

MAH
MT

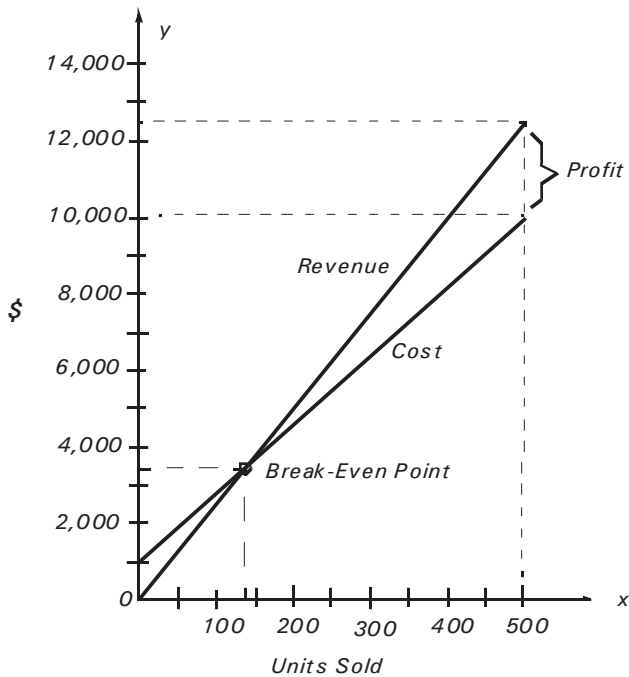
Mathematics

for Business, Science, and Technology

with MATLAB[®] and Excel[®] Computations

Third Edition

Steven T. Karris



*Includes a
Comprehensive
Treatment of Probability
and Statistics Illustrated
with Numerous
Examples*



Orchard Publications
www.orchardpublications.com

Mathematics for Business, Science, and Technology with MATLAB® and Excel® Computations, Third Edition

Copyright © 2007 Orchard Publications. All rights reserved. Printed in the United States of America. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written permission of the publisher.

Direct all inquiries to Orchard Publications, 39510 Paseo Padre Parkway, Fremont, California 94538, U.S.A.
URL: <http://www.orchardpublications.com>, e-mail: info@orchardpublications.com.

Product and corporate names are trademarks or registered trademarks of the MathWorks™, Inc., and Microsoft™ Corporation. They are used only for identification and explanation, without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Library of Congress Control Number: 2007922089

Copyright Number TX-5-471-563

ISBN-13: **978-1-934404-01-0**

ISBN-10: **1-934404-01-2**

Disclaimer

The publisher has used his best effort to prepare this text. However, the publisher and author makes no warranty of any kind, expressed or implied with regard to the accuracy, completeness, and computer codes contained in this book, and shall not be liable in any event for incidental or consequential damages in connection with, or arising out of, the performance or use of these programs.

Table of Contents

1	<i>Elementary Algebra</i>	1–1
1.1	Introduction.....	1–1
1.2	Algebraic Equations.....	1–2
1.3	Laws of Exponents.....	1–5
1.4	Laws of Logarithms.....	1–10
1.5	Quadratic Equations.....	1–13
1.6	Cubic and Higher Degree Equations.....	1–15
1.7	Measures of Central Tendency.....	1–15
1.8	Interpolation and Extrapolation.....	1–20
1.9	Infinite Sequences and Series.....	1–23
1.10	Arithmetic Series.....	1–23
1.11	Geometric Series.....	1–24
1.12	Harmonic Series.....	1–26
1.13	Proportions.....	1–27
1.14	Summary.....	1–29
1.15	Exercises.....	1–33
1.16	Solutions to End-of-Chapter Exercises.....	1–35

Excel Computations: Pages 1–7, 1–11, 1–16, 1–17, 1–33, 1–36

MATLAB Computations: Pages 1–7, 1–11

2	<i>Intermediate Algebra</i>	2–1
2.1	Systems of Two Equations.....	2–1
2.2	Systems of Three Equations.....	2–6
2.3	Matrices and Simultaneous Solution of Equations.....	2–7
2.4	Summary.....	2–26
2.5	Exercises.....	2–30
2.6	Solutions to End-of-Chapter Exercises.....	2–32

Excel Computations: Pages 2–23, 2–24, 2–30, 2–31

MATLAB Computations: Pages 2–15, 2–16, 2–24, 2–25, 2–32, 2–36

3	<i>Fundamentals of Geometry</i>	3–1
3.1	Introduction.....	3–1
3.2	Plane Geometry Figures.....	3–1
3.3	Solid Geometry Figures.....	3–16
3.4	Using Spreadsheets to Find Areas of Irregular Polygons.....	3–20
3.5	Summary.....	3–23

3.6 Exercises	3–27
3.7 Solutions to End-of-Chapter Exercises.....	3–29

Excel Computations: Pages 3–20, 3–30

4 *Fundamentals of Plane Geometry* 4–1

4.1 Introduction.....	4–1
4.2 Trigonometric Functions	4–2
4.3 Trigonometric Functions of an Acute Angle	4–2
4.4 Trigonometric Functions of an Any Angle	4–3
4.5 Fundamental Relations and Identities	4–6
4.6 Triangle Formulas.....	4–12
4.7 Inverse Trigonometric Functions	4–13
4.8 Area of Polygons in Terms of Trigonometric Functions	4–14
4.9 Summary	4–16
4.10 Exercises	4–18
4.11 Solutions to End-of-Chapter Exercises	4–19

Excel Computations: Pages 4–9 through 4–11, 4–13, 4–19 through 4–21

MATLAB Computations: Page 4–15

5 *Fundamentals of Calculus* 5–1

5.1 Introduction.....	5–1
5.2 Differential Calculus	5–1
5.3 The Derivative of a Function	5–3
5.4 Maxima and Minima.....	5–11
5.5 Integral Calculus	5–15
5.6 Indefinite Integrals.....	5–16
5.7 Definite Integrals	5–16
5.8 Summary	5–21
5.9 Exercises.....	5–23
5.10 Solutions to End-of-Chapter Exercises	5–24

Excel Computations: Pages 5–8, 5–11, 5–18

MATLAB Computations: Pages 5–24 through 5–26

6 *Mathematics of Finance and Economics* 6–1

6.1 Common Terms.....	6–1
6.1.1 Bond.....	6–1
6.1.2 Corporate Bond	6–1
6.1.3 Municipal Bond	6–1

6.1.4	Treasury Bond	6-1
6.1.5	Perpetuity	6-1
6.1.6	Perpetual Bond.....	6-1
6.1.7	Convertible Bond	6-2
6.1.8	Treasury Note	6-2
6.1.9	Treasury Bill	6-2
6.1.10	Face Value	6-2
6.1.11	Par Value.....	6-2
6.1.12	Book Value.....	6-3
6.1.13	Coupon Bond	6-3
6.1.14	Zero Coupon Bond.....	6-3
6.1.15	Junk Bond.....	6-3
6.1.16	Bond Rating Systems.....	6-3
6.1.17	Promissory Note	6-4
6.1.18	Discount Rate.....	6-4
6.1.19	Prime Rate	6-4
6.1.20	Mortgage Loan	6-4
6.1.21	Predatory Lending Practices.....	6-5
6.1.22	Annuity	6-5
6.1.23	Ordinary Annuity.....	6-5
6.1.24	Sinking Fund	6-6
6.2	Interest	6-6
6.2.1	Simple Interest.....	6-6
6.2.2	Compound Interest.....	6-8
6.2.3	Effective Interest Rate	6-22
6.3	Sinking Funds.....	6-24
6.4	Annuities.....	6-29
6.5	Amortization	6-34
6.6	Perpetuities.....	6-36
6.7	Valuation of Bonds.....	6-39
6.7.1	Calculating the Purchase Price of a Bond.....	6-40
6.7.2	Total Periodic Bond Disbursement.....	6-42
6.7.3	Calculation of Interest Rate of Bond	6-44
6.8	Spreadsheet Financial Functions	6-45
6.8.1	PV Function	6-46
6.8.2	FV Function	6-48
6.8.3	PMT Function.....	6-49
6.8.4	RATE Function.....	6-50
6.8.5	NPER Function.....	6-51
6.8.6	NPV Function	6-52
6.8.7	IIR Function.....	6-54

6.8.8	MIIR Function.....	6–56
6.8.9	IPMT Function.....	6–58
6.8.10	PPMT Function	6–59
6.8.11	ISPMT Function.....	6–60
6.9	The MATLAB Financial Toolbox	6–61
6.9.1	irr MATLAB Function.....	6–61
6.9.2	effrr MATLAB Function.....	6–62
6.9.3	pvfix MATLAB Function	6–62
6.9.4	pvvar MATLAB Function	6–63
6.9.5	fvfix MATLAB Function	6–64
6.9.6	fvvar MATLAB Function	6–65
6.9.7	annurate MATLAB Function.....	6–66
6.9.8	amortize MATLAB Function	6–67
6.10	Comparison of Alternate Proposals	6–68
6.11	Kelvin’s Law.....	6–71
6.12	Summary	6–75
6.13	Exercises.....	6–78
6.14	Solutions to End-of-Chapter Exercises	6–81

Excel Computations: Pages 6–10 through 6–14, 6–28, 6–29, 6–33, 6–35,
6–45 through 6–52, 6–54, 6–56, 6–58 through 6–61, 6–81

MATLAB Computations: Pages 6–61, 6–68

7	<i>Depreciation, Impairment, and Depletion</i>	7–1
7.1	Depreciation Defined.....	7–1
7.1.1	Items that Can Be Depreciated	7–2
7.1.2	Items that Cannot Be Depreciated	7–2
7.1.3	Depreciation Rules.....	7–2
7.1.4	When Depreciation Begins and Ends	7–3
7.1.5	Methods of Depreciation	7–3
7.1.5.1	Straight-Line (SL) Depreciation Method.....	7–4
7.1.5.2	Sum of the Years Digits (SYD) Method	7–5
7.1.5.3	Fixed-Declining Balance (FDB) Method	7–6
7.1.5.4	The 125%, 150%, and 200% General Declining Balance Methods... ..	7–8
7.1.5.5	The Variable Declining Balance Method.....	7–9
7.1.5.6	The Units of Production Method	7–10
7.1.5.7	Depreciation Methods for Income Tax Reporting	7–11
7.2	Impairments	7–18
7.3	Depletion	7–19
7.4	Valuation of a Depleting Asset	7–20
7.5	Summary.....	7–25

7.6 Exercises	7–27
7.7 Solutions to End-of-Chapter Exercises.....	7–28

Excel Computations: Pages 7–4, 7–6, 7–7, 7–9, 7–10, 7–29 through 7–32

MATLAB Computations: Pages 7–30, 7–31, 7–33

8 *Introduction to Probability and Statistics* 8–1

8.1 Introduction	8–1
8.2 Probability and Random Experiments	8–1
8.3 Relative Frequency	8–2
8.4 Combinations and Permutations	8–4
8.5 Joint and Conditional Probabilities.....	8–7
8.6 Bayes' Rule	8–11
8.7 Summary	8–13
8.8 Exercises.....	8–15
8.9 Solutions to End-of-Chapter Exercises.....	8–16

Excel Computations: Pages 8–6, 8–16

9 *Random Variables* 9–1

9.1 Definition of Random Variables	9–1
9.2 Probability Function	9–2
9.3 Cumulative Distribution Function.....	9–2
9.4 Probability Density Function.....	9–9
9.5 Two Random Variables	9–11
9.6 Statistical Averages	9–12
9.7 Summary	9–20
9.8 Exercises	9–23
9.9 Solutions to End-of-Chapter Exercises	9–24

Excel Computations: Page 9–3

10 *Common Probability Distributions and Tests of Significance* 10–1

10.1 Properties of Binomial Coefficients	10–1
10.2 The Binomial (Bernoulli) Distribution.....	10–2
10.3 The Uniform Distribution	10–6
10.4 The Exponential Distribution	10–10
10.5 The Normal (Gaussian) Distribution.....	10–12
10.6 Percentiles.....	10–32
10.7 The Student's t-Distribution	10–36
10.8 The Chi-Square Distribution.....	10–40

10.9	The F Distribution	10–43
10.10	Chebyshev’s Inequality	10–46
10.11	Law of Large Numbers	10–46
10.12	The Poisson Distribution	10–47
10.13	The Multinomial Distribution	10–52
10.14	The Hypergeometric Distribution	10–53
10.15	The Bivariate Normal Distribution	10–56
10.16	The Rayleigh Distribution	10–56
10.17	Other Probability Distributions	10–59
	10.17.1 The Cauchy Distribution.....	10–59
	10.17.2 The Geometric Distribution	10–59
	10.17.3 The Pascal Distribution.....	10–59
	10.17.4 The Weibull Distribution.....	10–60
	10.17.5 The Maxwell Distribution.....	10–61
	10.17.6 The Lognormal Distribution	10–61
10.18	Critical Values of the Binomial Distribution	10–62
10.19	Sampling Distribution of Means	10–63
10.20	Z –Score	10–64
10.21	Tests of Hypotheses and Levels of Significance.....	10–65
10.22	Tests of Significance	10–72
	10.22.1 The z–Test	10–72
	10.22.2 The t–Test	10–73
	10.22.3 The F–Test.....	10–74
	10.22.4 The Chi–Square Test.....	10–75
10.23	Summary.....	10–80
10.24	Exercises	10–89
10.24	Solutions to End–of–Chapter Exercises.....	10–91
Excel Computations: Pages 10–28, 10–29, 10–35, 10–36, 10–39, 10–40, 10–43, 10–45, 10–51, 10–55, 10–60, 10–62, 10–69 through 10–75, 10–77, 10–78, 10–93, 10–95		
MATLAB Computations: Pages 10–21, 10–28, 10–31, 10–32, 10–37, 10–41		

11	<i>Curve Fitting, Regression, and Correlation</i>	11–1
11.1	Curve Fitting	11–1
11.2	Linear Regression	11–2
11.3	Parabolic Regression	11–7
11.4	Covariance	11–10
11.5	Correlation Coefficient	11–12
11.6	Summary	11–17
11.7	Exercises	11–20

11.8 Solutions to End-of-Chapter Exercises	11–22
Excel Computations: Pages 11–3, 11–4, 11–6, 11–7, 11–8, 11–11 through 11–13, 11–15, 11–22 through 11–25	

12 *Analysis of Variance (ANOVA)* 12–1

12.1 Introduction.....	12–1
12.2 One-way ANOVA.....	12–1
12.3 Two-way ANOVA.....	12–8
12.3.1 Two-factor without Replication ANOVA.....	12–8
12.3.2 Two-factor with Replication ANOVA.....	12–14
12.4 Summary	12–25
12.5 Exercises.....	12–29
12.6 Solutions to End-of-Chapter Exercises	12–31
Excel Computations: Pages 12–6 through 12–8, 12–13, 12–14, 12–22 through 12–24, 12–31, 12–33, 12–35, 12–37	

A *Numbers and Arithmetic Operations* A–1

A.1 Number Systems.....	A–1
A.2 Positive and Negative Numbers.....	A–1
A.3 Addition and Subtraction	A–2
A.4 Multiplication and Division	A–7
A.5 Integer, Fractional, and Mixed Numbers.....	A–10
A.6 Reciprocals of Numbers	A–11
A.7 Arithmetic Operations with Fractional Numbers.....	A–13
A.8 Exponents.....	A–22
A.9 Scientific Notation	A–25
A.10 Operations with Numbers in Scientific Notation.....	A–27
A.11 Square and Cubic Roots.....	A–30
A.12 Common and Natural Logarithms	A–32
A.13 Decibel	A–33
A.14 Percentages.....	A–34
A.15 International System of Units (SI).....	A–35
A.16 Graphs	A–40
A.17 Summary.....	A–44
A.18 Exercises	A–50
A.19 Solutions to End-of-Appendix Exercises	A–51
Excel Computations: Pages A–13 through A–18, A–26, A–30 through A–35, A–37, A–39, A–51	

B	<i>Introduction to MATLAB®</i>	B-1
B.1	MATLAB® and Simulink®	B-1
B.2	Command Window	B-1
B.3	Roots of Polynomials	B-3
B.4	Polynomial Construction from Known Roots	B-4
B.5	Evaluation of a Polynomial at Specified Values	B-5
B.6	Rational Polynomials	B-8
B.7	Using MATLAB to Make Plots	B-10
B.8	Subplots	B-18
B.9	Multiplication, Division, and Exponentiation	B-18
B.10	Script and Function Files	B-26
B.11	Display Formats	B-31

MATLAB Computations: Entire Appendix

C	<i>The Gamma and Beta Functions and Distributions</i>	C-1
C.1	The Gamma Function	C-1
C.2	The Gamma Distribution	C-15
C.3	The Beta Function	C-17
C.4	The Beta Distribution	C-20

Excel Computations: Pages C-5, C-11, C-16, C-17, C-19

MATLAB Computations: Pages C-3 through C-5, C-10, C-16, C-19, C-20

D	<i>Introduction to Markov Chains</i>	D-1
D.1	Stochastic Processes	D-1
D.2	Stochastic Matrices	D-1
D.3	Transition Diagrams	D-4
D.4	Regular Stochastic Matrices	D-5
D.5	Some Practical Examples	D-7

Excel Computations: Pages D-8, D-11, D-14

E	<i>The Lambda Index</i>	E-1
E.1	Introduction	E-1
E.2	The Lambda Index Defined	E-1

E.3 Spreadsheet Construction	E-2
Excel Computations: Page E-2	

F *The Black–Scholes Stock Options Model* F-1

F.1 Stock Options	F-1
F.2 The Black-Scholes Model Equations	F-2
F.3 Spreadsheet for the Black–Scholes Model.....	F-3
Excel Computations: Page F-3	

G *Forecasting Bankruptcy* G-1

G.1 The Financial Ratios.....	G-1
G.2 Interpretation of The Altman Z-Score	G-2
G.3 Spreadsheet for The Altman Z-Score.....	G-2
Excel Computations: Page G-2	

References R-1

Index IN-1

Preface

This text is written for

- a. high school graduates preparing to take business or science courses at community colleges or universities
- b. working professionals who feel that they need a math review from the very beginning
- c. young students and working professionals who are enrolled in continued education institutions, and majoring in business related topics, such as business administration and accounting, and those pursuing a career in science, electronics, and computer technology.

