\|P\|\to 0}\sum_{i=1}^{n}\left[4\left(2+\frac{8i}{n}\right)+3\right]\frac{1}{n} = 8\lim_{\|P\|\to 0}\left(\sum_{i=1}^{n}\left[\frac{32}{n^2}i+\frac{11}{n}\right]\right) = \lim_{\|P\|\to 0}\left(\frac{256}{n^2}\sum_{i=1}^{n} i + \frac{8}{n}\sum_{i=1}^{n} 11\right)$$

$$= \lim_{\|P\|\to 0}\left(\frac{256}{n^2}(1+2+3+\ldots+n)+\frac{8}{n}(11n)\right) = \lim_{n\to\infty}\left(\frac{256}{n^2}\frac{n(n+1)}{2}+88\right) \quad \left[\begin{matrix}\text{Using}\\(5.39)\end{matrix}\right]$$

$$= \lim_{n\to\infty}\left(128\left(1+\frac{1}{n}\right)+88\right) = 128(1+0)+88 = 216.$$

Since $\lim_{\|P\|\to 0}\sum_{i=1}^{n} f(\xi_i)\delta_i = 216$ so by definition 2 of Riemann integration f is Riemann integrable and $\int_2^{10} f(x)\,dx = 216$.

Remark 1: From school mathematics you know that

$\int_2^{10} f(x)\,dx = \int_2^{10}(4x+3)\,dx = \left[2x^2+3x\right]_2^{10} = 200+30-8-6 = 230-14 = 216.$ But do not evaluate integral like this if it is given that evaluate using definition 1 or definition 2 of Riemann integration.

Example 2: Give an example of a function which is not Riemann integrable.

Solution: Let us consider the Dirichlet function on the interval [0, 1] given as follows.

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q} \cap [0, 1] \\ 0, & x \notin \mathbb{Q} \cap [0, 1] \end{cases}$$

In usual notations we are given

a = 0, b = 1, f(x) as defined, length of the interval [0, 1] is 1 – 0 = 1 = b – a.

For the finite and closed interval [a, b] we form the partition

$$P = \{a = a+0h,\ a+h,\ a+2h,\ a+3h,\ \ldots,\ a+nh = b\},\ h = \frac{b-a}{n}$$

So, in our case partition will be

$$P = \{0 = 0+0h,\ 0+h,\ 0+2h,\ 0+3h,\ \ldots,\ 0+nh = 1\},\ \text{where } h = \frac{1-0}{n} = \frac{1}{n}$$

$$\Rightarrow P = \{0h,\ h,\ 2h,\ 3h,\ \ldots,\ nh = 1\},\ \text{where } h = \frac{1}{n}$$

Therefore, i^{th} sub interval $= [(i-1)h,\ ih]$, so

$$x_{i-1} = (i-1)h = \frac{(i-1)}{n},\ x_i = ih = \frac{i}{n},\ \text{as } h = \frac{1}{n} \text{ and } \delta_i = h = \frac{1}{n}$$

We claim that given function is not Riemann integrable. To prove our claim, we shall show that value of two limits used in definition 1 of Riemann integration

do not match.

Let $m_i = \text{glb of } f \text{ on } I_i = \left[\frac{(i-1)}{n}, \frac{i}{n}\right], \ i = 1, 2, 3, \ldots, n$

$M_i = \text{lub of } f \text{ on } I_i = \left[\frac{(i-1)}{n}, \frac{i}{n}\right], \ i = 1, 2, 3, \ldots, n$

Since every sub-interval $\left[\frac{(i-1)}{n}, \frac{i}{n}\right], \ i = 1, 2, 3, \ldots, n$ contains rational as well as irrational numbers so, we have

$m_i = 0, \ M_i = 1, \ i = 1, 2, 3, \ldots, n,$

So, $L(f, P) = \sum_{i=1}^{n} m_i \, \delta_i = \sum_{i=1}^{n} (0) \frac{1}{n} = \sum_{i=1}^{n} 0 = 0$

and $U(f, P) = \sum_{i=1}^{n} M_i \, \delta_i = \sum_{i=1}^{n} (1) \frac{1}{n} = \sum_{i=1}^{n} \frac{1}{n} = \frac{1}{n}(n) = 1$

By definition of lower Riemann integration refer (5.42), we have

$$\underline{\int_0^1} f(x)\,dx = \text{lub}\{L(f, P)\}_{P \in P[0,1] \cap \mathbb{Q}} = \lim_{\|P\| \to 0} (0) = 0 \qquad \ldots (5.48)$$

By definition of upper Riemann integration refer (5.43), we have

$$\overline{\int_0^1} f(x)\,dx = \text{glb}\{U(f, P)\}_{P \in P[0,1] \cap \mathbb{Q}} = \lim_{\|P\| \to 0} (1) = \lim_{n \to \infty} (1) = 1 \qquad \ldots (5.49)$$

Since $\underline{\int_0^1} f(x)\,dx \neq \overline{\int_0^1} f(x)\,dx$. Hence, given function is not Riemann integrable.

Before doing the next example let us define a notation.

Notation: We shall denote the set of all Riemann integrable functions over the closed and finite interval [a, b] by R[a, b].

Example 3: If c is a fixed real number and $f : [a, b] \to \mathbb{R}$ is defined by

$f(x) = c, \ x \in [a, b]$, then prove that $f \in R[a, b]$, and $\int_a^b f(x)\,dx = c(b - a)$. (5.50)

Hence, evaluate $\int_{10}^{1000} 100\,dx$.

Solution: In usual notations we are given

a = a, b = b, f(x) = c, length of the interval [a, b] is b – a.

For the finite and closed interval [a, b] we form the partition

$$P = \{a = a + 0h, a + h, a + 2h, a + 3h, \ldots, a + nh = b\}, \ h = \frac{b - a}{n}$$

Therefore, i^{th} sub interval $= [a + (i-1)h, a + ih]$, so

$x_{i-1} = a + (i-1)h, \ x_i = a + ih$, as $h = \frac{b-a}{n}$ and $\delta_i = h = \frac{b-a}{n}$

Since f is a constant function so it will take value c at all points of its domain. Hence, if

$m_i = \text{glb of } f \text{ on } I_i = [a+(i-1)h,\ a+ih],\ \ i = 1, 2, 3, \ldots, n$ and

$M_i = \text{lub of } f \text{ on } I_i = [a+(i-1)h,\ a+ih],\ \ i = 1, 2, 3, \ldots, n$ then

$m_i = c,\ \ M_i = c,\ \ i = 1, 2, 3, \ldots, n,$

So, $L(f, P) = \sum_{i=1}^{n} m_i\,\delta_i = \sum_{i=1}^{n} ch = \sum_{i=1}^{n} c\left(\frac{b-a}{n}\right) = c\left(\frac{b-a}{n}\right)n = c(b-a)$

and $U(f, P) = \sum_{i=1}^{n} M_i\,\delta_i = \sum_{i=1}^{n} ch = \sum_{i=1}^{n} c\left(\frac{b-a}{n}\right) = c\left(\frac{b-a}{n}\right)n = c(b-a)$

By definition of lower Riemann integration refer (5.42), we have

$$\underline{\int_a^b} f(x)\,dx = \text{lub}\{L(f, P)\}_{P\in P[a,b]} = \lim_{\|P\|\to 0}\left(c(b-a)\right) = c(b-a) \qquad \ldots (5.51)$$

By definition of upper Riemann integration refer (5.43), we have

$$\overline{\int_a^b} f(x)\,dx = \text{glb}\{U(f, P)\}_{P\in P[a,b]} = \lim_{\|P\|\to 0}\left(c(b-a)\right) = c(b-a) \qquad \ldots (5.52)$$

Since due to (5.51) and (5.52) $\underline{\int_a^b} f(x)\,dx = c(b-a) = \overline{\int_a^b} f(x)\,dx$. Hence, given function is Riemann integrable and $\int_a^b f(x)\,dx = c(b-a)$.

Comparing $\int_{10}^{1000} 100\,dx$ with $\int_a^b c\,dx$, we get a = 10, b = 1000, c = 100, hence

$\int_{10}^{1000} 100\,dx = 100(1000-10) = 100(990) = 99000.$

Example 4: If [a, b] is finite and closed interval and $f : [a, b] \to \mathbb{R}$ is defined by $f(x) = x,\ \ x \in [a, b]$, i.e., f is an identity function then prove that $f \in R[a, b]$.

Solution: In usual notations we are given

a = a, b = b, f(x) = x, length of the interval [a, b] is b – a.

For the finite and closed interval [a, b] we form the partition

$P = \{a = a+0h,\ a+h,\ a+2h,\ a+3h,\ \ldots,\ a+nh = b\},\ h = \frac{b-a}{n}$

Therefore, i^{th} sub interval $= [a+(i-1)h,\ a+ih]$, so

$x_{i-1} = a+(i-1)h,\ \ x_i = a+ih,$ as $h = \frac{b-a}{n}$ and $\delta_i = h = \frac{b-a}{n}$

Since f is an identity function so it will take value equal to its input from the domain. Hence, if

$m_i = \text{glb of } f \text{ on } I_i = [a+(i-1)h,\ a+ih],\ \ i = 1, 2, 3, \ldots, n$ and

$M_i = \text{lub of } f \text{ on } I_i = [a+(i-1)h,\ a+ih],\ \ i = 1, 2, 3, \ldots, n$ then

$m_i = a + (i-1)h, \quad M_i = a + ih, \quad i = 1, 2, 3, \ldots, n,$

So,

$$L(f, P) = \sum_{i=1}^{n} m_i\,\delta_i = \sum_{i=1}^{n}\big(a + (i-1)h\big)h = \sum_{i=1}^{n}\big(ah + (i-1)h^2\big) = \sum_{i=1}^{n} ah + \sum_{i=1}^{n}(i-1)h^2$$

$$= nah + h^2(0 + 1 + 2 + 3 + \ldots + n - 1) = anh + h^2(1 + 2 + 3 + \ldots + n - 1)$$

$$= anh + h^2\,\frac{n(n-1)}{2} \qquad \left[\begin{array}{l} \because 1 + 2 + 3 + \ldots + n = \dfrac{n(n+1)}{2}, \text{ so} \\ 1 + 2 + 3 + \ldots + n - 1 = \dfrac{n(n-1)}{2} \end{array}\right]$$

$$= a(b-a) + \left(\frac{b-a}{n}\right)^2 \frac{n(n-1)}{2} \qquad \left[\because h = \frac{b-a}{n}\right]$$

$$\Rightarrow L(f, P) = a(b-a) + \frac{1}{2}(b-a)^2\left(1 - \frac{1}{n}\right)$$

$$U(f, P) = \sum_{i=1}^{n} M_i\,\delta_i = \sum_{i=1}^{n}(a + ih)h = \sum_{i=1}^{n}\big(ah + ih^2\big) = \sum_{i=1}^{n} ah + \sum_{i=1}^{n} ih^2$$

$$= nah + h^2(1 + 2 + 3 + \ldots + n)$$

$$= anh + h^2\,\frac{n(n+1)}{2} \qquad \left[\because 1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}\right]$$

$$= a(b-a) + \left(\frac{b-a}{n}\right)^2 \frac{n(n+1)}{2} \qquad [\text{Using } (5.52)]$$

$$\Rightarrow U(f, P) = a(b-a) + \frac{1}{2}(b-a)^2\left(1 + \frac{1}{n}\right)$$

By definition of lower Riemann integration refer (5.42), we have

$$\underline{\int_a^b} f(x)\,dx = \text{lub}\{L(f, P)\}_{P \in P[a, b]} = \lim_{\|P\| \to 0}\left(a(b-a) + \frac{1}{2}(b-a)^2\left(1 - \frac{1}{n}\right)\right)$$

$$= \lim_{n \to \infty}\left(a(b-a) + \frac{1}{2}(b-a)^2\left(1 - \frac{1}{n}\right)\right) \qquad [\text{Using } (5.39)]$$

$$= a(b-a) + \frac{1}{2}(b-a)^2(1-0) = ab - a^2 + \frac{1}{2}b^2 + \frac{1}{2}a^2 - ab = \frac{1}{2}\left(b^2 - a^2\right)$$

By definition of upper Riemann integration refer (5.43), we have

$$\overline{\int_a^b} f(x)\,dx = \text{glb}\{U(f, P)\}_{P \in P[a, b]} = \lim_{\|P\| \to 0}\left(a(b-a) + \frac{1}{2}(b-a)^2\left(1 + \frac{1}{n}\right)\right)$$

$$= \lim_{n \to \infty}\left(a(b-a) + \frac{1}{2}(b-a)^2\left(1 + \frac{1}{n}\right)\right) \qquad \left[\because h = \frac{b-a}{n}\right]$$

$$= a(b-a) + \frac{1}{2}(b-a)^2(1+0) = ab - a^2 + \frac{1}{2}b^2 + \frac{1}{2}a^2 - ab = \frac{1}{2}\left(b^2 - a^2\right)$$

Since $\underline{\int_a^b} f(x)\,dx = \frac{1}{2}\left(b^2 - a^2\right) = \overline{\int_a^b} f(x)\,dx$. Hence, given function is Riemann integrable and $\int_a^b f(x)\,dx = \frac{1}{2}\left(b^2 - a^2\right)$.