Chapter 1 is an introduction to the basics of algebra.

Chapter 2 is a continuation of Chapter 1 and presents some practical examples with systems of two and three equations.

Chapters 3 and 4 discuss the fundamentals of geometry and trigonometry respectively. These treatments are not exhaustive; these chapters contain basic concepts that are used in science and technology.

Chapter 5 is an abbreviated, yet a practical introduction to calculus.

Chapters 6 and 7 serve as an introduction to the mathematics of finance and economics and the concepts are illustrated with numerous real-world applications and examples.

Chapters 8 through 12 are devoted to probability and statistics. Many practical examples are given to illustrate the importance of this branch of mathematics. The topics that are discussed, are especially important in management decisions and in reliability. Some readers may find certain topics hard to follow; these may be skipped without loss of continuity.

In all chapters, numerous examples are given to teach the reader how to obtain quick answers to some complicated problems using computer tools such as MATLAB[®] and Microsoft Excel.[®]

Appendix A contains a review of the basic arithmetic operations, introduces the SI system of units, and discusses different types of graphs. It is written for the reader who needs a review of the very basics of arithmetic.

Appendix B is intended to teach the interested reader how to use MATLAB. Many practical examples are presented. The Student Edition of MATLAB is an inexpensive software package; it can be found in many college bookstores, or can be obtained directly from

The MathWorks™ Inc., 3 Apple Hill Drive, Natick, MA 01760–2098

Phone: 508 647–7000, Fax: 508 647–7001

http://www.mathworks.com

e-mail: info@mathwork.com

Appendix C introduces the gamma and beta functions. These appear in the gamma and beta distributions and find many applications in business, science, and engineering. For instance, the Erlang distributions, which are a special case of the gamma distribution, form the basis of queuing theory.

Appendix D is an introduction to Markov chains. A few practical examples illustrate their application in making management decisions.

Appendices E, F, and G are introductions to the Lambda Index, Black–Scholes stock options pricing, and the Altman Z–score bankruptcy prediction respectively.

Every chapter and appendix in this text is supplemented with Excel and / or MATLAB scripts to verify the computations and to construct relevant plots. The pages where the Excel and MATLAB scripts appear are listed in the Table of Contents.

New to the Second Edition

This is a refined revision of the first edition. The most notable changes are the addition of the new Chapters 6 and 7, chapter–end summaries, and detailed solutions to all exercises. The latter is in response to many students and working professionals who expressed a desire to obtain the author’s solutions for comparison with their own.

New to the Third Edition

This is a refined revision of the second edition. The most notable is the addition of Appendices E, F, and G. All chapters and Appendices A through D have been rewritten, and graphs have been redrawn with the latest MATLAB® Student Version, Release 14.

All feedback for typographical errors and comments will be most welcomed and greatly appreciated.

Orchard Publications

www.orchardpublications.com

info@orchardpublications.com

This chapter is an introduction to algebra and algebraic equations. A review of the basic arithmetic concepts is provided in Appendix A. It also introduces some financial functions that are used with Excel. Throughout this text, a left justified horizontal bar will denote the beginning of an example, and a right justified horizontal bar will denote the end of the example. These bars will not be shown whenever an example begins at the top of a page or at the bottom of a page. Also, when one example follows immediately after a previous example, the right justified bar will be omitted.

1.1 Introduction

Algebra is the branch of mathematics in which letters of the alphabet represent numbers or a set of numbers. *Equations* are equalities that indicate how some quantities are related to others. For example, the equation

$$^{\circ}\text{C} = \frac{5}{9}(^{\circ}\text{F} - 32) \quad (1.1)$$

is a relation or formula that enables us to convert degrees Fahrenheit, $^{\circ}\text{F}$, to degrees Celsius, $^{\circ}\text{C}$. For instance, if the temperature is 77°F , the equivalent temperature in $^{\circ}\text{C}$ is

$$^{\circ}\text{C} = \frac{5}{9}(77 - 32) = \frac{5}{9} \times 45 = 25$$

We observe that the mathematical operation in parentheses, that is, the subtraction of 32 from 77, was performed first. This is because numbers within parentheses have *precedence* (priority) over other operations.

In algebra, the *order of precedence*, where 1 is the highest and 4 is the lowest, is as follows:

1. Quantities inside parentheses ()
2. Exponentiation
3. Multiplication and Division
4. Addition or Subtraction

Chapter 1 Elementary Algebra

Example 1.1

Simplify the expression

$$a = \frac{-2^3}{(8-3)^2 \times 4}$$

Solution:

This expression is reduced in steps as follows:

Step 1. Subtracting 3 from 8 inside the parentheses yields 5. Then, $a = \frac{-2^3}{5^2 \times 4}$

Step 2. Performing the exponentiation operations we get $a = \frac{-8}{25 \times 4}$

Step 3. Multiplying 25 by 4 we get $a = \frac{-8}{100}$

Step 4. Dividing -8 by 100 we get $a = -0.08$ (Simplest form)

An *equality* is a mathematical expression where the left side is equal to the right side. For example, $7 + 3 = 10$ is an equality since the left and right sides are equal to each other.

1.2 Algebraic Equations

An *algebraic equation* is an equality that contains one or more unknown quantities, normally represented by the last letters of the alphabet such as x , y , and z . For instance, the expressions $x + 5 = 15$, $3 \times y = 24$, and $x \times y = 12$ are algebraic equations. In the last two equations, the multiplication sign between 3 and y and x and y is generally omitted; thus, they are written as $3y = 24$ and $xy = 12$. Henceforth, terms such as $3y$ will be interpreted as $3 \times y$, xy as $x \times y$, and, in general, XY as $X \times Y$.

Solving an equation means finding the value of the unknown quantity that will make the equation an equality with no unknowns. The following properties of algebra enable us to find the numerical value of the unknown quantity in an equation.

Property 1

The same number may be added to, or subtracted from both sides of an equation.

Example 1.2

Given the equality

$$7 + 5 = 12$$

prove that Property 1 holds if we

- a. add 3 to both sides
- b. subtract 2 from both sides

Proof:

- a. Adding 3 to both sides we get

$$7 + 5 + 3 = 12 + 3 \text{ or } 15 = 15$$

and thus, the equality holds.

- b. Subtracting 2 from both sides we get

$$7 + 5 - 2 = 12 - 2 \text{ or } 10 = 10$$

and thus, the equality holds.

Example 1.3

Solve the following equation, that is, find the value of x .

$$x - 5 = 15$$

Solution:

We need to find a value for x so that the equation will still hold after the unknown x has been replaced by the value that we have found. For this type of equations, a good approach is to find a number that when added to or subtracted from both sides of the equation, the left side will contain only the unknown x . For this example, this will be accomplished if we add 5 to both sides of the given equation. When we do this, we get

$$x - 5 + 5 = 15 + 5$$

and after simplification, we obtain the value of the unknown x as

$$x = 20$$

We can check this answer by substitution of x into the given equation. Thus, $20 - 5 = 15$ or $15 = 15$. Therefore, our answer is correct.

Chapter 1 Elementary Algebra

Example 1.4

Solve the following equation, that is, find the value of x .

$$x + 5 = 15$$

Solution:

Again, we need to find a number such that when added to or subtracted from both sides of the equation, the left side will have the unknown x only. In this example, this will be accomplished if we subtract 5 from both sides of the given equation. Doing this, we get

$$x + 5 - 5 = 15 - 5$$

and after simplification, we obtain the value of the unknown x as

$$x = 10$$

To verify that this is the correct answer, we substitute this value into the given equation and we get $10 + 5 = 15$ or $15 = 15$.

Property 2

Each side of an equation can be multiplied or divided* by the same number.

Example 1.5

Solve the following equation, that is, find the value of y .

$$\frac{y}{3} = 24 \tag{1.2}$$

Solution:

We need to eliminate the denominator 3 on the left side of the equation. This is done by multiplying both sides of the equation by 3. Then,

$$\frac{y}{3} \times 3 = 24 \times 3$$

and after simplification we get

$$y = 72 \tag{1.3}$$

To verify that this is correct, we divide 72 by 3; this yields 24, and this is equal to the right side of

* Division by zero is meaningless; therefore, it must be avoided.

the given equation.

Note 1.1

In an algebraic term such as $4x$ or $(a + b)x$, the number or symbol multiplying a variable or an unknown quantity is called the *coefficient* of that term. Thus, in the term $4x$, the number 4 is the coefficient of that term, and in $(a + b)x$, the coefficient is $(a + b)$. Likewise, the coefficient of y in (1.2) is $1/3$, and the coefficient of y in (1.3) is 1 since $1y = y$. In other words, every algebraic term has a coefficient, and if it is not shown, it is understood to be 1 since, in general, $1x = x$.

Example 1.6

Solve the following equation, that is, find the value of z .

$$3z = 24$$

Solution:

We need to eliminate the coefficient 3 of z on the left side of the equation. This is done by dividing both sides of the equation by 3. Then,

$$\frac{3z}{3} = \frac{24}{3}$$

and after simplification

$$z = 8$$

Other algebraic equations may contain exponents and logarithms. For those equations we may need to apply the laws of exponents and the laws of logarithms. These are discussed next.

1.3 Laws of Exponents

For any number a and for a positive integer m , the exponential number a^m is defined as

$$\underbrace{a \cdot a \cdot a \cdot \dots \cdot a}_m = \text{number of } a\text{'s} \tag{1.4}$$

where the dots between the a 's in (1.4) denote multiplication.

The *laws of exponents* state that

$$a^m \cdot a^n = a^{m+n} \tag{1.5}$$

Also,

$$\frac{a^m}{a^n} = \begin{cases} a^{m-n} & \text{if } m > n \\ 1 & \text{if } m = n \\ \frac{1}{a^{n-m}} & \text{if } m < n \end{cases} \quad (1.6)$$

$$(a^m)^n = a^{mn} \quad (1.7)$$

The n th root is defined as the inverse of the n th exponent, that is, if

$$b^n = a \quad (1.8)$$

then

$$b = \sqrt[n]{a} \quad (1.9)$$

If, in (1.9) n is an odd positive integer, there will be a unique number satisfying the definition for $\sqrt[n]{a}$ for any value of a^* . For instance, $\sqrt[3]{512} = 8$ and $\sqrt[3]{-512} = -8$.

If, in (1.9) n is an even positive integer, for positive values of a there will be two values, one positive and one negative. For instance, $\sqrt{121} = \pm 11$. For additional examples, the reader may refer to the section on square and cubic roots in Appendix A.

If, in (1.8) n is even positive integer and a is negative, $\sqrt[n]{a}$ cannot be evaluated. For instance, $\sqrt{-3}$ cannot be evaluated; it results in a complex number.

It is also useful to remember the following definitions from Appendix A.

$$a^0 = 1 \quad (1.10)$$

$$a^{p/q} = \sqrt[q]{a^p} \quad (1.11)$$

$$a^{-1} = \frac{1}{a^1} = \frac{1}{a} \quad (1.12)$$

These relations find wide applications in business, science, technology, and engineering. We will consider a few examples to illustrate their application.

Example 1.7

The equation for calculating the present value (PV) of an ordinary annuity is

* In advance mathematics, there is a restriction. However, since in this text we are only concerned with real numbers, no restriction is imposed.

$$PV = \text{Payment} \times \frac{1 - (1 + \text{Interest})^{-n}}{\text{Interest}} \quad (1.13)$$

where

PV = cost of the annuity at present

$Payment$ = amount paid at the end of the year

$Interest$ = Annual interest rate which the deposited money earns

n = number of years the $Payment$ will be received.

Example 1.8

Suppose that an insurance company offers to pay us \$12,000 at the end of each year for the next 25 years provided that we pay the insurance company \$130,000 now, and we are told that our money will earn 8% per year. Should we accept this offer?

Solution:

Payment = \$12,000, Interest = 0.08 (8%), and $n = 25$ years. Let us calculate PV using (1.13).

$$PV = 12000 \times \frac{1 - (1 + 0.08)^{-25}}{0.08} \quad (1.14)$$

We will use Excel to compute (1.14). We select any cell and we type in the following expression where the *asterisk* (*) denotes multiplication, the *slash* (/) division, and the *caret* (^) exponentiation.

$$=12000*(1-(1+0.08)^{-25})/0.08$$

Alternately, we can use MATLAB[®] with the statement

$$PV=12000*(1-(1+0.08)^{-25})/0.08$$

Both Excel and MATLAB return \$128,097 and this is the fair amount that if deposited now at 8% interest, will earn \$12,000 per year for 25 years. This is less than \$130,000 which the insurance company asks for, and therefore, we should reject the offer.

Excel has a “build-in” function, called PV that computes the present value directly, that is, without the formula of (1.13). To invoke and use this function, we perform the sequential steps $f_x > \text{Financial} > PV > Rate = 0.08 > Nper = 25 > Pmt = 12000 > OK$. We observe that the numer-

* An introduction to MATLAB is provided in Appendix B. MATLAB applications may be skipped without loss of continuity. However, it is highly recommended since MATLAB now has become the standard for advanced computation. The reader may begin learning it with the inexpensive MATLAB Student Version.

Chapter 1 Elementary Algebra

cal value is the same as before, but it is displayed in red and within parentheses, that is, as negative value. It is so indicated because Excel interprets outgoing money as negative cash.

Note 2.2

Excel has several financial functions that apply to annuities.* It is beyond the scope of this text to discuss all of them. The interested reader may invoke the *Help* feature in Excel to get the description of these functions.

Property 3

Each side of an equation can be raised to the same power.

Example 1.9

Solve the following equation, that is, find the value of x .

$$\sqrt{x} = 5 \quad (1.15)$$

Solution:

As a first step, we use the alternate designation of the square root; this is discussed in Appendix A, Page A–31. Then,

$$x^{1/2} = 5 \quad (1.16)$$

Next, we square (raise to power 2) both sides of (1.16), and we get

$$(x^{1/2})^2 = 5^2 \quad (1.17)$$

Now, multiplication of the exponents on the left side of (1.17) using (1.7) yields

$$x^1 = 5^2$$

and after simplification, the answer is

$$x = 25$$

Example 1.10

Solve the following equation, that is, find the value of z .

$$\sqrt[3]{z} = 8 \quad (1.18)$$

* Annuities are discussed in Chapter 6, Page 6–6.

Solution:

As a first step, we use the alternate designation of the cubic root. This is discussed in Appendix A. Then,

$$z^{1/3} = 8 \tag{1.19}$$

Next, we cube (raise to power 3) both sides of (1.19) and we obtain

$$(z^{1/3})^3 = 8^3 \tag{1.20}$$

Now, multiplication of the exponents on the left side using relation (1.7), Page 1–6, yields

$$z^1 = 8^3$$

and after simplification, the answer is

$$z = 512$$

Other equations can be solved using combinations of Properties (1.5 through (1.7), Page 1–6.

Example 1.11

Let us consider the temperature conversion from degrees Fahrenheit to degrees Celsius. This is discussed in Appendix A, Page A–31. The relation (1.21) below converts degrees Fahrenheit to degrees Celsius. We wish to solve for °F so that the conversion will be degrees Celsius to degrees Fahrenheit.

Solution:

We begin with

$$^{\circ}\text{C} = \frac{5}{9}(^{\circ}\text{F} - 32) \tag{1.21}$$

Next, we multiply both sides of (1.21) by $9/5$. Then,

$$\left(\frac{9}{5}\right) \cdot ^{\circ}\text{C} = \left(\frac{9}{5}\right) \cdot \frac{5}{9}(^{\circ}\text{F} - 32) \tag{1.22}$$

and after simplifying the right side, we get

$$\frac{9}{5}^{\circ}\text{C} = ^{\circ}\text{F} - 32 \tag{1.23}$$

Now, we add 32 to both sides of (1.23) to eliminate -32 from the right side. Then,

$$\frac{9}{5}^{\circ}\text{C} + 32 = ^{\circ}\text{F} - 32 + 32 \quad (1.24)$$

Finally, after simplification and interchanging the positions of the left and right sides, we obtain

$$^{\circ}\text{F} = \frac{9}{5}^{\circ}\text{C} + 32$$

1.4 Laws of Logarithms

Common and natural logarithms are introduced in Appendix A. By definition, if

$$x = a^y \quad (1.25)$$

where a is the base and y is the exponent (power), then

$$y = \log_a x \quad (1.26)$$

The *laws of logarithms* state that

$$\log_a(xy) = \log_a x + \log_a y \quad (1.27)$$

$$\log_a\left(\frac{x}{y}\right) = \log_a x - \log_a y \quad (1.28)$$

$$\log_a(x^n) = n\log_a x \quad (1.29)$$

Logarithms are very useful because the laws of (1.27) through (1.29), allow us to replace multiplication, division and exponentiation by addition, subtraction and multiplication respectively. We used logarithms in Appendix A to define the decibel.

As we mentioned in Appendix A, the *common (base 10) logarithm* and the *natural (base e) logarithm*, where e is an irrational (endless) number whose value is 2.71828, are the most widely used. Both occur in many formulas of probability, statistics, and financial formulas. For instance, the formula (1.30) below computes the number of periods required to accumulate a specified future value by making equal payments at the end of each period into an interest-bearing account.

$$n = \frac{\ln(1 + (\text{Interest} \times \text{FutureValue})/\text{Payment})}{\ln(1 + \text{Interest})} \quad (1.30)$$

where:

n = number of periods (months or years)

\ln = natural logarithm

Interest = earned interest

Future Value = desired amount to be accumulated

Payment = equal payments (monthly or yearly) that must be made.

Example 1.12

Compute the number of months required to accumulate Future Value = \$100,000 by making a monthly payment of \$500 into a savings account paying 6% annual interest compounded monthly.

Solution:

The monthly interest rate is 6%/12 or 0.5%, and thus in (1.30), Interest = 0.5%=0.005 , Future Value = \$100,000, and Payment = \$500 . Then,

$$n = \frac{\ln(1 + (0.005 \times 100,000)/500)}{\ln(1 + 0.005)} \tag{1.31}$$

We will use Excel to find the value of *n*. In any cell, we enter the following formula:

$$=LN(1+(0.005*100000)/500)/LN(1+0.005)$$

Alternately, we can use the MATLAB statement*

$$n=\log(1+(0.005*100000)/500)/\log(1+0.005)$$

Both Excel and MATLAB return 138.9757 ; we round this to 139 . This represents the number of the months that a payment of \$500 is required to be deposited at the end of each month. This number, representing months, is equivalent to 11 years and 7 months.

Excel has a build-in function named NPER and its *syntax* (orderly arrangement) is

NPER(rate,pmt,pv,fv,type) where, for this example,

rate = 0.005

pmt = -500 (minus because it is a cash outflow)

pv = 0 (present value is zero)

fv = 100000

type = 0 which means that payments are made at the end of each period.

* With MATLAB, the function **log(x)** is used for the natural logarithm, and the command **log10(x)** is used for the common logarithm.

Chapter 1 Elementary Algebra

Then, in any cell we enter the formula

$$=NPER(0.005,-500,0,100000,0)$$

and we observe that Excel returns 138.9757. This is the same as the value that we found with the formula of (1.31).

The formula of (1.30) uses the natural logarithm. Others use the common logarithm.

Although we can use a PC, or a calculator to find the log of a number, it is useful to know the following facts that apply to common logarithms.

- I. Common logarithms consist of an integer, called the *characteristic* and an endless decimal called the *mantissa* *.
- II. If the decimal point is located immediately to the right of the msd of a number, the characteristic is zero; if the decimal point is located after two digits to the right of the msd, the characteristic is 1; if after three digits, it is 2 and so on. For instance, the characteristic of 1.9 is 0, of 58.3 is 1, and of 476.5 is 2.
- III. If the decimal point is located immediately to the left of the msd of a number, the characteristic is -1 , if located two digits to the left of the msd, it is -2 , if after three digits, it is -3 and so on. Thus, the characteristic of 0.9 is -1 , of 0.0583 is -2 , and of 0.004765 is -3 .
- IV. The mantissa cannot be determined by inspection; it must be extracted from tables of common logarithms.
- V. Although the common logarithm of a number less than one is negative, it is written with a *negative* characteristic and a *positive* mantissa. This is because the mantissas in tables are given as positive numbers.
- VI. Because mantissas are given in math tables as positive numbers, the negative sign is written above the characteristic. This is to indicate that the negative sign does not apply to the mantissa. For instance, $\log 0.00319 = \bar{3}.50379$, and since $-3 = 7 - 10$, this can be written as $\log 0.00319 = 7.50379 - 10 = -2.4962$.

A convenient method to find the characteristic of logarithms is to first express the given number in scientific notation;† the characteristic then is the exponent.

* Mantissas for common logarithms appear in books of mathematical tables. No such tables are provided here since we will not use them in our subsequent discussion. A good book with mathematical tables is *Handbook of Mathematical Functions* by Dover Publications.

† Scientific notation is discussed in Appendix A.

For negative numbers, the mantissa is the 9's complement* of the mantissa given in math tables.

Example 1.13

Given that $x = 73,000,000$ and $y = 0.0000000073$, find

- a. $\log x$ b. $\log y$

Solution:

a.
$$x = 73,000,000 = 7.3 \times 10^7$$

$$\log x = 7.8633$$

b.
$$y = 0.0000000073 = 7.3 \times 10^{-10}$$

$$\log y = -10.8633 = 0.8633 - 10 = -9.1367$$

1.5 Quadratic Equations

Quadratic equations are those that contain equations of second degree. The general form of a quadratic equation is

$$ax^2 + bx + c = 0 \tag{1.32}$$

where a , b and c are real constants (positive or negative). Let x_1 and x_2 be the roots† of (1.33). These can be found from the formulas

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \tag{1.33}$$

The quantity $b^2 - 4ac$ under the square root is called the *discriminant* of a quadratic equation.

Example 1.14

Find the roots of the quadratic equation

$$x^2 - 5x + 6 = 0 \tag{1.34}$$

* The 9's complement of a number is obtained by subtraction of that number from a number consisting of 9's with the same number of digits as the number. For the example cited above, we subtract 50379 from 99999 and we obtain 49620 and this is the mantissa of the number.

† Roots are the values which make the left and right sides of an equation equal to each other.

Chapter 1 Elementary Algebra

Solution:

Using the quadratic formulas of (1.33), we obtain

$$x_1 = \frac{-(-5) + \sqrt{(-5)^2 - 4 \times 1 \times 6}}{2 \times 1} = \frac{5 + \sqrt{25 - 24}}{2} = \frac{5 + \sqrt{1}}{2} = \frac{5 + 1}{2} = 3$$

and

$$x_2 = \frac{-(-5) - \sqrt{(-5)^2 - 4 \times 1 \times 6}}{2 \times 1} = \frac{5 - \sqrt{25 - 24}}{2} = \frac{5 - \sqrt{1}}{2} = \frac{5 - 1}{2} = 2$$

We observe that this equation has two unequal positive roots.

Example 1.15

Find the roots of the quadratic equation

$$x^2 + 4x + 4 = 0 \tag{1.35}$$

Solution:

Using the quadratic formulas of (1.33), we get

$$x_1 = \frac{-4 + \sqrt{4^2 - 4 \times 1 \times 4}}{2 \times 1} = \frac{-4 + \sqrt{16 - 16}}{2} = \frac{-4 + 0}{2} = -2$$

and

$$x_2 = \frac{-4 - \sqrt{4^2 - 4 \times 1 \times 4}}{2 \times 1} = \frac{-4 - \sqrt{16 - 16}}{2} = \frac{-4 - 0}{2} = -2$$

We observe that this equation has two equal negative roots.

Example 1.16

Find the roots of the quadratic equation

$$2x^2 + 4x + 5 = 0 \tag{1.36}$$

Solution:

Using (1.33), we obtain

$$x_1 = \frac{-4 + \sqrt{4^2 - 4 \times 1 \times 5}}{2 \times 2} = \frac{-4 + \sqrt{16 - 20}}{4} = \frac{-4 + \sqrt{-4}}{4} \quad (\text{value cannot be determined})$$

and

$$x_2 = \frac{-4 - \sqrt{4^2 - 4 \times 1 \times 5}}{2 \times 2} = \frac{-4 - \sqrt{16 - 20}}{4} = \frac{-4 - \sqrt{-4}}{4} \quad (\text{value cannot be determined})$$

Here, the square root of -4 , i.e., $\sqrt{-4}$ is undefined. It is an imaginary number and, as stated earlier, imaginary numbers will not be discussed in this text.

In general, if the coefficients a , b and c are real constants (known numbers), then:

- I. If $b^2 - 4ac$ is positive, as in Example 1.14, the roots are real and unequal.
- II. If $b^2 - 4ac$ is zero, as in Example 1.15, the roots are real and equal.
- III. If $b^2 - 4ac$ is negative, as in Example 1.16, the roots are imaginary.

1.6 Cubic and Higher Degree Equations

A *cubic equation* has the form

$$ax^3 + bx^2 + cx + d = 0 \tag{1.37}$$

and higher degree equations have similar forms. Formulas and procedures for solving these equations are included in books of mathematical tables. We will not discuss them here since their applications to business, basic science, and technology are limited. They are useful in higher mathematics applications and in engineering.

1.7 Measures of Central Tendency

Measures of central tendency are very important in probability and statistics and these are discussed in detail in Chapters 8 through 12. The intent here is to become familiar with terminologies used to describe data.

When we analyze data, we begin with the calculation of a single number, which we assume represents all the data. Because data often have a central cluster, this number is called a *measure of central tendency*. The most widely used are the *mean*, *median*, and *mode*. These are described below.

The *arithmetic mean* is the value obtained by dividing the sum of a set of quantities by the number of quantities in the set. It is also referred to as the *average*.

The arithmetic mean or simply the mean, is denoted with the letter x with a bar above it, and is computed from the equation

$$\bar{x} = \frac{\sum x}{n} \tag{1.38}$$

Chapter 1 Elementary Algebra

where the symbol Σ stands for summation and n is the number of data, usually called *sample*.

Example 1.17

The ages of 15 college students in a class are

24, 26, 27, 23, 31, 29, 25, 28, 21, 23, 32, 25, 30, 24, 26

Compute the mean (average) age of this group of students.

Solution:

Here, the sample is $n = 15$ and using (1.38) we obtain

$$\bar{x} = \frac{24 + 26 + 27 + 23 + 31 + 29 + 25 + 28 + 21 + 23 + 32 + 25 + 30 + 24 + 26}{15}$$

$$\bar{x} = \frac{389}{25} = 26.27 \approx 26$$

where the symbol \approx stands for approximately equal to.

We will check out answer with the Excel **AVERAGE** function to become familiar with it.

We start with a blank worksheet and we enter the given numbers in Cells **A1** through **A15**. In **A16** we type **=AVERAGE(A1:A15)**. Excel displays the answer 26.26667. We will use these values for the next example; therefore, it is recommended that they should not be erased.

The *median* of a sample is the value that separates the lower half of the data, from the upper half. To find the median, we arrange the values of the sample in increasing (ascending) order. If the number of the sample is odd, the median is in the middle of the list; if even, the median is the mean (average) of the two values closest to the middle of the list. We denote the median as M_d .

Example 1.18

Given the sample of Example 1.17, find the median.

Solution:

The given sample is repeated here for convenience.

24, 26, 27, 23, 31, 29, 25, 28, 21, 23, 32, 25, 30, 24, 26

We can arrange this sample in ascending (increasing) order with pencil and paper; however, we will let Excel do the work for us. Unless this list has been erased, it still exists in **A1:A15**. Now, we

erase the value in A16 by pressing the *Delete* key. We highlight the range A1:A15 and click on **Data>Sort>Column A>Ascending>OK**. We observe that the numbers now appear in ascending order, the median appears in A8 and has the value of 26, thus, for this example, $M_d = 26$.

Excel can find the median without first sorting the data. To illustrate the procedure, we undo sort by clicking on **Edit>Undo Sort** and we observe that the list now appears as entered the first time. We select any cell, we type `=MEDIAN(A1:A15)`, and we observe that Excel displays 26. Again, we will use these values for the next example; therefore, it is recommended that they should not be erased.

The *mode* is the value in a sample that occurs most often. If, in a sequence of numbers, no number appears two or more times, the sample has no mode. The mode, if it exists, may or may not be unique. If two such values exist, we say that the sample is *bimodal*, and if three values exist, we call it *trimodal*. We will denote the mode as M_o .

Example 1.19

Find the mode for the sample of Example 1.17.

Solution:

We assume that the data appear in the original order, that is, as

24, 26, 27, 23, 31, 29, 25, 28, 21, 23, 32, 25, 30, 24, 26

Let us sort these values as we did in Example 1.17. When this is done, we observe that the values 24, 25, and 26 each appear twice in the sample. Therefore, we say that this sample is trimodal.

Excel has also a function that computes the mode; however, if the sample has no unique mode, it displays only the first, and gives no indication that the sample is bimodal or trimodal. To verify this, we select any cell, and we type `=MODE(A1:A15)`. Excel displays the value 24.

Note 2.3

Textbooks in statistics provide formulas for the computation of the median and mode. We do not provide them here because, for our purposes, these are not as important as the arithmetic mean. In Chapters 9 and 10 we will discuss other important quantities such as the expected value, variance, standard deviation, and probability distributions. We will also present numerous practical applications.

Another useful measure of central tendency is the *moving average*. The following discussion will help us understand the meaning of a weighted moving average.

Chapter 1 Elementary Algebra

Suppose that the voltages displayed by an electronic instrument in a 5-day period, Monday through Friday, were 23.5, 24.2, 24.0, 23.9 and 24.1 volts respectively. The average of those five readings is

$$\text{Average} = \frac{23.5 + 24.2 + 24.0 + 23.9 + 24.1}{5} = 23.94$$

Now, suppose that on the following Monday the reading was found to be 24.2 volts. Then, the new 5-day average based on the last five days, Tuesday through Monday is

$$\frac{24.2 + 24.0 + 23.9 + 24.1 + 24.2}{5} = 24.08$$

We observe that the 5-day average has changed from 23.94 to 24.08 volts. In other words, the average has “moved” from 23.94 to 24.08 volts. Hence, the name **moving average**.

However, a more meaningful moving average can be obtained if we assign weights to each reading where the most recent reading carries the most weight. Thus, using a 5-day moving average we could take the reading obtained on the 5th day and multiply it by 5, the 4th day by 4, the 3rd day by 3, the 2nd day by 2, and the 1st day by 1. We could now add these numbers and divide the sum by the sum of the multipliers, i.e., $5+4+3+2+1=15$. Thus, the 5-day weighted moving average would be

$$\frac{1 \times 24.2 + 2 \times 24.0 + 3 \times 23.9 + 4 \times 24.1 + 5 \times 24.2}{15} = 24.09$$

and the value 24.09 is referred to as the **Weighted Moving Average (WMA)**.

An **Exponential Moving Average (EMA)** takes a percentage of the most recent value and adds in the previous value's exponential moving average times 1 minus that percentage. For instance, suppose we wanted a 10% EMA. We would take the most recent value and multiply it by 10% then add that figure to the previous value's EMA multiplied by the remaining percent, that is,

$$\text{Most Recent Value} \times 0.1 + \text{Previous Value's EMA} \times (1 - 0.1) \quad (1.39)$$

Alternately, we can use the following formula to determine the percentage to be used in the calculation:

$$\text{Exponential Percentage} = \frac{2}{\text{Time Periods} + 1} \quad (1.40)$$

For example, if we wanted a 20 period EMA, we would use

$$\frac{2}{20 + 1} = 9.52 \% \quad (1.41)$$

For the example* below, we will use the Simulink **Weighted Moving Average** block.

Example 1.20

The price of a particular security (stock) over a 5–day period is as follows:

77 80 82 85 90

where the last value is the most recent. We will create single–input / single–output (SISO) model with a Weighted Moving Average block to simulate the weighted moving average over this 5–day period.

For this example, we will represent the SISO output as follows:

$$y_1(k) = a_1u(k) + b_1u(k - 1) + c_1u(k - 2) + d_1u(k - 3) + e_1u(k - 4) \tag{1.42}$$

where

$$u(k) = 5/15 \quad u(k - 1) = 4/15 \quad u(k - 2) = 3/15 \quad u(k - 3) = 2/15 \quad u(k - 4) = 1/15 \tag{1.43}$$

The model is shown in Figure 1.1 where in the Function Block Parameters dialog box for the Weighted Moving Average block we have entered:

Weights:

[5/15 4/15 3/15 2/15 1/15]

Initial conditions:

[85 82 80 77]

Constant block – Output scaling value: [1.25 3]

Weighted Moving Average block – Parameter data types: sfix(16), Parameter scaling: 2⁻⁴

Signal data types: sfix(16), Parameter scaling: 2⁻⁶

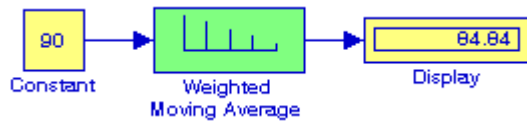


Figure 1.1. Model for Example 1.20

* This example can be skipped without loss of continuity. For an introduction to Simulink, please refer to Introduction to Simulink with Engineering Applications, ISBN 0-9744239-7-1.

1.8 Interpolation and Extrapolation

Let us consider the points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ shown on Figure 1.2 where, in general, the designation $P_n(x_n, y_n)$ is used to indicate the intersection of the lines parallel to the x – axis and y – axis .