Now, you can try the following Self-Assessment Question.

SAQ 1

If $f:[1, 3] \to \mathbb{R}$ be a function defined by $f(x) = x^2 + 1,\ x \in [1, 3]$, then using definition 1 and definition 2 show that f is Riemann integrable and $\int_1^3 f(x)\,dx = \frac{32}{3}$.

5.4 PROPERTIES OF RIEMANN INTEGRATION

In the previous section you have understood two definitions of Riemann integration and learnt to evaluate integral using them. In the same section we have proved that constant and identity functions are Riemann integrable. There is a long list of functions which are Riemann integrable. But due to scope of the course we did not discuss their proof. For example, every continuous function is Riemann integrable, we have stated this as a theorem refer (5.63). You know that polynomial, exponential, logarithm, sine, cosine, etc. all are continuous in their domain and hence Riemann integrable. Properties of Riemann integration help in creating more Riemann integrable functions from given Riemann integrable functions. In this section we will list some properties of Riemann integration. Proofs of these properties are beyond the scope of this course. Before listing these properties let define a notation. We will denote the set of all Riemann integrable function on the finite and closed interval [a, b] by R[a, b].

Property 1: Let $f, g \in R[a, b]$ such that $f(x) \le g(x)\ \ \forall\ x \in [a, b]$ then

$$\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx \qquad \ldots (5.53)$$

Also, if $f(x) \ge 0\ \ \forall\ x \in [a, b]$ then $\int_a^b f(x)\,dx \ge 0.$ … (5.54)

Property 2: Let $f \in R[a, b]$ and c is any real number then

$$\int_a^b (cf)(x)\,dx = c\int_a^b f(x)\,dx \qquad \ldots (5.55)$$

Property 3: Let $f, g \in R[a, b]$ and $\alpha, \beta \in \mathbb{R}$, then

(a) $f + g \in R[a, b]$ and $\int_a^b (f+g)(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$ … (5.56)

(b) $f - g \in R[a, b]$ and $\int_a^b (f-g)(x)\,dx = \int_a^b f(x)\,dx - \int_a^b g(x)\,dx$ … (5.57)

(c) $\alpha f + \beta g \in R[a, b]$ and $\int_a^b (\alpha f + \beta g)(x)\,dx = \alpha\int_a^b f(x)\,dx + \beta\int_a^b g(x)\,dx$ … (5.58)

(d) $fg \in R[a, b]$ … (5.59)

(e) $f^2 \in R[a, b]$ … (5.60)

(f) $|f| \in R[a, b]$ and $\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f|(x)\,dx$ … (5.61)

Property 4: For some $c \in [a, b]$, $f \in R[a, b]$, if and only if $f \in R[a, c]$ and $f \in R[c, b]$. Also, $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$ … (5.62)

We will end this section by proving four important theorems (Theorems 4 to 7) which will not only give you inside of calculus, connection between differentiation and integration but will also help you in understanding many concepts of probability theory of continuous world. But to prove these four theorems we need three more theorems. Due to the scope of this course, we are just stating these three theorems. After stating these three theorems we will state and prove four theorems as mentioned earlier. So, let us start discussion of last four results of this unit.

Theorem 1: Let [a, b] be a finite and closed interval and $f : [a, b] \to \mathbb{R}$ be a continuous function then f is integrable or $f \in R[a, b]$. … (5.63)

Theorem 2: If a function is continuous on a finite and closed interval [a, b] then f is bounded. … (5.64)

Theorem 3: Let [a, b] be a finite and closed interval and $f : [a, b] \to \mathbb{R}$ be a bounded and integrable function on [a, b] then $|f|$ is also bounded and integrable function on [a, b]. Moreover, $\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f|(x)\,dx$. … (5.65)

Theorem 4: Let [a, b] be a finite and closed interval and $f : [a, b] \to \mathbb{R}$ be a continuous function and $F(x) = \int_a^x f(t)\,dt$, where $x \in [a, b]$, then F is continuous on [a, b]. … (5.66)

Proof: We are given that $F(x) = \int_a^x f(t)\,dt$ where $x \in [a, b]$ … (5.67)

It is a function of x. As x increases F(x) also increases refer Fig. 5.7 (a) to (d).

Also, it is given that f is a continuous function on finite and closed interval [a, b].

$\Rightarrow$ f is a bounded function on [a, b]. … (5.68)

$\Rightarrow \exists$ a positive real number k such that

$|f(t)| \le k \qquad \forall\ t \in [a, b]$ … (5.69)

Let $x, y \in [a, b]$ such that $a \le x < y \le b$ then due to (5.67), we have

$$F(x) = \int_a^x f(t)\,dt \text{ and } F(y) = \int_a^y f(t)\,dt$$

$$F(y) - F(x) = \int_a^y f(t)\,dt - \int_a^x f(t)\,dt$$

$$= \int_a^x f(t)\,dt + \int_x^y f(t)\,dt - \int_a^x f(t)\,dt \quad [\because x < y]$$

$$= \int_x^y f(t)\,dt$$

$$\therefore\ |F(y)-F(x)| = \left|\int_x^y f(t)\,dt\right|$$

$$\leq \int_x^y |f(t)|\,dt \qquad \left[\begin{array}{l}\because \text{ If } f \text{ is bounded and integrable on } [a, b] \text{ then} \\ |f| \text{ is also bounded and integrable and} \\ \left|\int_a^b f(t)\,dt\right| \leq \int_a^b |f(t)|\,dt\end{array}\right]$$

$$\therefore\ |F(y)-F(x)| \leq \int_x^y k\,|dt| \qquad [\text{Using } (5.69)]$$

$$= k\int_x^y |dt| = k|y-x| < \varepsilon \quad \text{whenever } |y-x| < \frac{\varepsilon}{k}$$

$$\Rightarrow\ |F(y)-F(x)| < \varepsilon \quad \text{whenever } |y-x| < \frac{\varepsilon}{k}$$

$\Rightarrow$ F is uniformly continuous on [a, b]. [Refer (2.35)]

Hence, F is continuous on [a, b]. $[\because \text{Uniform continuity} \Rightarrow \text{continuity}]$

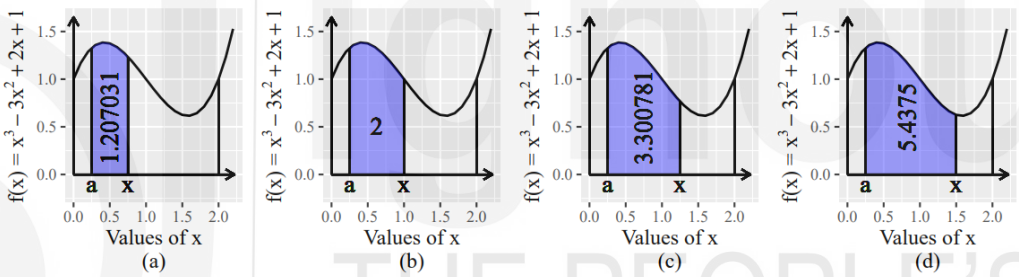


Fig. 5.7: Visualisation of F(x) increases as x increases (a) to (d)