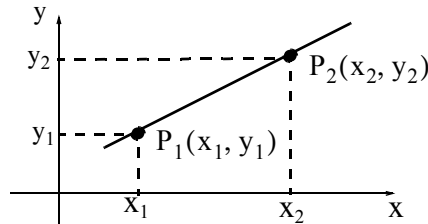


Figure 1.2. Graph to define interpolation and extrapolation

Let us assume that the values x_1 , y_1 , x_2 , and y_2 are known. Next, let us suppose that a known value x_i lies between x_1 and x_2 and we want to find the value y_i that corresponds to the known value of x_i . We must now make a decision whether the unknown value y_i lies on the straight line segment that connects the points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ or not. In other words, we must decide whether the new point $P_i(x_i, y_i)$ lies on the line segment P_1P_2 , above it, or below it. *Linear interpolation* implies that the point $P_i(x_i, y_i)$ lies on the segment P_1P_2 between points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$, and *linear extrapolation* implies that the point $P_i(x_i, y_i)$ lies to the left of point $P_1(x_1, y_1)$ or to the right of $P_2(x_2, y_2)$ but on the same line segment which may be extended either to the left or to the right. Interpolation and extrapolation methods other than linear are discussed in numerical analysis* textbooks where polynomials are used very commonly as functional forms. Our remaining discussion and examples will be restricted to linear interpolation.

Linear interpolation and extrapolation can be simplified if the first calculate the *slope* of the straight line segment. The *slope*, usually denoted as m , is the *rise* in the vertical (y -axis) direction over the *run* in the abscissa (x -axis) direction. Stated mathematically, the slope is defined as

$$\text{slope} = m = \frac{\text{rise}}{\text{run}} = \frac{y_2 - y_1}{x_2 - x_1} \quad (1.44)$$

Example 1.21

Compute the slope of the straight line segment that connects the points $P_1(3, 2)$ and $P_2(7, 4)$

* Refer, for example, to *Numerical Analysis Using MATLAB and Spreadsheets*, ISBN 0-9709511-1-6.

shown in Figure 1.3.

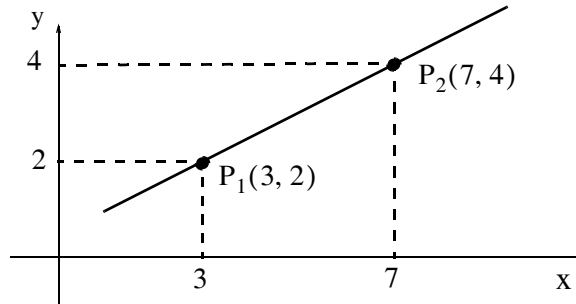


Figure 1.3. Graph for example 1.21

Solution:

$$\text{slope} = m = \frac{\text{rise}}{\text{run}} = \frac{4-2}{7-3} = \frac{2}{4} = 0.5$$

In the graph of Figure 1.3, if we know the value x_i and we want to find the value y_i , we use the formula

$$\frac{y_i - y_1}{x_i - x_1} = \text{slope} \tag{1.45}$$

or

$$y_i = \text{slope} \times (x_i - x_1) + y_1 = m(x_i - x_1) + y_1 \tag{1.46}$$

and if we know the value of y_i and we want to find the value x_i , we solve (1.46) for x_i and we obtain

$$x_i = \frac{1}{\text{slope}} \times (y_i - y_1) + x_1 = \frac{1}{m} \times (y_i - y_1) + x_1 \tag{1.47}$$

We observe that, if in (1.47) we let $x_1 = 0$, $x_i = x$, $y_i = y$, and $y_1 = b$, we obtain the equation of any straight line

$$y = mx + b \tag{1.48}$$

which will be introduced on Chapter 2.

Example 1.22

Given the graph of Figure 1.4, perform linear interpolation to compute the value of y that corresponds to the value $x = 7.5$

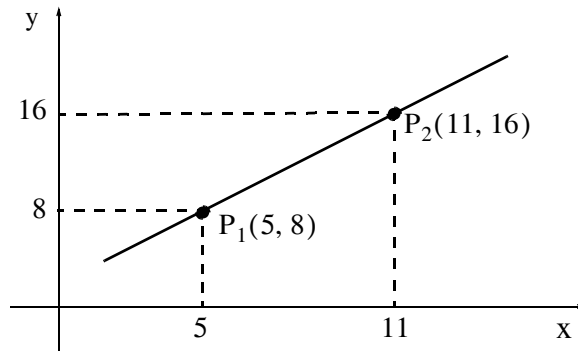


Figure 1.4. Graph for example 1.22

Solution:

Using (1.47) we get

$$y|_{x=7.5} = \text{slope} \times (x_i - x_1) + y_1 = \frac{16-8}{11-5} \times (7.5-5) + 8 = \frac{8}{6} \times 2.5 + 8 = \frac{10}{3} + 8 = 11.33$$

Note 2.4

The smaller the interval, the better the approximation will be obtained by linear interpolation.

Note 2.5

It is highly recommended that the data points are plotted so that we can assess how reasonable our approximation will be.

Note 2.6

We must exercise good judgement when we use linear interpolation since we may obtain unrealistic values. As an example, let us consider the following table where x represents the indicated numbers and y represents the square of x , that is, $y = x^2$.

x	y
1	1
.....
6	36

If we use linear interpolation to find the square of 5 with the data of the above table we will find that

$$y_{x=5} = \text{slope} \times (x_i - x_1) + y_1 = \frac{36-1}{6-1} \times (5-1) + 1 = \frac{35}{5} \times 4 + 1 = 29$$

and obviously this is gross error since the square of 5 is 25, not 29.

1.9 Infinite Sequences and Series

An infinite sequence is a function whose domain is the set of positive integers. For example, when x is successively assigned the values 1, 2, 3, ..., the function defined as

$$f(x) = \frac{1}{1+x}$$

yields the infinite sequence $1/2, 1/3, 1/4, \dots$ and so on. This sequence is referred to as infinite sequence to indicate that there is no last term. We can create a sequence $\{s_n\}$ by addition of numbers. Let us suppose that the numbers to be added are

$$u_1, u_2, u_3, \dots, u_n, \dots$$

We let

$$\begin{aligned} s_1 &= u_1 \\ s_2 &= u_1 + u_2 \\ &\dots \\ s_n &= u_1 + u_2 + u_3 + \dots + u_n = \sum_{k=1}^n u_k \end{aligned} \tag{1.49}$$

An expression such as (1.49) is referred to as an *infinite series*. There are many forms of infinite series with practical applications. In this text, we will discuss only the arithmetic series, geometric series, and harmonic series.

1.10 Arithmetic Series

An *arithmetic series* (or *arithmetic progression*) is a sequence of numbers such that each number differs from the previous number by a constant amount, called the *common difference*.

If a_1 is the first term, a_n is the n th term, d is the common difference, n is the number of terms, and s_n is the sum of n terms, then

$$a_n = a_1 + (n-1)d \tag{1.50}$$

$$s_n = \frac{n}{2}(a_1 + a_n) \tag{1.51}$$

$$s_n = \frac{n}{2}[2a_1 + (n-1)d] \tag{1.52}$$

The expression

$$\sum_{k=1}^n k = \frac{1}{2}n(n+1) \quad (1.53)$$

is known as the *sum identity*.

Example 1.23

Compute the sum of the integers from 1 to 100

Solution:

Using the sum identity, we get

$$\sum_{k=1}^{100} k = \frac{1}{2}100(100+1) = 50 \times 101 = 5050$$

1.11 Geometric Series

A *geometric series* (or *geometric progression*) is a sequence of numbers such that each number bears a constant ratio, called the *common ratio*, to the previous number.

If a_1 is the first term, a_n is the n th term, r is the common ratio, n is the number of terms, and s_n is the sum of n terms, then

$$a_n = a_1 r^{n-1} \quad (1.54)$$

and for $r \neq 1$,

$$\begin{aligned} s_n &= a_1 \frac{1-r^n}{1-r} \\ &= \frac{a_1 - ra_n}{1-r} \\ &= \frac{ra_n - a_1}{r-1} \end{aligned} \quad (1.55)$$

The first sum equation in (1.55) is derived as follows:

The general form of a geometric series is

$$a_1 + a_1 r + a_1 r^2 + a_1 r^3 + \dots + a_1 r^{n-1} + \dots \quad (1.56)$$

The sum of the first n terms of (1.56) is

$$s_n = a_1 + a_1r + a_1r^2 + a_1r^3 + \dots + a_1r^{n-1} \quad (1.57)$$

Multiplying both sides of (1.57) by r we obtain

$$rs_n = a_1r + a_1r^2 + a_1r^3 + \dots + a_1r^{n-1} + a_1r^n \quad (1.58)$$

If we subtract (1.58) from (1.57) we will find that all terms on the right side cancel except the first and the last leaving

$$(1-r)s_n = a_1(1-r^n) \quad (1.59)$$

Provided that $r \neq 1$, division of (1.59) by $1-r$ yields

$$s_n = a_1 \frac{1-r^n}{1-r} \quad (1.60)$$

It is shown in advanced mathematics textbooks that if $|r| < 1$, the geometric series

$$a_1 + a_1r + a_1r^2 + a_1r^3 + \dots + a_1r^{n-1} + \dots$$

converges (approaches a limit) to the sum

$$s_n = \frac{a_1}{1-r} \quad (1.61)$$

Example 1.24

A ball is dropped from x feet above a flat surface. Each time the ball hits the ground after falling a distance h , it rebounds a distance rh where $r < 1$. Compute the total distance the ball travels.

Solution:

The path and the distance the ball travels is shown on the sketch of Figure 1.5. The total distance s is computed by the geometric series

$$s = a_1 + 2a_1r + 2a_1r^2 + 2a_1r^3 + \dots \quad (1.62)$$

By analogy to equation of (1.61), the second and subsequent terms in (1.62) can be expressed as the sum of

$$\frac{2a_1r}{1-r} \quad (1.63)$$

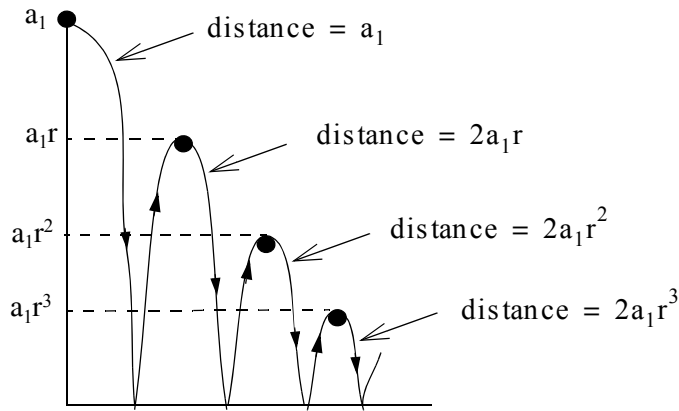


Figure 1.5. Sketch for Example 1.24

Adding the first term of (1.62) with (1.63) we form the total distance as

$$s = a_1 + \frac{2a_1 r}{1-r} = a_1 \frac{1+r}{1-r} \quad (1.64)$$

For example, if $a_1 = 6$ ft and $r = 2/3$, the total distance the ball travels is

$$a = 6 \times \frac{1+2/3}{1-2/3} = 30 \text{ ft}$$

1.12 Harmonic Series

A sequence of numbers whose reciprocals form an arithmetic series is called an *harmonic series* (or *harmonic progression*). Thus

$$\frac{1}{a_1}, \frac{1}{a_1 + d}, \frac{1}{a_1 + 2d}, \dots, \frac{1}{a_1 + (n-1)d}, \dots \quad (1.65)$$

where

$$\frac{1}{a_n} = \frac{1}{a_1 + (n-1)d} \quad (1.66)$$

forms an harmonic series.

Harmonic series provide quick answers to some very practical applications such as the example that follows.

Example 1.25

Suppose that we have a list of temperature numbers on a particular day and time for 50 years. Let

us assume the numbers are random, that is, the temperature on a particular day and time of one year has no relation to the temperature on the same day and time of any subsequent year.

The first year was undoubtedly a record year. In the second year, the temperature could equally likely have been more than, or less than, the temperature of the first year. So there is a probability of $1/2$ that the second year was a record year. The expected number of record years in the first two years of record-keeping is therefore $1 + 1/2$. For the third year, the probability is $1/3$ that the third observation is higher than the first two, so the expected number of record temperatures in three years is $1 + 1/2 + 1/3$. Continuing this line of reasoning leads to the conclusion that the expected number of records in the list of n observations is

$$1 + 1/2 + 1/3 + \dots + 1/n$$

For our example, $n = 50$ and thus the sum of the first 50 terms of harmonic series is 4.499 or 5 record years. Of course, this is the expected number of record temperatures, not the actual.

The partial sums of this harmonic series are plotted on Figure 1.6.

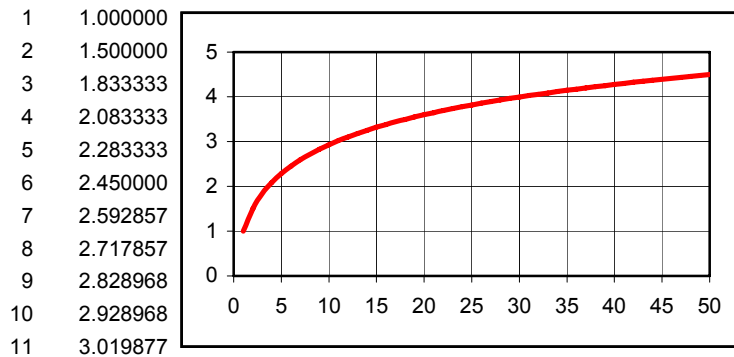


Figure 1.6. Plot for Example 1.25

1.13 Proportions

A *proportion* is defined as two ratios being equal to each other. If

$$\frac{a}{b} = \frac{c}{d} \tag{1.67}$$

where $a, b, c,$ and d are non-zero values, then

$$\frac{a + b}{b} = \frac{c + d}{d} \tag{1.68}$$

$$\frac{a - b}{b} = \frac{c - d}{d} \tag{1.69}$$

Chapter 1 Elementary Algebra

$$\frac{a-b}{a+b} = \frac{c-d}{c+d} \quad (1.70)$$

Consider a line segment with a length of one unit. Let us divide this segment into two unequal segments and let

x = shorter segment

$1 - x$ = longer segment

The *golden proportion* is the one where the ratio of the shorter segment to the longer segment, that is, $x/(1-x)$ is equal to the ratio of the longer segment to the whole segment, that is, $(1-x)/1$. In other words, the golden proportion states that

$$\frac{x}{(1-x)} = \frac{(1-x)}{1} \quad (1.71)$$

Solving for x we get

$$\begin{aligned}(1-x)^2 &= x \\ 1 - 2x + x^2 &= x \\ x^2 - 3x + 1 &= 0\end{aligned}$$

Solution of this quadratic equation yields

$$x_1, x_2 = \frac{3 \pm \sqrt{9-4}}{2} = \frac{3 \pm \sqrt{5}}{2} = \frac{3 \pm 2.236}{2}$$

and since x cannot be more than unity, we reject the largest value and we accept x as $x = 0.382$. Then, $1 - x = 0.618$ and thus the golden proportion is

$$\frac{0.382}{0.618} = \frac{0.618}{1} \quad (1.72)$$

The *golden rectangle* is defined as one whose length is 1 and its width is 0.618 as shown in Figure 1.7.

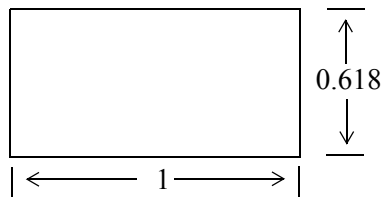


Figure 1.7. The golden rectangle

1.14 Summary

- *Algebra* is the branch of mathematics in which letters of the alphabet represent numbers or a set of numbers.
- *Equations* are equalities that indicate how some quantities are related to others.
- In algebra, the *order of precedence*, where 1 is the highest and 4 is the lowest, is as follows:
 1. Quantities inside parentheses ()
 1. Exponentiation
 3. Multiplication and Division
 4. Addition or Subtraction
- An *algebraic equation* is an equality that contains one or more unknown quantities, normally represented by the last letters of the alphabet such as x , y , and z .
- Solving an equation means finding the value of the unknown quantity that will make the equation an equality with no unknowns.
- The same number may be added to, or subtracted from both sides of an equation.
- Each side of an equation can be multiplied or divided by the same number. Division by zero must be avoided.
- The *laws of exponents* state that

$$a^m \cdot a^n = a^{m+n}$$

$$\frac{a^m}{a^n} = \begin{cases} a^{m-n} & \text{if } m > n \\ 1 & \text{if } m = n \\ \frac{1}{a^{n-m}} & \text{if } m < n \end{cases}$$

$$(a^m)^n = a^{mn}$$

- The n th root is defined as the inverse of the n th exponent, that is, if

$$b^n = a, \quad \text{then } b = \sqrt[n]{a}$$

- The *laws of logarithms* state that

$$\log_a(xy) = \log_a x + \log_a y$$

$$\log_a \left(\frac{x}{y} \right) = \log_a x - \log_a y$$

$$\log_a (x^n) = n \log_a x$$

- Logarithms consist of an integer, called the *characteristic* and an endless decimal called the *mantissa*.
- If the decimal point is located immediately to the right of the msd of a number, the characteristic is zero; if the decimal point is located after two digits to the right of the msd, the characteristic is 1; if after three digits, it is 2 and so on.
- If the decimal point is located immediately to the left of the msd of a number, the characteristic is -1 , if located two digits to the left of the msd, it is -2 , if after three digits, it is -3 and so on.
- The mantissa cannot be determined by inspection; it must be extracted from tables of common logarithms.
- Although the common logarithm of a number less than one is negative, it is written with a *negative* characteristic and a *positive* mantissa. This is because the mantissas in tables are given as positive numbers.
- Because mantissas are given in math tables as positive numbers, the negative sign is written above the characteristic. This is to indicate that the negative sign does not apply to the mantissa.
- A convenient method to find the characteristic of logarithms is to first express the given number in scientific notation; the characteristic then is the exponent.
- *Quadratic equations* are those that contain equations of second degree. The general form of a quadratic equation is

$$ax^2 + bx + c = 0$$

where a , b and c are real constants (positive or negative).