Theorem 5: Let [a, b] be a finite and closed interval and $f : [a, b] \to \mathbb{R}$ be a continuous function then there exists $c \in [a, b]$, such that $\int_a^b f(x)\,dx = (b-a)f(c)$. … (5.70)

Proof: It is given that f is a continuous function on finite and closed interval [a, b]. … (5.71)

$\Rightarrow$ f is a bounded function on [a, b]. … (5.72)

Also, continuity of f implies f is Riemann integrable. … (5.73)

Due to (5.72) and (5.40), we have

$$m(b-a) \leq L(f, P) \leq U(f, P) \leq M(b-a) \quad \forall\ P \in P[a, b] \qquad \ldots (5.74)$$

We know that lower and upper integrals are the lub and glb respectively of the lower and upper sums, therefore

$$L(f, P) \leq \underline{\int_a^b} f(x)\,dx = \int_a^b f(x)\,dx \qquad [\text{Using } (5.63) \text{ and } (5.71)] \qquad \ldots (5.75)$$

$$U(f, P) \geq \overline{\int_a^b} f(x)\,dx = \int_a^b f(x)\,dx \qquad [\text{Using } (5.63) \text{ and } (5.71)] \qquad \ldots (5.76)$$

Combining (5.74), (5.75) and (5.76), we have

$$m(b-a) \le \int_a^b f(x)\,dx \le M(b-a)$$

$$\Rightarrow m \le \frac{1}{(b-a)}\int_a^b f(x)\,dx \le M$$

$$\Rightarrow m \le \lambda \le M, \quad \text{where} \quad \lambda = \frac{1}{(b-a)}\int_a^b f(x)\,dx$$

Or $\dfrac{1}{(b-a)}\displaystyle\int_a^b f(x)\,dx = \lambda, \quad m \le \lambda \le M$

$$\Rightarrow \int_a^b f(x)\,dx = \lambda(b-a), \quad \lambda \in [m, M] \qquad \ldots (5.77)$$

Since f is continuous on [a, b], so it will attain every value between its bounds m and M.

$\therefore \lambda \in [m, M] \Rightarrow \exists$ a number $c \in [a, b]$ such that

$$f(c) = \lambda \qquad \ldots (5.78)$$

Using (5.78) in (5.77), we get

$$\int_a^b f(x)\,dx = (b-a)f(c). \text{ where } c \in [a, b]$$

Remark 2: If we replace b by a + h in above integral then we have

$$\int_a^{a+h} f(x)\,dx = h\,f(a+\theta h), \text{ where } a+\theta h \in [a, a+h], \ 0 \le \theta \le 1 \qquad \ldots (5.79)$$

Theorem 6: Let [a, b] be a finite and closed interval and $f : [a, b] \to \mathbb{R}$ be a continuous function and $F(x) = \displaystyle\int_a^x f(t)\,dt$, where $x \in [a, b]$, then F is differentiable on [a, b] and $F' = f$. … (5.80)

Proof: It is given that f is a continuous function on finite and closed interval [a, b]. … (5.81)

Also, $F(x) = \displaystyle\int_a^x f(t)\,dt, \ x \in [a, b]$ … (5.82)

Let x_0 be an arbitrary point of [a, b] and h be a small positive real number such that $x_0 + h \in [a, b]$. So, due to (5.82), we have

$$F(x_0) = \int_a^{x_0} f(t)\,dt \quad \text{and} \quad F(x_0+h) = \int_a^{x_0+h} f(t)\,dt$$

$$\therefore \ F(x_0+h) - F(x_0) = \int_a^{x_0+h} f(t)\,dt - \int_a^{x_0} f(t)\,dt$$

$$= \int_a^{x_0} f(t)\,dt + \int_{x_0}^{x_0+h} f(t)\,dt - \int_a^{x_0} f(t)\,dt \quad \left[\text{Using (5.62) in the first integral}\right]$$

$$= \int_{x_0}^{x_0+h} f(t)\,dt$$

$$= hf(x_0 + \theta h), \text{ where } 0 \le \theta \le 1 \quad [\text{Using } (5.79)]$$

$$\Rightarrow \frac{F(x_0+h) - F(x_0)}{h} = f(x_0 + \theta h), \text{ where } 0 \le \theta \le 1$$

$$\Rightarrow \lim_{h\to 0} \frac{F(x_0+h) - F(x_0)}{h} = \lim_{h\to 0} f(x_0 + \theta h), \text{ where } 0 \le \theta \le 1$$

$$\Rightarrow F'(x_0) = f(x_0), \; x_0 \in [a, b]$$

Since x_0 is an arbitrary point of [a, b] so F is differentiable on [a, b] and $F' = f$.

Theorem 7: (**Fundamental Theorem of Calculus**): Let [a, b] be a finite and closed interval and $f : [a, b] \to \mathbb{R}$ be Riemann integrable (or we can simply say integrable) on [a, b] and F is antiderivative of the function f then

$$\int_a^b f(x)\,dx = F(b) - F(a). \qquad \ldots (5.83)$$

Proof: Since F is antiderivative of f on [a, b], therefore

$$F'(x) = f(x) \;\; \forall \; x \in [a, b] \qquad \ldots (5.84)$$

Consider the partition $P = \{a = x_0, x_1, x_2, x_3, \ldots, x_n = b\}$ of [a, b], then in usual notations, we have

$$I_i = [x_{i-1}, x_i], \; \delta_i = x_i - x_{i-1}, \; i = 1, 2, 3, \ldots, n \qquad \ldots (5.85)$$

Since F is differentiable on [a, b] $\Rightarrow$ F is continuous on [a, b].

$\Rightarrow$ F is continuous on each sub-interval $I_i = [x_{i-1}, x_i]$, $i = 1, 2, 3, \ldots, n$

Applying Lagrange's mean value theorem to F on each sub-interval $I_i = [x_{i-1}, x_i]$, $i = 1, 2, 3, \ldots, n$ we have

$$F(x_i) - F(x_{i-1}) = (x_i - x_{i-1})F'(\xi_i) \;\; \xi_i \in [x_{i-1}, x_i]$$

$$= f(\xi_i)\delta_i \qquad [\text{Using } (5.84) \text{ and } (5.85)]$$

or $f(\xi_i)\delta_i = F(x_i) - F(x_{i-1})$

$$\Rightarrow \sum_{i=1}^{n} f(\xi_i)\delta_i = \sum_{i=1}^{n} \left(F(x_i) - F(x_{i-1})\right)$$

$$= \left(F(x_1) - F(x_0)\right) + \left(F(x_2) - F(x_1)\right) + \left(F(x_3) - F(x_2)\right) + \ldots + \left(F(x_n) - F(x_{n-1})\right)$$

$$\Rightarrow \sum_{i=1}^{n} f(\xi_i)\delta_i = F(x_n) - F(x_0) \quad \text{All other terms in RHS cancel out in pairs}$$

$$\Rightarrow \sum_{i=1}^{n} f(\xi_i)\delta_i = F(b) - F(a) \qquad [\because x_0 = a, \; x_n = b]$$

$$\Rightarrow \lim_{\|P\|\to 0} \left(\sum_{i=1}^{n} f(\xi_i)\delta_i \right) = \lim_{\|P\|\to 0} \left(F(b) - F(a)\right)$$

$$\Rightarrow \int_a^b f(x)\,dx = F(b) - F(a) \qquad [\text{Using } (5.46 \text{ D2}) \text{ in LHS}]$$

Now, you can try the following Self Assessment Question.

SAQ 2

Find $\frac{d}{dx}(F(x))$, where $F(x) = \int_5^x e^{t^2} dt$.

5.5 SUMMARY

A brief summary of what we have covered in this unit is given as follows:

- Summation notation fails to work in the case of uncountable numbers. So, what is the solution? Solution is **integration**.
- Definite integration gives area bounded by four things (i) under the curve of the function to which we are integrating (ii) equation corresponding to the lower limit of integration (iii) equation corresponding to the upper limit of integration (iv) equation corresponding to the axis of integration.
- If I = [a, b] be a finite and closed interval then by a **partition P of I**, we mean a finite ordered set $\{x_0, x_1, x_2, x_3, \ldots, x_n\}$ of points of [a, b] such that

  $a = x_0 < x_1 < x_2 < x_3 < \ldots < x_{n-1} < x_n = b.$

  The n + 1 points $x_0, x_1, x_2, x_3, \ldots, x_n$ are called partition points of P and the n sub-intervals
  $I_1 = [x_0, x_1], I_2 = [x_1, x_2], I_3 = [x_2, x_3], \ldots, I_i = [x_{i-1}, x_i], \ldots, I_n = [x_{n-1}, x_n]$
  determined by n + 1 points of P are called the segments of the partition P.
- The norm of the partition P is denoted by $\|P\|$ and defined as follows.

  $\|P\| = \max\{x_i - x_{i-1} : i = 1, 2, 3, 4, \ldots, n\}$
- If I = [a, b] be a finite and closed interval and P_1 be a partition of [a, b], then another partition P_2 of [a, b] is called refinement of the partition P_1 if $P_1 \subset P_2$.
- **Lower Darboux sum** of f over the partition P is denoted by $L(f, P)$ and given by $L(f, P) = \sum_{i=1}^{n} m_i \delta_i$, where $m_i = \text{glb of } f \text{ on } [a, b]$
- **Upper Darboux sum** of f over the partition P is denoted by $U(f, P)$ and given by $U(f, P) = \sum_{i=1}^{n} M_i \delta_i$. where $M_i = \text{lub of } f \text{ on } [a, b]$
- **Oscillatory sum** of f over the partition P is denoted by $\omega(f, P)$ and given by $\omega(f, P) = \sum_{i=1}^{n} (M_i - m_i)\delta_i$.
- On addition of a new point in a given partition it contributes some additional area in the previous lower sum. So, addition of a new point in a given partition increases its lower sum.
- On addition of a new point in a given partition it deletes some area from the

previous upper sum. So, addition of a new point in a given partition decreases its upper sum.

- **Lower Riemann integral** of f on [a, b] is denoted by $\underline{\int_a^b} f(x)\,dx$ and is defined by $\underline{\int_a^b} f(x)\,dx = \text{lub}\{L(f, P)\}_{P \in P[a, b]}$
- **Upper Riemann integral** of f on [a, b] is denoted by $\overline{\int_a^b} f(x)\,dx$ and is defined by $\overline{\int_a^b} f(x)\,dx = \text{glb}\{U(f, P)\}_{P \in P[a, b]}$
- **Definition 1**: We say that a function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if lower Riemann integral and upper Riemann integral are equal. i.e., if

  $\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx$
- **Definition 2**: We say that a function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if

  $\lim_{\|P\| \to 0} \sum_{i=1}^{n} f(\xi_i)\delta_i$ exists and is independent of the choice of sub-interval $[x_{i-1}, x_i]$ and of the point $\xi_i \in [x_{i-1}, x_i]$. If this limit exists then value of this limit is called the Riemann integral of f on [a, b] and is denoted by $\int_a^b f(x)\,dx$.
- We have also listed some properties of Riemann integration. Some theorems on Riemann integration are also discussed in this unit.

5.6 TERMINAL QUESTIONS

1. If $f : [0, 4] \to \mathbb{R}$ be a function defined by $f(x) = x^3 + 7,\ x \in [0, 4]$. Show that f is Riemann integrable using definition 1 and definition 2 and $\int_0^4 f(x)\,dx = 92$.

2. Without integrating, evaluate $\int_5^{25} 30\,dx$.

3. If $f : [0, 0.5] \to \mathbb{R}$ be a function defined by $f(x) = 2x + 3,\ x \in [0, 0.5]$ and let P = {0, 0.1, 0.4, 0.5} be the partition of [0, 0.5] then find $L(f, P)$ and $U(f, P)$.

5.7 SOLUTIONS/ANSWERS

Self-Assessment Questions (SAQs)

1. In usual notations we are given

   a = 1, b = 3, f(x) = x^2 + 1, length of the interval [1, 3] is 3 – 1 = 2 = b – a.

   For the finite and closed interval [a, b] we form the partition

$$P = \{a = a + 0h,\ a + h,\ a + 2h,\ a + 3h,\ \ldots,\ a + nh = b\},\ h = \frac{b-a}{n}$$

So, in our case partition will be

$$P = \{1 + 0h,\ 1 + h,\ 1 + 2h,\ 1 + 3h,\ \ldots,\ 1 + nh = 3\},\ \text{ where } h = \frac{3-1}{n} = \frac{2}{n}$$

Therefore, i^{th} sub interval $= [1 + (i-1)h,\ 1 + ih]$, so

$$x_{i-1} = 1 + (i-1)h = 1 + \frac{2(i-1)}{n},\ x_i = 1 + ih = 1 + \frac{2i}{n},\ \text{ as } h = \frac{2}{n} \text{ and } \delta_i = h = \frac{2}{n}$$

Now, we proceed to apply definition 1 and definition 2 of Riemann integration as follows.