- The roots of a quadratic equations can be found from the formulas

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

- The quantity $b^2 - 4ac$ under the square root is called the *discriminant* of a quadratic equation.

- A *cubic equation* has the form

$$ax^3 + bx^2 + cx + d = 0$$

and higher degree equations have similar forms. Formulas and procedures for solving these equations are included in books of mathematical tables.

- The *arithmetic mean* is the value obtained by dividing the sum of a set of quantities by the number of quantities in the set. It is also referred to as the *average*. The arithmetic mean or simply mean, is denoted with the letter x with a bar above it, and is computed from the equation

$$\bar{x} = \frac{\sum x}{n}$$

where the symbol Σ stand for summation and n is the number of data, usually called *sample*. The mean can also be found with the Excel =AVERAGE function.

- The *median* of a sample is the value that separates the lower half of the data, from the upper half. To find the median, we arrange the values of the sample in increasing (ascending) order. If the number of the sample is odd, the median is in the middle of the list; if even, the median is the mean (average) of the two values closest to the middle of the list. The median can also be found with the Excel =MEDIAN function.
- The *mode* is the value in a sample that occurs most often. If, in a sequence of numbers, no number appears two or more times, the sample has no mode. The mode, if it exists, may or may not be unique. If two such values exist, we say that the sample is *bimodal*, and if three values exist, we call it *trimodal*. The mode can also be found with the Excel =MODE function.
- *Linear interpolation* implies that the point $P_i(x_i, y_i)$ lies on the segment P_1P_2 between points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ on a straight line
- *Linear extrapolation* implies that the point $P_i(x_i, y_i)$ lies to the left of point $P_1(x_1, y_1)$ or to the right of $P_2(x_2, y_2)$ but on the same line segment which may be extended either to the left or to the right.
- An infinite sequence is a function whose domain is the set of positive integers.
- An expression such as the one shown below is referred to as an *infinite series*.

$$s_n = u_1 + u_2 + u_3 + \dots + u_n = \sum_{k=1}^n u_k$$

- An *arithmetic series* (or *arithmetic progression*) is a sequence of numbers such that each number differs from the previous number by a constant amount, called the *common difference*.

Chapter 1 Elementary Algebra

- A *geometric series* (or *geometric progression*) is a sequence of numbers such that each number bears a constant ratio, called the *common ratio*, to the previous number.
- A sequence of numbers whose reciprocals form an arithmetic series is called an *harmonic series* (or *harmonic progression*).
- A *proportion* is defined as two ratios being equal to each other.
- The *golden proportion* states that

$$\frac{x}{(1-x)} = \frac{(1-x)}{1}$$

- The *golden rectangle* is defined as one whose length is 1 and its width is 0.618.

1.15 Exercises

1. In the following equations V , I , R and P are the unknowns. Solve the following equations, that is, find the value of each unknown.

$$\text{a. } R + 50 = 150 \quad \text{b. } V - 25 = 15 \quad \text{c. } 5I - 10 = 45$$

$$\text{d. } \frac{1}{R} = 0.1 \quad \text{e. } \sqrt{5P} = 10 \quad \text{f. } (100 + R)10^{-3} = 1$$

2. The notation $\sqrt[r]{x}$ stands for the r th root of a number x . Thus $\sqrt[4]{256}$ stands for the fourth root of 256, $\sqrt[5]{3125}$ for the fifth root of 3125 and so on. Solve the following equations.

$$\text{a. } \sqrt[6]{262144} = x \quad \text{b. } \sqrt[9]{1000000000} = y$$

3. Mathematical tables show that the mantissa of the common logarithm of 0.000417 is 0.62014. However, Excel displays the number -3.37986 rounded to five decimal places when we use the function `=LOG(0.000417)`. Is there something wrong with the formula that we used, or is Excel giving us the wrong answer? Explain.

4. The formula below computes the terms of an annuity with an initial payment and periodic payments.

$$n = \frac{\log\left(\frac{1 + fv \times (\text{int}) / (\text{pmt})}{1 + pv \times (\text{int}) / (\text{pmt})}\right)}{\log(1 + \text{int})}$$

fv = future value, that is, the cash value we want to receive after the last payment is made

int = interest rate per period

pmt = payment made at the end of each period

pv = initial payment.

Use this formula to compute the number of months required to accumulate \$1,000,000 by making an initial payment of \$10,000 now and monthly payments of \$2,000 into a savings account paying 6% annual interest compounded monthly. Verify your answer with Excel's `NPER` function remembering that the initial and monthly payments are cash outflows and, as such, must be entered as negative numbers. However, in the formula above, these must be entered as positive values.

5. Solve the following quadratic equations.

$$\text{a. } x^2 + 5x + 6 \quad \text{b. } x^2 - 6x + 9$$

Chapter 1 Elementary Algebra

6. Twenty houses were sold for the prices listed below where, for simplicity, the dollar (\$) sign has been omitted.

547,000	670,000	458,000	715,000	812,000
678,000	595,000	760,000	490,000	563,000
805,000	715,000	635,000	518,000	645,000
912,000	873,000	498,000	795,000	615,000

Compute the mean, median, and mode.

1.16 Solutions to End-of-Chapter Exercises

Dear Reader:

The remaining pages on this chapter contain the solutions to the exercises.

You must, for your benefit, make an honest effort to find the solutions to the exercises without first looking at the solutions that follow. It is recommended that first you go through and work out those you feel that you know. For the exercises that you are uncertain, review this chapter and try again. Refer to the solutions as a last resort and rework those exercises at a later date.

You should follow this practice with the rest of the exercises of this book.

Chapter 1 Elementary Algebra

1.
 - a. $R + 50 = 150$, $R + 50 - 50 = 150 - 50$, $R = 50$
 - b. $V - 25 = 15$, $V - 25 + 25 = 15 + 25$, $V = 40$
 - c. $5I - 10 = 45$, $5I - 10 + 10 = 45 + 10$, $5I = 55$, $(5I)/5 = 55/5$, $I = 11$
 - d. $\frac{1}{R} = 0.1$, $\frac{1}{R} \times R = 0.1 \times R$, $1 = 0.1 \times R$, $1/0.1 = R$, $R = 10$
 - e. $\sqrt{5P} = 10$, $(5P)^{(1/2)} = 10$, $(5P)^{(1/2) \times 2} = 10^2$, $5P = 100$, $(5P)/5 = 100/5$, $P = 20$
 - f. $(100 + R)10^{-3} = 1$, $(100 + R)10^{-3} \times 10^3 = 1 \times 10^3$, $(100 + R)10^0 = 1000$,
 $(100 + R) \times 1 = 1000$, $100 + R - 100 = 1000 - 100$, $R = 900$

2.
 - a. $\sqrt[6]{262144} = x$, $262144^{(1/6)} = x$, and in Excel we enter the formula `=262144^(1/6)` which yields 8 and thus $x = 8$
 - b. $\sqrt[9]{1000000000} = y$, $1000000000^{(1/9)} = y$, and in Excel we enter the formula `=1000000000^(1/9)` which yields 10 and thus $y = 10$

3. Mathematical tables show only the mantissa which is the fractional part of the logarithm, and as explained in Section 1.4, mantissas in tables are given as positive numbers. Review Example 1.13 to verify that the answer displayed by Excel is correct.

4. We construct the spreadsheet below.

	A4	fx =LOG((1+D2*A2/B2)/(1+C2*A2/B2))/LOG(1+A2)					
	A	B	C	D	E	F	G
1	int	pmt	pv	fv			
2	0.005	2000	10000	1000000			
3							
4	246.23						

and in Cell A4 we type the formula `=LOG((1+D2*A2/B2)/(1+C2*A2/B2))/LOG(1+A2)`

This yields 246.23 months.

The same result is obtained with Excel's NPER function as shown below.

`=NPER(0.06/12,-2000,-10000,1000000,0)`

yields 246.23 months.

5.

a.

$$x_1 = \frac{-5 + \sqrt{5^2 - 4 \times 1 \times 6}}{2 \times 1} = \frac{-5 + \sqrt{25 - 24}}{2} = \frac{-5 + 1}{2} = -2$$

$$x_2 = \frac{-5 - \sqrt{5^2 - 4 \times 1 \times 6}}{2 \times 1} = \frac{-5 - \sqrt{25 - 24}}{2} = \frac{-5 - 1}{2} = -3$$

b.

$$x_1 = \frac{6 + \sqrt{6^2 - 4 \times 1 \times 9}}{2 \times 1} = \frac{6 + \sqrt{36 - 36}}{2} = \frac{6 + 0}{2} = 3$$

$$x_2 = \frac{6 - \sqrt{6^2 - 4 \times 1 \times 9}}{2 \times 1} = \frac{6 - \sqrt{36 - 36}}{2} = \frac{6 - 0}{2} = 3$$

6.

We find the mean, median, and mode using the procedures of Examples 1.17, 1.18, and 1.19 respectively. Then we verify the answers with the spreadsheet below where the values were entered in Cells A1 through A20 in the order given, and they were copied in B1 through B20 and were sorted.

The mean was found with the formula =AVERAGE(B1:B20) entered in D1.

The median was found with the formula =MEDIAN(B1:B20) entered in D2.

The mode was found with the formula =MODE(B1:B20) entered in D3.

Chapter 1 Elementary Algebra

	A	B	C	D
1	547,000	458,000	Mean=	664,950
2	678,000	490,000	Median=	657,500
3	805,000	498,000	Mode=	715,000
4	912,000	518,000		
5	670,000	547,000		
6	595,000	563,000		
7	715,000	595,000		
8	873,000	815,000		
9	458,000	635,000		
10	760,000	645,000		
11	635,000	670,000		
12	498,000	678,000		
13	715,000	715,000		
14	490,000	715,000		
15	518,000	760,000		
16	795,000	795,000		
17	812,000	805,000		
18	583,000	812,000		
19	645,000	873,000		
20	615,000	912,000		

This chapter is a continuation of Chapter 1. It is intended for readers who need the knowledge for more advanced topics in mathematics. It serves as a prerequisite for probability and statistics. Readers with a strong mathematical background may skip this chapter. Others will find it useful, as well as a convenient source for review.

2.1 Systems of Two Equations

Quite often, we are faced with problems that seem impossible to solve because they involve two or more unknowns. However, a solution can be found if we write two or more linear independent equations* with the same number of unknowns, and solve these simultaneously. Before we discuss generalities, let us consider a practical example.

Example 2.1

Jeff Simpson plans to start his own business, manufacturing and selling bicycles. He wants to compute the *break-even* point; this is defined as the point where the revenues are equal to costs. In other words, it is the point where Jeff neither makes nor loses money.

Jeff estimates that his *fixed costs* (rent, electricity, gas, water, telephone, insurance etc.), would be around \$1,000 per month. Other costs such as material, production and payroll are referred to as *variable costs* and will increase linearly (in a straight line fashion). Preliminary figures show that the variable costs for the production of 500 bicycles will be \$9,000 per month. Then,

$$\text{Total Costs} = \text{Fixed} + \text{Variable} = 1000 + 9000 = 10,000 \quad (2.1)$$

Let us plot cost (*y-axis*), versus number of bicycles sold (*x-axis*) as shown in Figure 2.1.

* Linear equations are those in which the unknowns, such as x and y , have exponent 1. Thus, the equation $7x + 5y = 24$ is linear since the exponents of x and y are both unity. However, the equation $2x^2 + y = 8$ is non-linear because the exponent of x is 2. By independent equations we mean that these they do not depend on each other. For instance, the equations $3x + 2y = 5$ and $6x + 4y = 10$ are not independent since the second can be formed from the first by multiplying both sides by 2.

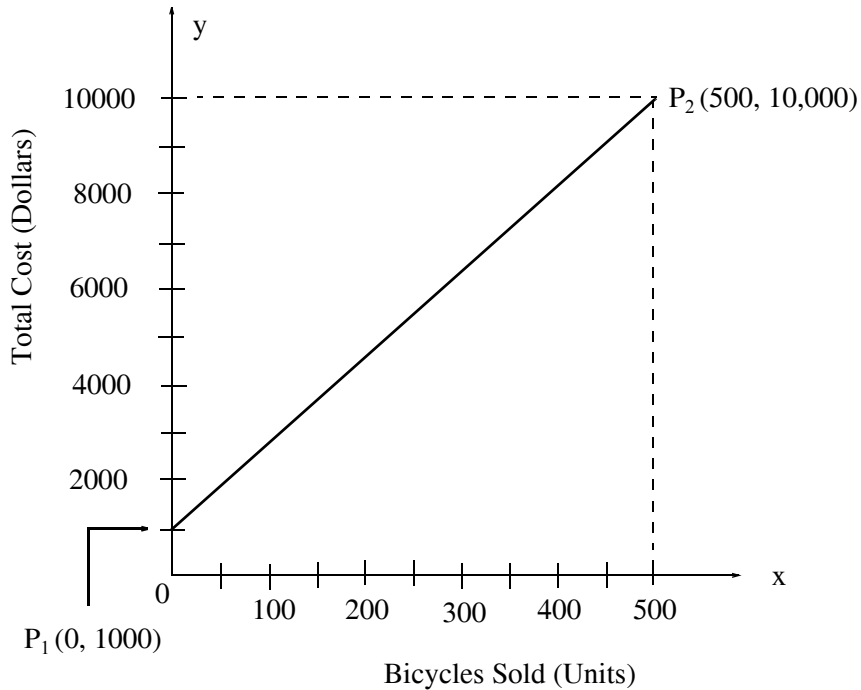


Figure 2.1. Total cost versus bicycles sold.

In Figure 2.1, the x -axis is the *abscissa* and the y -axis is the *ordinate*. Together, these axes constitute a *Cartesian coordinate system*; it is also mentioned in Appendix A. The straight line drawn from point P_1 to point P_2 , is a graphical presentation of the total costs for the production of the bicycles. It starts at the \$1,000 point because this represents a fixed cost; it occurs even though no bicycles are sold. We denote this point as $P_1(0, 1000)$ where the first number 0, within the parentheses, denotes the x -axis value at this point, and the second number (1000) denotes the y -axis value at that point. For simplicity, the dollar sign has been omitted. Thus, we say that the coordinates of point P_1 are $(0, 1000)$. Likewise, the coordinates of the end point P_2 are denoted as $P_2(500, 10000)$ since the total costs to produce 500 bicycles is \$10,000. We will now derive the equation that describes this straight line.

In general, a straight line is represented by the equation

$$y = mx + b \tag{2.2}$$

where m is the *slope*, x is the *abscissa*, y is the *ordinate*, and b is the *y-intercept* of the straight line, that is, the point where the straight line crosses the y -axis.

As we indicated in Chapter 1, the *slope* m is the *rise* in the vertical (y -axis) direction over the *run* in the *abscissa* (x -axis) direction. We recall from Chapter 1 that the slope is defined as

$$m = \frac{\text{rise}}{\text{run}} = \frac{y_2 - y_1}{x_2 - x_1} \quad (2.3)$$

In Figure 2.1, $y_2 = 10000$, $y_1 = 1000$, $x_2 = 500$ and $x_1 = 0$. Therefore, the slope is

$$m = \frac{\text{rise}}{\text{run}} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{10000 - 1000}{500 - 0} = \frac{9000}{500} = 18$$

and, by inspection, the y-intercept is 1000. Then, in accordance with (2.2), the equation of the line that represents the total costs is

$$y = 18x + 1000 \quad (2.4)$$

It is customary, and convenient, to show the unknowns of an equation on the left side and the known values on the right side. Then, (2.4) is written as

$$y - 18x = 1000^* \quad (2.5)$$

This equation has two unknowns x and y and thus, no unique solution exists; that is, we can find an infinite number of x and y combinations that will satisfy this equation. We need a second equation, and we will find it by making use of additional facts. These are stated below.

Jeff has determined that if he sells 500 bicycles at \$25 each, he will generate a revenue of $\$25 \times 500 = \12500 . This fact can be represented by another straight line which is added to the previous figure. It is shown in Figure 2.2.

The new line starts at $P_3(0, 0)$ because there will be no revenue if no bicycles are sold. It ends at $P_4(500, 12500)$ which represents the condition that Jeff will generate \$12500 when 500 bicycles are sold.

The intersection of the two lines shown as a small circle, establishes the *break-even point* and this is what Jeff is looking for. The projection of the break-even point on the x -axis, shown as a broken line, indicates that approximately 140 bicycles must be sold just to break even. Also, the projection on the y -axis shows that at this break-even point, the generated revenue is approximately \$3,500. This procedure is known as *graphical solution*. A graphical solution gives *approximate* values.

We will now proceed to find the so-called *analytical solution*. This solution will produce *exact* values.

* We assume that the reader has become proficient with the properties which we discussed in the previous chapter, and henceforth, the details for simplifying or rearranging an equation will be omitted.