Using Definition 1

$$\text{Let } m_i = \text{glb of } f \text{ on } I_i = \left[1 + \frac{2(i-1)}{n},\ 1 + \frac{2i}{n}\right],\ i = 1, 2, 3, \ldots, n$$

$$M_i = \text{lub of } f \text{ on } I_i = \left[1 + \frac{2(i-1)}{n},\ 1 + \frac{2i}{n}\right],\ i = 1, 2, 3, \ldots, n$$

Since f is an increasing function over the interval [1, 3] and hence over each sub-interval $I_i = \left[1 + \frac{2(i-1)}{n},\ 1 + \frac{2i}{n}\right],\ i = 1, 2, 3, \ldots, n,$ therefore

$$m_i = f\left(1 + \frac{2(i-1)}{n}\right),\ M_i = f\left(1 + \frac{2i}{n}\right),\ i = 1, 2, 3, \ldots, n,$$

So,

$$L(f, P) = \sum_{i=1}^{n} m_i\,\delta_i = \sum_{i=1}^{n} f\left(1 + \frac{2(i-1)}{n}\right)\frac{2}{n} = 2\sum_{i=1}^{n}\left[\left(1 + \frac{2(i-1)}{n}\right)^2 + 1\right]\frac{1}{n}$$

$$= 2\sum_{i=1}^{n}\left[\left(1 + \frac{2^2(i-1)^2}{n^2} + 4\frac{i-1}{n}\right) + 1\right]\frac{1}{n} = \frac{8}{n^3}\sum_{i=1}^{n}(i-1)^2 + \frac{8}{n^2}\sum_{i=1}^{n}(i-1) + \frac{2}{n}\sum_{i=1}^{n}2$$

$$= \frac{8}{n^3}\left(0^2 + 1^2 + 2^2 + \ldots + (n-1)^2\right) + \frac{8}{n^2}(0 + 1 + 2 + \ldots + n - 1) + \frac{2}{n}(2n)$$

$$= \frac{8}{n^3}\left(1^2 + 2^2 + 3^2 + \ldots + (n-1)^2\right) + \frac{8}{n^2}(1 + 2 + 3 + \ldots + n - 1) + 4$$

$$= \frac{8}{n^3}\frac{n(n-1)(2n-1)}{6} + \frac{8}{n^2}\frac{n(n-1)}{2} + 4$$

$$\left[\begin{array}{l} \because 1 + 2 + 3 + \ldots + n = \dfrac{n(n+1)}{2},\ \text{ so } 1 + 2 + 3 + \ldots + n - 1 = \dfrac{n(n-1)}{2} \\ 1^2 + 2^2 + 3^2 + \ldots + n^2 = \dfrac{n(n+1)(2n+1)}{6},\ \text{ so} \\ 1^2 + 2^2 + 3^2 + \ldots + (n-1)^2 = \dfrac{n(n-1)(2n-1)}{6} \end{array}\right]$$

$$\Rightarrow L(f, P) = \frac{4}{3}\left(1 - \frac{1}{n}\right)\left(2 - \frac{1}{n}\right) + 4\left(1 - \frac{1}{n}\right) + 4$$

$$U(f, P) = \sum_{i=1}^{n} M_i\,\delta_i = \sum_{i=1}^{n} f\left(1 + \frac{2i)}{n}\right)\frac{2}{n} = 2\sum_{i=1}^{n}\left[\left(1 + \frac{2i}{n}\right)^2 + 1\right]\frac{1}{n}$$

$$= 2\sum_{i=1}^{n}\left[\left(1+\frac{2^2 i^2}{n^2}+4\frac{i}{n}\right)+1\right]\frac{1}{n} = \frac{8}{n^3}\sum_{i=1}^{n} i^2 + \frac{8}{n^2}\sum_{i=1}^{n} i + \frac{2}{n}\sum_{i=1}^{n} 2$$

$$= \frac{8}{n^3}\left(1^2+2^2+3^2+\ldots+n^2\right)+\frac{8}{n^2}(1+2+3+\ldots+n)+\frac{2}{n}(2n)$$

$$= \frac{8}{n^3}\frac{n(n+1)(2n+1)}{6}+\frac{8}{n^2}\frac{n(n+1)}{2}+4$$

$$\Rightarrow U(f, P) = \frac{4}{3}\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right)+4\left(1+\frac{1}{n}\right)+4$$

By definition of lower Riemann integration refer (5.42), we have

$$\underline{\int_1^3} f(x)\,dx = \text{lub}\{L(f, P)\}_{P\in P[1,3]} = \lim_{\|P\|\to 0}\left(\frac{4}{3}\left(1-\frac{1}{n}\right)\left(2-\frac{1}{n}\right)+4\left(1-\frac{1}{n}\right)+4\right)$$

$$= \lim_{n\to\infty}\left(\frac{4}{3}\left(1-\frac{1}{n}\right)\left(2-\frac{1}{n}\right)+4\left(1-\frac{1}{n}\right)+4\right) \qquad [\text{Using } (5.39)]$$

$$= \frac{4}{3}(1-0)(2-0)+4(1-0)+4 = \frac{8}{3}+8 = \frac{32}{3}$$

By definition of upper Riemann integration refer (5.43), we have

$$\overline{\int_1^3} f(x)\,dx = \text{lub}\{L(f, P)\}_{P\in P[1,3]} = \lim_{\|P\|\to 0}\left(\frac{4}{3}\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right)+4\left(1+\frac{1}{n}\right)+4\right)$$

$$= \lim_{n\to\infty}\left(\frac{4}{3}\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right)+4\left(1+\frac{1}{n}\right)+4\right) \qquad [\text{Using } (5.39)]$$

$$= \frac{4}{3}(1+0)(2+0)+4(1+0)+4 = \frac{8}{3}+8 = \frac{32}{3}$$

Since $\underline{\int_1^3} f(x)\,dx = \frac{32}{3} = \overline{\int_1^3} f(x)\,dx$, so by definition 1 of Riemann integration f is Riemann integrable and $\int_1^3 f(x)\,dx = \frac{32}{3}$.

Using Definition 2

Let $\xi_i = 1+\frac{2i}{n}$, $i = 1, 2, 3, \ldots, n$ [Refer (5.46)]

Now,

$$\lim_{\|P\|\to 0}\sum_{i=1}^{n} f(\xi_i)\delta_i = \lim_{\|P\|\to 0}\sum_{i=1}^{n} f\left(1+\frac{2i}{n}\right)\frac{2}{n} \qquad \left[\because \xi_i = 1+\frac{2i}{n} \text{ and } \delta_i = \frac{2}{n}\right]$$

$$= 2\lim_{\|P\|\to 0}\sum_{i=1}^{n}\left[\left(1+\frac{2i}{n}\right)^2+1\right]\frac{1}{n} = 2\lim_{\|P\|\to 0}\sum_{i=1}^{n}\left[\left(1+\frac{2^2 i^2}{n^2}+4\frac{i}{n}\right)+1\right]\frac{1}{n}$$

$$= \frac{8}{n^3}\lim_{\|P\|\to 0}\sum_{i=1}^{n} i^2 + \frac{8}{n^2}\sum_{i=1}^{n} i + \frac{2}{n}\sum_{i=1}^{n} 2$$

$$= \lim_{\|P\|\to 0}\left(\frac{8}{n^3}\left(1^2+2^2+3^2+\ldots+n^2\right)+\frac{8}{n^2}(1+2+3+\ldots+n)+\frac{2}{n}(2n)\right)$$

$$= \lim_{\|P\|\to 0}\left(\frac{8}{n^3}\frac{n(n+1)(2n+1)}{6}+\frac{8}{n^2}\frac{n(n+1)}{2}+4\right)$$

$$\Rightarrow \lim_{\|P\|\to 0}\sum_{i=1}^{n} f(\xi_i)\delta_i = \lim_{\|P\|\to 0}\left(\frac{4}{3}\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right)+4\left(1+\frac{1}{n}\right)+4\right)$$

$$= \lim_{n\to\infty}\left(\frac{4}{3}\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right)+4\left(1+\frac{1}{n}\right)+4\right) \quad [\text{Using } (5.39)]$$

$$= \frac{4}{3}(1+0)(2+0)+4(1+0)+4 = \frac{8}{3}+8 = \frac{32}{3}$$

Since $\lim_{\|P\|\to 0}\sum_{i=1}^{n} f(\xi_i)\delta_i = \frac{32}{3}$ so by definition 2 of Riemann integration f is Riemann integrable and $\int_1^3 f(x)\,dx = \frac{32}{3}$.

2. We are given $F(x) = \int_5^x e^{t^2}dt$.

Let $f(t) = e^{t^2}$ which is an exponential function. We know that exponential function is continuous. Therefore, f(t) is a continuous function. So, using (5.80), we get

$$\frac{d}{dx}(F(x)) = f(x) = e^{x^2}.$$

Terminal Questions

1. In usual notations we are given

a = 0, b = 4, f(x) = x^3 + 7, length of the interval [0, 4] is 4 – 0 = 4 = b – a.

For the finite and closed interval [a, b] we form the partition

$$P = \{a = a+0h,\ a+h,\ a+2h,\ a+3h,\ \ldots,\ a+nh = b\},\ h = \frac{b-a}{n}$$

So, in our case partition will be

$$P = \{0+0h,\ 0+h,\ 0+2h,\ 0+3h,\ \ldots,\ 0+nh = 4\},\ \text{where } h = \frac{4-0}{n} = \frac{4}{n}$$

$$\Rightarrow P = \{0h,\ h,\ 2h,\ 3h,\ \ldots,\ nh = 4\},\ \text{where } h = \frac{4}{n}$$

Therefore, i^{th} sub interval $= [(i-1)h,\ ih]$, so

$$x_{i-1} = (i-1)h = \frac{4(i-1)}{n},\ x_i = ih = \frac{4i}{n},\ \text{as } h = \frac{4}{n} \text{ and } \delta_i = h = \frac{4}{n}$$

Now, we proceed to apply definition 1 and definition 2 of Riemann integration as follows.

Using Definition 1

Let $m_i = \text{glb of } f \text{ on } I_i = \left[\frac{4(i-1)}{n}, \frac{4i}{n}\right],\ i = 1, 2, 3, \ldots, n$

$M_i = \text{lub of } f \text{ on } I_i = \left[\frac{4(i-1)}{n}, \frac{4i}{n}\right],\ i = 1, 2, 3, \ldots, n$

Since f is an increasing function over the interval [0, 4] and hence over each sub-interval $I_i = \left[\frac{4(i-1)}{n}, \frac{4i}{n}\right],\ i = 1, 2, 3, \ldots, n,$ therefore

$$m_i = f\left(\frac{4(i-1)}{n}\right), \quad M_i = f\left(\frac{4i}{n}\right), \quad i = 1, 2, 3, \ldots, n,$$

So,

$$L(f, P) = \sum_{i=1}^{n} m_i\,\delta_i = \sum_{i=1}^{n} f\left(\frac{4(i-1)}{n}\right)\frac{4}{n} = 4\sum_{i=1}^{n}\left[\left(\frac{4(i-1)}{n}\right)^3 + 7\right]\frac{1}{n}$$

$$= 4\sum_{i=1}^{n}\left(\frac{4^3(i-1)^3}{n^3} + 7\right)\frac{1}{n} = \frac{256}{n^4}\sum_{i=1}^{n}(i-1)^3 + \frac{4}{n}\sum_{i=1}^{n}7$$

$$= \frac{256}{n^4}\left(0^3 + 1^3 + 2^3 + \ldots + (n-1)^3\right) + \frac{4}{n}(7n)$$

$$= \frac{256}{n^4}\left(1^3 + 2^3 + 3^3 + \ldots + (n-1)^3\right) + 28$$

$$= \frac{256}{n^4}\left(\frac{n(n-1)}{2}\right)^2 + 28$$

$$\left[\because 1^3 + 2^3 + 3^3 + \ldots + n^3 = \left(\frac{n(n+1)}{2}\right)^2, \text{ so } 1^3 + 2^3 + 3^3 + \ldots + (n-1)^3 = \left(\frac{n(n-1)}{2}\right)^2\right]$$

$$\Rightarrow L(f, P) = 64\left(1 - \frac{1}{n}\right)^2 + 28$$

$$U(f, P) = \sum_{i=1}^{n} M_i\,\delta_i = \sum_{i=1}^{n} f\left(\frac{4i)}{n}\right)\frac{4}{n} = 4\sum_{i=1}^{n}\left[\left(\frac{4i)}{n}\right)^3 + 7\right]\frac{1}{n}$$

$$= 4\sum_{i=1}^{n}\left(\frac{4^3 i^3}{n^3} + 7\right)\frac{1}{n} = \frac{256}{n^4}\sum_{i=1}^{n} i^3 + \frac{4}{n}\sum_{i=1}^{n}7$$

$$= \frac{256}{n^4}\left(1^3 + 2^3 + \ldots + n^3\right) + \frac{4}{n}(7n) = \frac{256}{n^4}\left(\frac{n(n+1)}{2}\right)^2 + 28$$