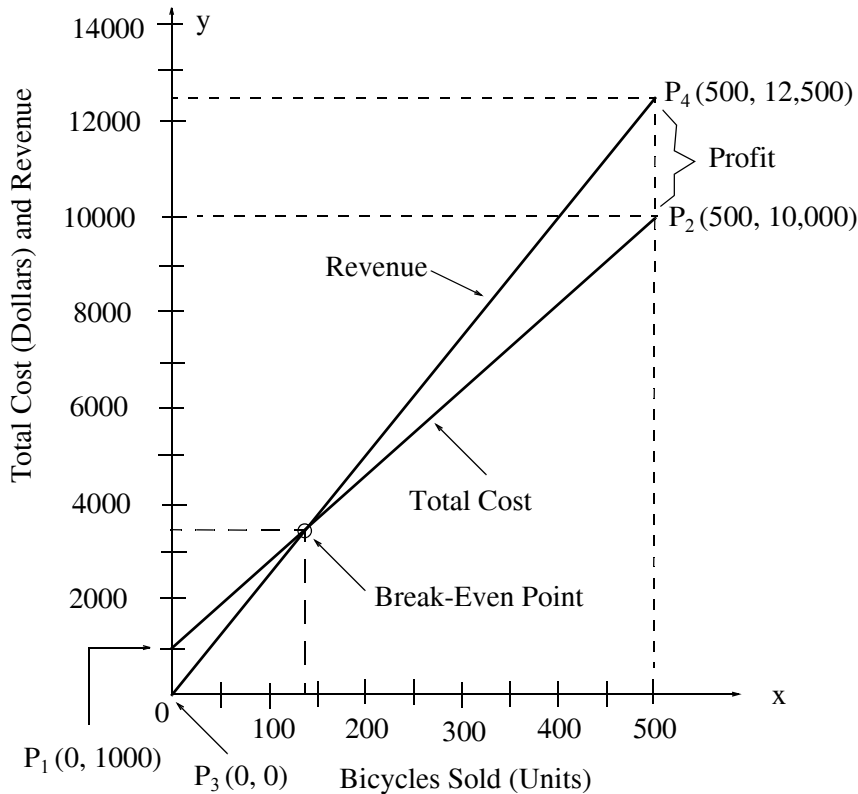


Figure 2.2. Graph showing the intersection of two straight lines.

As stated above, we need a second equation. This is obtained from the straight line in Figure 2.2 which represents the revenue. To obtain the equation of this line, we start with the equation of (2.2) which is repeated here for convenience.

$$y = mx + b \quad (2.6)$$

The slope m of the revenue line is

$$m = \frac{\text{rise}}{\text{run}} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{12500 - 0}{500 - 0} = \frac{125}{5} = 25 \quad (2.7)$$

and, by inspection, b , that is, the y -intercept, is 0. Therefore, the equation that describes the revenue line is

$$y = 25x \quad (2.8)$$

By grouping equations (2.5) and (2.8), we get the system of equations

$$\begin{aligned} y - 18x &= 1000 \\ y &= 25x \end{aligned}$$

(2.9)

Now, we must solve the equations of (2.9) simultaneously. This means that we must find unique values for x and y , so that both equations of (2.9) will be satisfied at the same time. An easy way to solve these equations is by substitution of the second into the first. This eliminates y from the first equation which now becomes

$$25x - 18x = 1000 \quad (2.10)$$

or

$$7x = 1000$$

or

$$x = \frac{1000}{7}^* \quad (2.11)$$

To find the unknown y , we substitute (2.11) into either the first, or the second equation of (2.9). Thus, by substitution into the second equation, we get

$$y = 25 \times \frac{1000}{7} = \frac{25000}{7} \quad (2.12)$$

Therefore, the *exact* solutions for x and y are

$$\begin{aligned} x &= \frac{1000}{7} \\ y &= \frac{25000}{7} \end{aligned} \quad (2.13)$$

Since no further substitutions are needed, we simplify the values of (2.13) by division and we get $x = 1000/7 = 142.86$ and $y = 25000/7 = 3571.40$. Of course, bicycles must be expressed as whole numbers and therefore, the break-even point for Jeff's business is 143 bicycles which, when sold, will generate a revenue of

$$y = \$25 \times 143 = \$3575$$

We recall that for the break-even point, the graphical solution produced approximate values of 140 bicycles which would generate a revenue of \$3,500. Although the graphical solution is not as accurate as the analytical solution, it nevertheless provides us with some preliminary information, which can be used as a check for the answers found using the analytical solution.

Note 2.1

The system of equations of (2.9) was solved by the so-called *substitution method*, also known as *Gauss's elimination method*. In Section 2.3, we will discuss this method in more detail, and two additional methods, the *solution by matrix inversion*, and the *solution by Cramer's rule*.

* It is not advisable to divide 1000 by 7 at this time. This division will produce an irrational (endless) number which can only be approximated and, when substituted to find the unknown y , it will produce another approximation. Usually, we perform the division at the last step, that is, after the solutions are completed.

2.2 Systems of Three Equations

Systems of three or more equations also appear in practical applications. We will now consider another example consisting of three equations with three unknowns. It is imperative to remember that *we must have the same number of equations as the number of unknowns*.

Example 2.2

In an automobile dealership, the most popular passenger cars are *Brand A*, *B* and *C*. Because buyers normally bargain for the best price, the sales price for each brand is not the same. Table 2.1 shows the sales and revenues for a 3-month period.

TABLE 2.1 Sales and Revenues for Example 2.2

Month	Brand A	Brand B	Brand C	Revenue
1	25	60	50	\$2,756,000
2	30	40	60	2,695,000
3	45	53	58	3,124,000

Compute the average sales price for each of these brands of cars.

Solution:

In this example, the unknowns are the average sales prices for the three brands of automobiles. Let these be denoted as x for *Brand A*, y for *Brand B*, and z for *Brand C*. Then, the sales for each of the three month period can be represented by the following system of equations*.

$$\begin{aligned}25x + 62y + 54z &= 2756000 \\28x + 42y + 58z &= 2695000 \\45x + 53y + 56z &= 3124000\end{aligned}\tag{2.14}$$

The next task is to solve these equations simultaneously, that is, find the values of the unknowns x , y , and z at the same time. In the next section, we will discuss matrix theory and methods for solving systems of equations of this type. We digress to discuss the simultaneous solution methods of a system of equations such as that of (2.14) above, and will return to this example at the conclusion of the next section, Page 2–24.

* The reader should observe that these equations are consistent with each side, that is, both the left and the right side in all equations represent dollars. This observation should be made always when writing equations, since it is a very common mistake to add and equate unrelated quantities.

2.3 Matrices and Simultaneous Solution of Equations

A *matrix** is a rectangular array of numbers such as those shown below.

$$\begin{bmatrix} 2 & 3 & 7 \\ 1 & -1 & 5 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 3 & 1 \\ -2 & 1 & -5 \\ 4 & -7 & 6 \end{bmatrix}$$

In general form, a matrix A is denoted as

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} \quad (2.15)$$

The numbers a_{ij} are the *elements* of the matrix where the index i indicates the row and j indicates the column in which each element is positioned. Thus, a_{43} represents the element positioned in the fourth row and third column.

A matrix of m rows and n columns is said to be of $m \times n$ *order matrix*.

If $m = n$, the matrix is said to be a *square matrix of order* m (or n). Thus, if a matrix has five rows and five columns, it is said to be a square matrix of *order 5*.

In a square matrix, the elements $a_{11}, a_{22}, a_{33}, \dots, a_{nn}$ are called the *diagonal elements*. Alternately, we say that the elements $a_{11}, a_{22}, a_{33}, \dots, a_{nn}$ are located on the *main diagonal*.

A matrix in which every element is zero, is called a *zero matrix*.

Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are said to be equal, that is, $A = B$, if, and only if $a_{ij} = b_{ij}$ for $i = 1, 2, 3, \dots, m$ and $j = 1, 2, 3, \dots, n$.

Two matrices are said to be *conformable for addition (subtraction)* if they are of the same order, that is, both matrices must have the same number of rows and columns.

If $A = [a_{ij}]$ and $B = [b_{ij}]$ are conformable for addition (subtraction), their sum (difference) will be another matrix C with the same order as A and B , where each element of C represents the sum

* To some readers, this material may seem too boring and uninteresting. However, it should be read at least once to become familiar with matrix terminology. We will use it for some practical examples in later chapters.

Chapter 2 Intermediate Algebra

(difference) of the corresponding elements of A and B , that is, $C = A \pm B = [a_{ij} \pm b_{ij}]$

Example 2.3

Compute $A + B$ and $A - B$ given that

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

Solution:

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

and

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

If k is any scalar (a positive or negative number) and not $[k]$ which is a 1×1 matrix, then multiplication of a matrix A by the scalar k is the multiplication of every element of A by k .

Example 2.4

Multiply the matrix

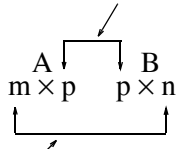
$$A = \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix} \text{ by } k = 5$$

Solution:

$$k \cdot A = 5 \times \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 5 \times 1 & 5 \times (-2) \\ 5 \times 2 & 5 \times 3 \end{bmatrix} = \begin{bmatrix} 5 & -10 \\ 10 & 15 \end{bmatrix}$$

Two matrices A and B are said to be *conformable for multiplication* $A \cdot B$ in that order, only when the number of columns of matrix A is equal to the number of rows of matrix B . The product $A \cdot B$, which is not the same as the product $B \cdot A$, is conformable for multiplication only if A is an $m \times p$ and matrix B is an $p \times n$ matrix. The product $A \cdot B$ will then be an $m \times n$ matrix. A convenient way to determine whether two matrices are conformable for multiplication, and the dimension of the resultant matrix, is to write the dimensions of the two matrices side-by-side as shown below.

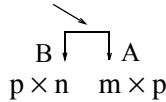
Shows that A and B are conformable for multiplication



Indicates the dimension of the product $A \cdot B$

For the product $B \cdot A$ we have:

Here, B and A are not conformable for multiplication since $n \neq m$



For matrix multiplication, the operation is row by column. Thus, to obtain the product $A \cdot B$, we multiply each element of a row of A by the corresponding element of a column of B and then we add these products.

Example 2.5

It is given that

$$C = [2 \ 3 \ 4] \text{ and } D = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

Compute the products $C \cdot D$ and $D \cdot C$

Solution:

The dimensions of matrices C and D are 1×3 and 3×1 respectively; therefore, the product $C \cdot D$ is feasible, and results in a 1×1 matrix as shown below.

$$C \cdot D = [2 \ 3 \ 4] \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = [(2) \cdot (1) + (3) \cdot (-1) + (4) \cdot (2)] = [7]$$

The dimensions for D and C are 3×1 and 1×3 respectively and therefore, the product $D \cdot C$ is also feasible. However, multiplication of these results in a 3×3 matrix as shown below.

Chapter 2 Intermediate Algebra

$$D \cdot C = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \begin{bmatrix} 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} (1) \cdot (2) & (1) \cdot (3) & (1) \cdot (4) \\ (-1) \cdot (2) & (-1) \cdot (3) & (-1) \cdot (4) \\ (2) \cdot (2) & (2) \cdot (3) & (2) \cdot (4) \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \\ -2 & -3 & -4 \\ 4 & 6 & 8 \end{bmatrix}$$

Division of one matrix by another is not defined. However, an equivalent operation exists. It will become apparent later when we discuss the inverse of a matrix.

An *identity matrix* I is a square matrix where all the elements on the main diagonal are ones, and all other elements are zero. Shown below are 2×2 , 3×3 , and 4×4 identity matrices.

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

Let matrix A be defined as the square matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} \quad (2.16)$$

then, the *determinant of* A , denoted as $\det A$, is defined as

$$\det A = a_{11}a_{22}a_{33}\dots a_{nn} + a_{12}a_{23}a_{34}\dots a_{n1} + a_{13}a_{24}a_{35}\dots a_{n2} + \dots - a_{n1}\dots a_{22}a_{13}\dots - a_{n2}\dots a_{23}a_{14} - a_{n3}\dots a_{24}a_{15} - \dots \quad (2.17)$$

The determinant of a square matrix of order n is referred to as *determinant of order* n .

Let A be a *determinant of order* 2, that is,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad (2.18)$$

Then, by (2.17), (2.18) reduces to

$$\det A = a_{11} a_{22} - a_{21} a_{12} \quad (2.19)$$

Example 2.6

Compute $\det A$ and $\det B$ given that

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

Solution:

Using (2.19), we obtain

$$\begin{aligned} \det A &= 1 \cdot 4 - 3 \cdot 2 = 4 - 6 = -2 \\ \det B &= 2 \cdot 0 - 2 \cdot (-1) = 0 - (-2) = 2 \end{aligned}$$

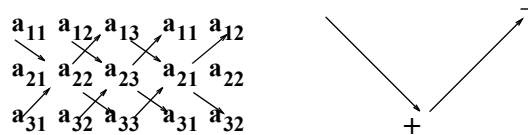
Let A be a *determinant of order 3*, that is,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

then,

$$\begin{aligned} \det A &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{11}a_{22}a_{33} \\ &\quad - a_{11}a_{22}a_{33} - a_{11}a_{22}a_{33} - a_{11}a_{22}a_{33} \end{aligned} \tag{2.20}$$

A convenient method to evaluate the determinant of order 3 is to write the first two columns to the right of the 3×3 matrix, and add the products formed by the diagonals from upper left to lower right; then subtract the products formed by the diagonals from lower left to upper right as shown below. When this is done properly, we obtain (2.20) above.



This method works only with second and third order determinants. To evaluate higher order determinants, we must first compute the *cofactors*; these will be defined shortly.

Example 2.7

Compute $\det A$ and $\det B$ given that

Chapter 2 Intermediate Algebra

$$A = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & -3 & -4 \\ 1 & 0 & -2 \\ 0 & -5 & -6 \end{bmatrix}$$

Solution:

$$\det A = \begin{array}{ccccccc} & 2 & 3 & 5 & 2 & 3 & \\ & 1 & 0 & 1 & 1 & 0 & \\ & 2 & 1 & 0 & 2 & 1 & \end{array}$$

or

$$\det A = (2 \times 0 \times 0) + (3 \times 1 \times 1) + (5 \times 1 \times 1) - (2 \times 0 \times 5) - (1 \times 1 \times 2) - (0 \times 1 \times 3) = 11 - 2 = 9$$

Likewise,

$$\det B = \begin{array}{ccccccc} & 2 & -3 & -4 & 2 & -3 & \\ & 1 & 0 & -2 & 1 & -2 & \\ & 0 & -5 & -6 & 2 & -6 & \end{array}$$

or

$$\det B = [(2 \times 0 \times (-6)) + [(-3) \times (-2) \times 0] + [(-4) \times 1 \times (-5)]] - [0 \times 0 \times (-4)] - [(-5) \times (-2) \times 2] - [(-6) \times 1 \times (-3)] = 20 - 38 = -18$$

Let matrix A be defined as the square matrix of order n as shown below.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}$$

If we remove the elements of its i th row and j th column, the determinant of the remaining $(n - 1)$ square matrix is called a *minor of determinant A*, and it is denoted as $[M_{ij}]$.

The signed minor $(-1)^{i+j}[M_{ij}]$ is called the *cofactor* of a_{ij} , and it is denoted as α_{ij} .

Example 2.8

Compute the minors $[M_{11}]$, $[M_{12}]$, $[M_{13}]$, and the cofactors α_{11} , α_{12} , and α_{13} given that

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Solution:

$$[M_{11}] = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, [M_{12}] = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, [M_{13}] = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

and

$$\alpha_{11} = (-1)^{1+1}[M_{11}] = [M_{11}], \quad \alpha_{12} = (-1)^{1+2}[M_{12}] = -[M_{12}]$$

$$\alpha_{13} = (-1)^{1+3}[M_{13}] = [M_{13}]$$

The remaining minors $[M_{21}], [M_{22}], [M_{23}], [M_{31}], [M_{32}], [M_{33}]$, and the remaining cofactors $\alpha_{21}, \alpha_{22}, \alpha_{23}, \alpha_{31}, \alpha_{32}, \alpha_{33}$ are defined similarly.

Example 2.9

Compute the cofactors of

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

Solution:

$$\alpha_{11} = (-1)^{1+1} \begin{bmatrix} -4 & 2 \\ 2 & -6 \end{bmatrix} = 20 \quad \alpha_{12} = (-1)^{1+2} \begin{bmatrix} 2 & 2 \\ -1 & -6 \end{bmatrix} = 10$$

$$\alpha_{13} = (-1)^{1+3} \begin{bmatrix} 2 & -4 \\ -1 & 2 \end{bmatrix} = 0 \quad \alpha_{21} = (-1)^{2+1} \begin{bmatrix} 2 & -3 \\ 2 & -6 \end{bmatrix} = 6$$

$$\alpha_{22} = (-1)^{2+2} \begin{bmatrix} 1 & -3 \\ -1 & -6 \end{bmatrix} = -9 \quad \alpha_{23} = (-1)^{2+3} \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix} = -4$$

$$\alpha_{31} = (-1)^{3+1} \begin{bmatrix} 2 & -3 \\ -4 & 2 \end{bmatrix} = -8 \quad \alpha_{32} = (-1)^{3+2} \begin{bmatrix} 1 & -3 \\ 2 & 2 \end{bmatrix} = -8$$

and

$$\alpha_{33} = (-1)^{3+3} \begin{bmatrix} 1 & 2 \\ 2 & -4 \end{bmatrix} = -8$$

Chapter 2 Intermediate Algebra

It is useful to remember that the signs of the cofactors follow the pattern

$$\begin{array}{cccccc} + & - & + & - & + & \\ & - & + & - & + & - \\ + & - & + & - & + & \\ & - & + & - & + & - \\ + & - & + & - & + & \end{array}$$

that is, the cofactors on the diagonals, have the same sign as their minors.