$$\Rightarrow U(f, P) = 64\left(1 + \frac{1}{n}\right)^2 + 28$$

By definition of lower Riemann integration refer (5.42), we have

$$\underline{\int_0^4} f(x)\,dx = \text{lub}\left\{L(f, P)\right\}_{P \in P[0, 4]} = \lim_{\|P\| \to 0}\left(64\left(1 - \frac{1}{n}\right)^2 + 28\right)$$

$$= \lim_{n \to \infty}\left(64\left(1 - \frac{1}{n}\right)^2 + 28\right) \qquad [\text{Using } (5.39)]$$

$$= 64(1 - 0) + 28 = 92$$

By definition of upper Riemann integration refer (5.43), we have

$$\overline{\int_0^4} f(x)\,dx = \text{glb}\left\{U(f, P)\right\}_{P \in P[0, 4]} = \lim_{\|P\| \to 0}\left(64\left(1 + \frac{1}{n}\right)^2 + 28\right)$$

$$= \lim_{n \to \infty}\left(64\left(1 + \frac{1}{n}\right)^2 + 28\right) \qquad [\text{Using } (5.39)]$$

$$= 64(1 + 0) + 28 = 92$$

Since $\underline{\int_0^4} f(x)\,dx = 92 = \overline{\int_0^4} f(x)\,dx,$ so by definition 1 of Riemann integration f is Riemann integrable and $\int_0^4 f(x)\,dx = 92.$

Using Definition 2

Let $\xi_i = \frac{4i}{n},\ \ i = 1, 2, 3, \ldots, n$ [Refer (5.46)]

Now,

$$\lim_{\|P\|\to 0}\sum_{i=1}^{n} f(\xi_i)\delta_i = \lim_{\|P\|\to 0}\sum_{i=1}^{n} f\left(\frac{4i}{n}\right)\frac{4}{n} \qquad \left[\because \xi_i = \frac{4i}{n} \text{ and } \delta_i = \frac{4}{n}\right]$$

$$= 4\lim_{\|P\|\to 0}\sum_{i=1}^{n}\left[\left(\frac{4i}{n}\right)^3 + 7\right]\frac{1}{n} = 4\lim_{\|P\|\to 0}\sum_{i=1}^{n}\left(\frac{4^3 i^3}{n^3} + 7\right)\frac{1}{n}$$

$$= \lim_{\|P\|\to 0}\left(\frac{256}{n^4}\sum_{i=1}^{n} i^3 + \frac{4}{n}\sum_{i=1}^{n} 7\right)$$

$$= \lim_{\|P\|\to 0}\left(\frac{256}{n^4}\left(1^3 + 2^3 + 3^3 + \ldots + n^3\right) + \frac{4}{n}(7n)\right)$$

$$= \lim_{\|P\|\to 0}\left(\frac{256}{n^4}\left(\frac{n(n+1)}{2}\right)^2 + 28\right)$$

$$\Rightarrow \lim_{\|P\|\to 0}\sum_{i=1}^{n} f(\xi_i)\delta_i = \lim_{\|P\|\to 0}\left(64\left(1+\frac{1}{n}\right)^2 + 28\right)$$

$$= \lim_{n\to\infty}\left(64\left(1+\frac{1}{n}\right)^2 + 28\right) \quad [\text{Using } (5.39)]$$

$$= 64(1+0)^2 + 28 = 64 + 28 = 92$$

Since $\lim_{\|P\|\to 0}\sum_{i=1}^{n} f(\xi_i)\delta_i = 92,$ so by definition 2 of Riemann integration f is Riemann integrable and $\int_0^4 f(x)\,dx = 92.$

2. Comparing $\int_5^{25} 30\,dx$ with $\int_a^b c\,dx,$ we get

a = 5, b = 25, c = 30, length of the interval [5, 25] is 25 – 5 = 20.

Now, using (5.50), we get

$$\int_5^{25} 30\,dx = 30(25-5) = 30(20) = 600.$$

3. In usual notations we are given

a = 0, b = 0.5, f(x) = 2x + 3, length of the interval [0, 0.5] is 0.5 – 0 = 0.5.

Given partition of the finite and closed interval [0, 0.5] is

$P = \{0, 0.1, 0.4, 0.5\}$

Sub-intervals are $I_1 = [0, 0.1],\ \ I_2 = [0.1, 0.4],\ \ I_3 = [0.4, 0.5].$

Length of these sub-intervals are
$\delta_1 = 0.1 - 0 = 0.1,\ \ \delta_2 = 0.4 - 0.1 = 0.3,\ \ \delta_2 = 0.5 - 0.4 = 0.1.$

Let $m_i = \text{glb of } f \text{ on } I_i,\ M_i = \text{lub of } f \text{ on } I_i,\ i = 1, 2, 3.$

Since f is an increasing function over the interval [0, 0.5] and hence over each sub-interval $I_i,\ \ i = 1, 2, 3.$ therefore

$m_1 = f(0) = 2(0) + 3 = 3,$ Similary, $m_2 = f(0.1) = 3.2,\ m_3 = f(0.4) = 3.8.$

$M_1 = f(0.1) = 2(0.1) + 3 = 3.2,$ Similary, $M_2 = f(0.4) = 3.8,\ M_3 = f(0.5) = 4$

Now, we find L(f, P) and U(f, P) as follows.

$$L(f, P) = \sum_{i=1}^{3} m_i\,\delta_i = m_1\delta_1 + m_2\delta_2 + m_3\delta_3$$
$$= (3)(0.1) + (3.2)(0.3) + (3.8)(0.1) = 1.64$$
$$U(f, P) = \sum_{i=1}^{3} M_i\,\delta_i = M_1\delta_1 + M_2\delta_2 + M_3\delta_3$$
$$= (3.2)(0.1) + (3.8)(0.3) + (4)(0.1) = 1.86$$

Remark 3: By Fundamental Theorem of Calculus, we know that

$\int_a^b f(x)\,dx = F(b) - F(a),$ where $F'(x) = f(x),\ a = 0, b = 0.5$

Now, $F'(x) = f(x) \Rightarrow F(x) = \int (2x + 3)\,dx = x^2 + 3x$

$\therefore F(b) = F(0.5) = (0.5)^2 + 3(0.5) = 0.25 + 1.5 = 1.75$

and $F(a) = F(0) = (0)^2 + 3(0) = 0$

So, $\int_0^{0.5} (2x + 3)\,dx = F(b) - F(a) = 1.75 - 0 = 1.75.$

Note that lower sum is 1.64 which is ≤ 1.75 and

upper sum is 1.86 which is ≥ 1.75 as expected.