It was shown earlier that the determinant of a 2×2 or 3×3 matrix can be found by the method of Example 2.7. This method cannot be used for matrices of higher order. In general, *the determinant of a matrix A of any order, is the sum of the products obtained by multiplying each element of any row or any column by its cofactor.*

Example 2.10

Compute the determinant of A from the elements of the first row and their cofactors given that

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

Solution:

$$\det A = 1 \begin{vmatrix} -4 & 2 \\ 2 & -6 \end{vmatrix} - 2 \begin{vmatrix} 2 & 2 \\ -1 & -6 \end{vmatrix} - 3 \begin{vmatrix} 2 & -4 \\ -1 & 2 \end{vmatrix} = 1 \times 20 - 2 \times (-10) - 3 \times 0 = 40$$

The determinant of a matrix of order 4 or higher, must be evaluated using the above procedure. Thus, the determinant of a fourth-order matrix can first be expressed as the sum of the products of the elements of its first row, by its cofactor as shown in (2.21) below.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = a_{11} \begin{bmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{bmatrix} - a_{21} \begin{bmatrix} a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{bmatrix} + a_{31} \begin{bmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{42} & a_{43} & a_{44} \end{bmatrix} - a_{41} \begin{bmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{bmatrix} \quad (2.21)$$

Example 2.11

Compute the determinant of

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

Solution:

Using the procedure of (2.21), we multiply each element of the first column by its cofactor. Then,

$$A=2 \underbrace{\begin{bmatrix} 1 & 0 & -1 \\ 0 & 3 & -2 \\ 0 & 0 & 1 \end{bmatrix}}_{[a]} - (-1) \underbrace{\begin{bmatrix} -1 & 0 & -3 \\ 0 & 3 & -2 \\ 0 & 0 & 1 \end{bmatrix}}_{[b]} + 4 \underbrace{\begin{bmatrix} -1 & 0 & -3 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}}_{[c]} - (-3) \underbrace{\begin{bmatrix} -1 & 0 & -3 \\ 1 & 0 & -1 \\ 0 & 3 & -2 \end{bmatrix}}_{[d]}$$

Next, using the procedure of Example 2.7 or Example 2.10, we find that

$$[a] = 6, \quad [b] = -3, \quad [c] = 0, \quad [d] = -36$$

and thus

$$\det A = [a] + [b] + [c] + [d] = 6 - 3 + 0 - 36 = -33$$

We can also use the MATLAB function **det(A)** to compute the determinant. For this example,

$$A=[2 \ -1 \ 0 \ -3; -1 \ 1 \ 0 \ -1; 4 \ 0 \ 3 \ -2; -3 \ 0 \ 0 \ 1]; \det(A)$$

ans =

$$-33$$

The determinants of matrices of order five or higher can be evaluated similarly.

Some useful properties of determinants are given below.

1. If all elements of one row or one column of a square matrix are zero, the determinant is zero. An example of this is the determinant of the cofactor [c] in Example 2.11.
2. If all the elements of one row (column) of a square matrix A are m times the corresponding elements of another row (column), detA is zero. For instance, if

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

then,

$$\det A = \begin{vmatrix} 2 & 4 & 1 \\ 3 & 6 & 1 \\ 1 & 2 & 1 \end{vmatrix} \begin{vmatrix} 2 & 4 \\ 3 & 6 \\ 1 & 2 \end{vmatrix} = 12 + 4 + 6 - 6 - 4 - 12 = 0$$

Here, $\det A$ is zero because the second column in A is 2 times the first column.

We can use the MATLAB function **det(A)** to compute the determinant.

Consider the systems of the three equations below.

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= A \\ a_{21}x + a_{22}y + a_{23}z &= B \\ a_{31}x + a_{32}y + a_{33}z &= C \end{aligned} \tag{2.22}$$

Cramer's rule states that the unknowns x , y , and z can be found from the relations

$$x = \frac{D_1}{\Delta} \quad y = \frac{D_2}{\Delta} \quad z = \frac{D_3}{\Delta} \tag{2.23}$$

where

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad D_1 = \begin{vmatrix} A & a_{12} & a_{13} \\ B & a_{22} & a_{23} \\ C & a_{32} & a_{33} \end{vmatrix} \quad D_2 = \begin{vmatrix} a_{11} & A & a_{13} \\ a_{21} & B & a_{23} \\ a_{31} & C & a_{33} \end{vmatrix} \quad D_3 = \begin{vmatrix} a_{11} & a_{12} & A \\ a_{21} & a_{22} & B \\ a_{31} & a_{32} & C \end{vmatrix}$$

provided that the determinant Δ (delta) is not zero.

We observe that the numerators of (2.23) are determinants that are formed from Δ by the substitution of the known values A , B , and C for the coefficients of the desired unknown.

Cramer's rule applies to systems of two or more equations.

If (2.22) is a homogeneous set of equations, that is, if $A = B = C = 0$, then D_1 , D_2 , and D_3 are all zero as we found in Property 1 above. In this case, $x = y = z = 0$ also.

Example 2.12

Use Cramer's rule to find the unknowns v_1 , v_2 , and v_3 if

$$\begin{aligned} 2v_1 - 5 - v_2 + 3v_3 &= 0 \\ -2v_3 - 3v_2 - 4v_1 &= 8 \\ v_2 + 3v_1 - 4 - v_3 &= 0 \end{aligned}$$

Solution:

Rearranging the unknowns v and transferring known values to the right side we get

$$\begin{aligned}2v_1 - v_2 + 3v_3 &= 5 \\ -4v_1 - 3v_2 - 2v_3 &= 8 \\ 3v_1 + v_2 - v_3 &= 4\end{aligned}$$

Now, by Cramer's rule,

$$\Delta = \begin{vmatrix} 2 & -1 & 3 \\ -4 & -3 & -2 \\ 3 & 1 & -1 \end{vmatrix} \begin{vmatrix} 2 & -1 \\ -4 & -3 \\ 3 & 1 \end{vmatrix} = 6 + 6 - 12 + 27 + 4 + 4 = 35$$

Also,

$$D_1 = \begin{vmatrix} 5 & -1 & 3 \\ 8 & -3 & -2 \\ 4 & 1 & -1 \end{vmatrix} \begin{vmatrix} 5 & -1 \\ 8 & -3 \\ 4 & 1 \end{vmatrix} = 15 + 8 + 24 + 36 + 10 - 8 = 85$$

$$D_2 = \begin{vmatrix} 2 & 5 & 3 \\ -4 & 8 & -2 \\ 3 & 4 & -1 \end{vmatrix} \begin{vmatrix} 2 & 5 \\ -4 & 8 \\ 3 & 4 \end{vmatrix} = -16 - 30 - 48 - 72 + 16 - 20 = -170$$

and

$$D_3 = \begin{vmatrix} 2 & -1 & 5 \\ -4 & -3 & 8 \\ 3 & 1 & 4 \end{vmatrix} \begin{vmatrix} 2 & -1 \\ -4 & -3 \\ 3 & 1 \end{vmatrix} = -24 - 24 - 20 + 45 - 16 - 16 = -55$$

Therefore,

$$v_1 = \frac{D_1}{\Delta} = \frac{85}{35} = \frac{17}{7}$$

$$v_2 = \frac{D_2}{\Delta} = \frac{-170}{35} = -\frac{34}{7}$$

$$v_3 = \frac{D_3}{\Delta} = \frac{-55}{35} = -\frac{11}{7}$$

We can verify the first equation as follows:

$$2v_1 - v_2 + 3v_3 = 2 \times \frac{17}{7} - \left(-\frac{34}{7}\right) + 3 \times \left(-\frac{11}{7}\right) = \frac{34 + 34 - 33}{7} = \frac{35}{7} = 5$$

We can also find the unknowns in a system of two or more equations by the *Gaussian Elimination Method*. With this method, the objective is to eliminate one unknown at a time. This is done by

Chapter 2 Intermediate Algebra

multiplying the terms of any of the given equations of the system, by a number such that we can add (or subtract) this equation to (from) another equation in the system, so that one of the unknowns is eliminated. Then, by substitution to another equation with two unknowns, we can find the second unknown. Subsequently, substitution of the two values that were found, can be made into an equation with three unknowns from which we can find the value of the third unknown. This procedure is repeated until all unknowns are found. This method is best illustrated with the following example which consists of the same equations as the previous example.

Example 2.13

Use the Gaussian elimination method to find v_1 , v_2 , and v_3 of

$$\begin{aligned}2v_1 - v_2 + 3v_3 &= 5 \\-4v_1 - 3v_2 - 2v_3 &= 8 \\3v_1 + v_2 - v_3 &= 4\end{aligned}\tag{2.24}$$

Solution:

As a first step, we add the first equation of (2.24) with the third, to eliminate the unknown v_2 . When this is done, we obtain the equation

$$5v_1 + 2v_3 = 9\tag{2.25}$$

Next, we multiply the third equation of (2.24) by 3 and we add it with the second to eliminate v_2 . When this step is done, we obtain the following equation.

$$5v_1 - 5v_3 = 20\tag{2.26}$$

Subtraction of (2.26) from (2.25) yields

$$7v_3 = -11 \text{ or } v_3 = -\frac{11}{7}^*\tag{2.27}$$

Now, we can find the unknown v_1 from either (2.25) or (2.26). Thus, by substitution of (2.27) into (2.25), we get

$$5v_1 + 2 \cdot \left(-\frac{11}{7}\right) = 9 \text{ or } v_1 = \frac{17}{7}\tag{2.28}$$

Finally, we can find the last unknown v_2 from any of the three equations of (2.24). Then, by substitution into the first equation, we obtain

* As stated earlier, it is advisable to leave the values in rational form until the last step.

$$v_2 = 2v_1 + 3v_3 - 5 = \frac{34}{7} - \frac{33}{7} - \frac{35}{7} = -\frac{34}{7}$$

These are the same values as those we found in Example 2.12.

Let A be an n square matrix, and α_{ij} be the cofactor of a_{ij} . Then, *the adjoint of A* , denoted as $\text{adj}A$, is defined as the n square matrix of (2.29) below.

$$\text{adj}A = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \alpha_{31} & \cdots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} & \cdots & \alpha_{n2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{1n} & \alpha_{2n} & \alpha_{3n} & \cdots & \alpha_{nn} \end{bmatrix} \quad (2.29)$$

We observe that the cofactors of the elements of the i th row (column) of A , are the elements of the i th column (row) of $\text{adj}A$.

Example 2.14

Given that

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

compute $\text{adj}A$.

Solution:

$$\text{adj}A = \begin{bmatrix} \begin{bmatrix} 3 & 4 \\ 4 & 3 \end{bmatrix} - \begin{bmatrix} 2 & 3 \\ 4 & 3 \end{bmatrix} & \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix} \\ - \begin{bmatrix} 1 & 4 \\ 1 & 3 \end{bmatrix} & \begin{bmatrix} 1 & 3 \\ 1 & 3 \end{bmatrix} - \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix} \\ \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} -7 & 6 & -1 \\ 1 & 0 & -1 \\ 1 & -2 & 1 \end{bmatrix}$$

An n square matrix A is called *singular* if $\det A = 0$; if $\det A \neq 0$, it is called *non-singular*.

Chapter 2 Intermediate Algebra

Example 2.15

Given that

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

determine whether this matrix is singular or non-singular.

Solution:

$$\det A = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 5 & 7 \end{vmatrix} \begin{vmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 5 \end{vmatrix} = 21 + 24 + 30 - 27 - 20 - 28 = 0$$

Therefore, matrix A is singular.

If A and B are n square matrices such that $AB = BA = I$ where I is the identity matrix, B is called the *inverse* of A, denoted as $B = A^{-1}$, and likewise, A is called the *inverse* of B, that is, $A = B^{-1}$.

If a matrix A is non-singular, we can compute its inverse from the relation

$$\boxed{A^{-1} = \frac{1}{\det A} \text{adj} A} \quad (2.30)$$

Example 2.16

Given that

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

compute its inverse, that is, find A^{-1} .

Solution:

For this example,

$$\det A = 9 + 8 + 12 - 9 - 16 - 6 = -2$$

and since $\det A \neq 0$, it is possible to compute the inverse of A using relation (2.30).

From Example 2.14,

$$\text{adj}A = \begin{bmatrix} -7 & 6 & -1 \\ 1 & 0 & -1 \\ 1 & -2 & 1 \end{bmatrix}$$

Then,

$$A^{-1} = \frac{1}{\det A} \text{adj}A = \frac{1}{-2} \begin{bmatrix} -7 & 6 & -1 \\ 1 & 0 & -1 \\ 1 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 3.5 & -3 & 0.5 \\ -0.5 & 0 & 0.5 \\ -0.5 & 1 & -0.5 \end{bmatrix}$$

Multiplication of a matrix A by its inverse A^{-1} produces the identity matrix I , that is,

$$AA^{-1} = I \quad \text{or} \quad A^{-1}A = I \tag{2.31}$$

Example 2.17

Prove the validity of (2.31) for

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

Proof:

$$\det A = 8 - 6 = 2 \quad \text{and} \quad \text{adj}A = \begin{bmatrix} 2 & -3 \\ -2 & 4 \end{bmatrix}$$

then,

$$A^{-1} = \frac{1}{\det A} \text{adj}A = \frac{1}{2} \begin{bmatrix} 2 & -3 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & -3/2 \\ -1 & 2 \end{bmatrix}$$

Therefore,

$$AA^{-1} = \begin{bmatrix} 4 & 3 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & -3/2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 4-3 & -6+6 \\ 2-2 & -3+4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Consider the relation

$$AX = B \tag{2.32}$$

where A and B are matrices whose elements are known, and X is a matrix whose elements are the unknowns. Here, we assume that A and X are conformable for multiplication. Multiplication of both sides of (2.32) by A^{-1} yields:

$$A^{-1}AX = A^{-1}B = IX = A^{-1}B$$

and since $IX = X$, we obtain the important matrix relation

$$\boxed{X=A^{-1}B} \quad (2.33)$$

We can use (2.33) to solve any set of simultaneous equations that have solutions. We call this method, the *inverse matrix method of solution*.

Example 2.18

For the system of equations

$$2x_1 + 3x_2 + x_3 = 9$$

$$x_1 + 2x_2 + 3x_3 = 6$$

$$3x_1 + x_2 + 2x_3 = 8$$

compute the unknowns x_1 , x_2 , and x_3 using the inverse matrix method.

Solution:

In matrix form, the given set of equations is $AX = B$ where

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad B = \begin{bmatrix} 9 \\ 6 \\ 8 \end{bmatrix}$$

and from (2.33),

$$X = A^{-1}B$$

or

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 6 \\ 8 \end{bmatrix}$$

Next, we compute the determinant and the adjoint of A . Following the procedures of Examples 2.7 or 2.10, and 2.14 we find that

$$\det A = 18 \quad \text{and} \quad \text{adj}A = \begin{bmatrix} 1 & -5 & 7 \\ 7 & 1 & -5 \\ -5 & 7 & 1 \end{bmatrix}$$

Then,

$$A^{-1} = \frac{1}{\det A} \text{adj}A = \frac{1}{18} \begin{bmatrix} 1 & -5 & 7 \\ 7 & 1 & -5 \\ -5 & 7 & 1 \end{bmatrix}$$

Now, using (2.33) we obtain the solution shown below.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{18} \begin{bmatrix} 1 & -5 & 7 \\ 7 & 1 & -5 \\ -5 & 7 & 1 \end{bmatrix} \begin{bmatrix} 9 \\ 6 \\ 8 \end{bmatrix} = \frac{1}{18} \begin{bmatrix} 35 \\ 29 \\ 5 \end{bmatrix} = \begin{bmatrix} 35/18 \\ 29/18 \\ 5/18 \end{bmatrix} = \begin{bmatrix} 1.94 \\ 1.61 \\ 0.28 \end{bmatrix}$$

We can also use Microsoft Excel's **MINVERSE** (Matrix Inversion) and **MMULT** (Matrix Multiplication) functions to obtain the values of the three unknowns of Example 2.18.

The procedure is as follows:

1. We start with a blank spreadsheet, and in a block of cells, say **B4:D6**, we enter the elements of matrix A as shown on the spreadsheet of Figure 2.2. Then, we enter the elements of matrix B in **G4:G6**.
2. Next, we compute and display the inverse of A, that is, A^{-1} . We choose **B8:D10** for the elements of this inverted matrix. We format this block for number display with three decimal places. With this range highlighted, and making sure that the selected cell is **B8**, we type the formula

=MINVERSE(B4:D6)

and we press the *Crtl-Shift-Enter* keys simultaneously. We observe that A^{-1} appears in these cells.

3. Now, we choose the block of cells **G8:G10** for the values of the unknowns. As before, we highlight them, and with the cell marker positioned in **G8**, we type the formula

=MMULT(B8:D10,G4:G6)

and we press the *Crtl-Shift-Enter* keys simultaneously. The values of X then appear in **G8:G10**.

	A	B	C	D	E	F	G	H
1	Spreadsheet for Matrix Inversion and Matrix Multiplication							
2	Example 3.17							
3								
4		2	3	1			9	
5	A=	1	2	3		B=	6	
6		3	1	2			8	
7								
8		0.056	-0.278	0.389			1.9444	
9	A^{-1} =	0.389	0.056	-0.278		X=	1.6111	
10		-0.278	0.389	0.056			0.2778	
11								

Figure 2.3. Spreadsheet for Example 2.18

We should not erase this spreadsheet; we will use it to finish Example 2.2, Page 2–6.

Chapter 2 Intermediate Algebra

To continue with Example 2.2, Page 2–6, we type-in (overwrite) the existing entries as follows:

1. Replace the contents of matrix A, in A4:D6 with the coefficients of the unknowns in (2.14), which is repeated here for convenience. These are shown in the spreadsheet of Figure 2.4.

$$\begin{aligned} 25x + 62y + 54z &= 2756000 \\ 28x + 42y + 58z &= 2695000 \\ 45x + 53y + 56z &= 3124000 \end{aligned} \quad (2.34)$$

2. In G4:G6, we enter the values that appear on the right side of (2.34). The values of the unknowns x, y, and z appear in G8:G10*.

The updated spreadsheet is shown in Figure 2.4.

The range G8:G10 shows that *Brand A* car was sold at an average price of \$20,904.99, *Brand B* at \$11,757.61, and *Brand C* at \$27,589.32.

The last step is to verify that these values satisfy all three equations in (2.34). This can be done easily with Excel as follows:

In any cell, say A11, we type the formula

$$=B4*G8+C4*G9+D4*G10$$

Spreadsheet for Matrix Inversion and Matrix Multiplication						
Example 2.2						
	25	62	54		2756000	
A=	28	42	58	B=	2695000	
	45	53	56		3124000	
	-0.029	-0.025	0.054		20904.99	
A ⁻¹ =	0.042	-0.042	0.003	X=	11757.61	
	-0.016	0.059	-0.028		27859.32	

Figure 2.4. Spreadsheet for Example 2.2

This formula instructs Excel to multiply the contents of B4 by the contents of G8, C4 by G9, D4 by G10, and add these. We observe that A11 displays 2756000; this verifies the first equation. The second and third equations can be verified similarly.

MATLAB[†], which is discussed in Appendix B, offers a more convenient method for matrix oper-

* Make sure that Excel is configured for automatic recalculation. This can be done with the sequential commands **Tool > Options > Calculation > Automatic**.

† For readers not familiar with it, it is highly recommended that Appendix B is studied next. The MATLAB Student Version is an inexpensive software package. For details, please visit www.mathworks.com

ations. To verify our results, we could use the MATLAB **inv(A)** function, and then multiply A^{-1} by B . However, it is easier to use the *matrix left division* operation $X = A \setminus B$; this is MATLAB's solution of $A^{-1}B$ for the matrix equation $A \cdot X = B$, where matrix X is the same size as matrix B . For this example, we enter

```
A=[25 62 54; 28 42 58; 45 53 56]; B=[2756000 2695000 3124000];
```

```
X=A\B
```

and MATLAB displays $x_1 = 2.0905 \times 10^4$, $x_2 = 1.1758 \times 10^4$, and $x_3 = 2.7859 \times 10^4$

2.4 Summary

- In general, a straight line is represented by the equation

$$y = mx + b$$

where m is the *slope*, x is the abscissa, y is the ordinate, and b is the y -*intercept* of the straight line.

- The *slope* m is the *rise* in the vertical (y -*axis*) direction over the *run* in the abscissa (x -*axis*) direction. In other words,

$$m = \frac{\text{rise}}{\text{run}} = \frac{y_2 - y_1}{x_2 - x_1}$$

- A *matrix* is a rectangular array of numbers. The general form of a matrix A is

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

- The numbers a_{ij} are the *elements* of the matrix where the index i indicates the row and j indicates the column in which each element is positioned. Thus, a_{43} represents the element positioned in the fourth row and third column.
- A matrix of m rows and n columns is said to be of $m \times n$ *order matrix*. If $m = n$, the matrix is said to be a *square matrix of order* m (or n).
- In a square matrix, the elements $a_{11}, a_{22}, a_{33}, \dots, a_{nn}$ are called the *diagonal elements*.
- A matrix in which every element is zero, is called a *zero matrix*.
- Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are said to be equal, that is, $A = B$, if, and only if $a_{ij} = b_{ij}$ for $i = 1, 2, 3, \dots, m$ and $j = 1, 2, 3, \dots, n$.
- Two matrices are said to be *conformable for addition (subtraction)* if they are of the same order, that is, both matrices must have the same number of rows and columns.
- If two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are conformable for addition (subtraction), their sum (difference) will be another matrix C with the same order as A and B , where each element of C represents the sum (difference) of the corresponding elements of A and B , that is,

$$C = A \pm B = [a_{ij} \pm b_{ij}]$$

- If k is any scalar (a positive or negative number) and not $[k]$ which is a 1×1 matrix, then multiplication of a matrix A by the scalar k is the multiplication of every element of A by k .
- Two matrices A and B are said to be *conformable for multiplication* $A \cdot B$ in that order, only when the number of columns of matrix A is equal to the number of rows of matrix B . That is, matrix product $A \cdot B$ is conformable for multiplication only if A is an $m \times p$ and matrix B is an $p \times n$ matrix. The matrix product $A \cdot B$ will then be an $m \times n$ matrix.
- The matrix product $A \cdot B$, is not the same as the matrix product $B \cdot A$.
- For matrix multiplication, the operation is row by column. Thus, to obtain the product $A \cdot B$, we multiply each element of a row of A by the corresponding element of a column of B , and then we add these products.
- Division of one matrix by another is not defined. An equivalent operation is performed with the inverse of a matrix.
- An *identity matrix* I is a square matrix where all the elements on the main diagonal are ones, and all other elements are zero.
- If a matrix A is defined as the square matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}$$

the *determinant* of A , denoted as $\det A$, is defined as

$$\det A = a_{11}a_{22}a_{33}\dots a_{nn} + a_{12}a_{23}a_{34}\dots a_{n1} + a_{13}a_{24}a_{35}\dots a_{n2} + \dots - a_{n1}\dots a_{22}a_{13}\dots - a_{n2}\dots a_{23}a_{14} - a_{n3}\dots a_{24}a_{15} - \dots$$

- The determinant of a square matrix of order n is referred to as *determinant of order n* .
- If a matrix A be defined as the square matrix of order n as shown below

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}$$

and we remove the elements of its i th row and j th column, the determinant of the remaining $(n - 1)$ square matrix is called a *minor of determinant A*, and it is denoted as $[M_{ij}]$.

- The signed minor $(-1)^{i+j}[M_{ij}]$ is called the *cofactor* of a_{ij} , and it is denoted as α_{ij} .
- In general, the determinant of a matrix A of any order, is the sum of the products obtained by multiplying each element of any row or any column by its cofactor.
- If all elements of one row or one column of a square matrix are zero, the determinant is zero.
- If all the elements of one row (column) of a square matrix A are m times the corresponding elements of another row (column), $\det A$ is zero.
- Cramer's rule states that the unknowns x , y , and z can be found from the relations

$$x = \frac{D_1}{\Delta} \quad y = \frac{D_2}{\Delta} \quad z = \frac{D_3}{\Delta}$$

where

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad D_1 = \begin{vmatrix} A & a_{12} & a_{13} \\ B & a_{22} & a_{23} \\ C & a_{32} & a_{33} \end{vmatrix} \quad D_2 = \begin{vmatrix} a_{11} & A & a_{13} \\ a_{21} & B & a_{23} \\ a_{31} & C & a_{33} \end{vmatrix} \quad D_3 = \begin{vmatrix} a_{11} & a_{12} & A \\ a_{21} & a_{22} & B \\ a_{31} & a_{32} & C \end{vmatrix}$$

provided that the determinant Δ (delta) is not zero.

- We can also find the unknowns in a system of two or more equations by the *Gaussian Elimination Method*. With this method we eliminate one unknown at a time. This is done by multiplying the terms of any of the given equations of the system, by a number such that we can add (or subtract) this equation to (from) another equation in the system, so that one of the unknowns is eliminated. Then, by substitution to another equation with two unknowns, we can find the second unknown. Subsequently, substitution of the two values that were found, can be made into an equation with three unknowns from which we can find the value of the third unknown. This procedure is repeated until all unknowns are found.
- If a matrix A is an n square matrix, and α_{ij} is the cofactor of a_{ij} , the *adjoint of A*, denoted as $\text{adj}A$, is defined as the n square matrix below.

$$\text{adj}A = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \alpha_{31} & \dots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} & \dots & \alpha_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{1n} & \alpha_{2n} & \alpha_{3n} & \dots & \alpha_{nn} \end{bmatrix}$$

- The cofactors of the elements of the i th row (column) of A , are the elements of the i th column (row) of $\text{adj}A$.
- An n square matrix A is called *singular* if $\det A = 0$; if $\det A \neq 0$, it is called *non-singular*.
- If A and B are n square matrices such that $AB = BA = I$ where I is the identity matrix, B is called the *inverse* of A , denoted as $B = A^{-1}$, and likewise, A is called the *inverse* of B , that is, $A = B^{-1}$.
- If a matrix A is non-singular, we can compute its inverse from the relation

$$A^{-1} = \frac{1}{\det A} \text{adj}A$$

- The relation

$$X = A^{-1}B$$

where A and B are matrices whose elements are known, X is a matrix whose elements are the unknowns, and A and X are conformable for multiplication, can be used to solve any set of simultaneous equations that have solutions. We call this method, the *inverse matrix method of solution*.

- We can also use Microsoft Excel's **MINVERSE** (Matrix Inversion) and **MMULT** (Matrix Multiplication) functions to obtain the values of the unknowns in a system of equations.
- The *matrix left division* operation $X = A \setminus B$ is MATLAB's solution of $A^{-1}B$ for the matrix equation $A \cdot X = B$, where matrix X is the same size as matrix B .

Chapter 2 Intermediate Algebra

2.5 Exercises

For Exercises 1 through 3 below, the matrices A, B, C and D are defined as:

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

- Perform the following computations, if possible. Verify your answers with Excel or MATLAB.
 - $A+B$
 - $A+C$
 - $B+D$
 - $C+D$
 - $A-B$
 - $A-C$
 - $B-D$
 - $C-D$
- Perform the following computations, if possible. Verify your answers with Excel or MATLAB.
 - $A \cdot B$
 - $A \cdot C$
 - $B \cdot D$
 - $C \cdot D$
 - $B \cdot A$
 - $C \cdot A$
 - $D \cdot A$
 - $D \cdot C$
- Perform the following computations, if possible. Verify your answers with Excel or MATLAB.
 - $\det A$
 - $\det B$
 - $\det C$
 - $\det D$
 - $\det(A \cdot B)$
 - $\det(A \cdot C)$
- Solve the following system of equations using Cramer's rule. Verify your answers with Excel or MATLAB.

$$\begin{aligned}x_1 - 2x_2 + x_3 &= -4 \\-2x_1 + 3x_2 + x_3 &= 9 \\3x_1 + 4x_2 - 5x_3 &= 0\end{aligned}$$

- Repeat Exercise 4 using the Gaussian elimination method.
- Use the MATLAB **det(A)** function to find the unknowns of the system of equations below.

$$\begin{aligned}-x_1 + 2x_2 - 3x_3 + 5x_4 &= 14 \\x_1 + 3x_2 + 2x_3 - x_4 &= 9 \\3x_1 - 3x_2 + 2x_3 + 4x_4 &= 19 \\4x_1 + 2x_2 + 5x_3 + x_4 &= 27\end{aligned}$$

- Solve the following system of equations using the inverse matrix method. Verify your answers with Excel or MATLAB.

$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & -2 \\ 2 & 3 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3 \\ -2 \\ 0 \end{bmatrix}$$

8. Use Excel to find the unknowns for the system

$$\begin{bmatrix} 2 & 4 & 3 & -2 \\ 2 & -4 & 1 & 3 \\ -1 & 3 & -4 & 2 \\ 2 & -2 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \\ -14 \\ 7 \end{bmatrix}$$

Verify your answers with the MATLAB left division operation.

2.6 Solutions to End-of-Chapter Exercises

1.

$$\text{a. } A + B = \begin{bmatrix} 1+5 & -1+9 & -4-3 \\ 5-2 & 7+8 & -2+2 \\ 3+7 & -5-4 & 6+6 \end{bmatrix} = \begin{bmatrix} 6 & 8 & -7 \\ 3 & 15 & 0 \\ 10 & -9 & 12 \end{bmatrix} \quad \text{b. } A + C \text{ not conformable for addition}$$

c. $B + D$ not conformable for addition d. $C + D$ not conformable for addition

$$\text{e. } A - B = \begin{bmatrix} 1-5 & -1-9 & -4+3 \\ 5+2 & 7-8 & -2-2 \\ 3-7 & -5+4 & 6-6 \end{bmatrix} = \begin{bmatrix} -4 & -10 & -1 \\ 7 & -1 & -4 \\ -4 & -1 & 0 \end{bmatrix} \quad \text{f. } A - C \text{ not conformable for subtraction}$$

g. $B - D$ not conformable for subtraction h. $C - D$ not conformable for subtraction

2.

$$\text{a. } A \cdot B = \begin{bmatrix} 1 \times 5 + (-1) \times (-2) + (-4) \times 7 & 1 \times 9 + (-1) \times 8 + (-4) \times (-4) & 1 \times (-3) + (-1) \times 2 + (-4) \times 6 \\ 5 \times 5 + 7 \times (-2) + (-2) \times 7 & 5 \times 9 + 7 \times 8 + (-2) \times (-4) & 5 \times (-3) + 7 \times 2 + (-2) \times 6 \\ 3 \times 5 + (-5) \times (-2) + 6 \times 7 & 3 \times 9 + (-5) \times 8 + 6 \times (-4) & 3 \times (-3) + (-5) \times 2 + 6 \times 6 \end{bmatrix}$$
$$= \begin{bmatrix} -21 & 17 & -29 \\ -3 & 109 & -13 \\ 67 & -37 & 17 \end{bmatrix}$$

Check with MATLAB:

 $A=[1 \ -1 \ -4; 5 \ 7 \ -2; 3 \ -5 \ 6]; B=[5 \ 9 \ -3; -2 \ 8 \ 2; 7 \ -4 \ 6]; A*B$

ans =

$$\begin{array}{ccc} -21 & 17 & -29 \\ -3 & 109 & -13 \\ 67 & -37 & 17 \end{array}$$

$$\text{b. } A \cdot C = \begin{bmatrix} 1 \times 4 + (-1) \times (-3) + (-4) \times 5 & 1 \times 6 + (-1) \times 8 + (-4) \times (-2) \\ 5 \times 4 + 7 \times (-3) + (-2) \times 5 & 5 \times 6 + 7 \times 8 + (-2) \times (-2) \\ 3 \times 4 + (-5) \times (-3) + 6 \times 5 & 3 \times 6 + (-5) \times 8 + 6 \times (-2) \end{bmatrix} = \begin{bmatrix} -13 & 6 \\ -11 & 90 \\ 57 & -34 \end{bmatrix}$$

c. $B \cdot D$ not conformable for multiplication

$$\text{d. } C \cdot D = \begin{bmatrix} 4 \times 1 + 6 \times (-3) & 4 \times (-2) + 6 \times 6 & 4 \times 3 + 6 \times (-4) \\ (-3) \times 1 + 8 \times (-3) & (-3) \times (-2) + 8 \times 6 & (-3) \times 3 + 8 \times (-4) \\ 5 \times 1 + (-2) \times (-3) & 5 \times (-2) + (-2) \times 6 & 5 \times 3 + (-2) \times (-4) \end{bmatrix} = \begin{bmatrix} -14 & 28 & -12 \\ -27 & 54 & -41 \\ 11 & -22 & 23 \end{bmatrix}$$

$$\begin{aligned}
 \text{e. } B \cdot A &= \begin{bmatrix} 5 \times 1 + 9 \times 5 + (-3) \times 3 & (-2) \times 1 + 8 \times 5 + 2 \times 3 & 7 \times 1 + (-4) \times 5 + 6 \times 3 \\ 5 \times (-1) + 9 \times 7 + (-3) \times (-5) & (-2) \times (-1) + 8 \times 7 + 2 \times (-5) & 7 \times (-1) + (-4) \times 7 + 6 \times (-5) \\ 5 \times (-4) + 9 \times (-2) + (-3) \times 6 & (-2) \times (-4) + 8 \times (-2) + 2 \times 6 & 7 \times (-4) + (-4) \times (-2) + 6 \times 6 \end{bmatrix} \\
 &= \begin{bmatrix} 41 & 73 & -56 \\ 44 & 48 & 4 \\ 5 & -65 & 16 \end{bmatrix}
 \end{aligned}$$

f. $C \cdot A$ not conformable for multiplication

$$\begin{aligned}
 \text{g. } D \cdot A &= \begin{bmatrix} 1 \times 1 + (-2) \times 5 + 3 \times 3 & 1 \times (-1) + (-2) \times 7 + 3 \times (-5) & 1 \times (-4) + (-2) \times (-2) + 3 \times 6 \\ (-3) \times 1 + 6 \times 5 + (-4) \times 3 & (-3) \times (-1) + 6 \times 7 + (-4) \times (-5) & (-3) \times (-4) + 6 \times (-2) + (-4) \times 6 \end{bmatrix} \\
 &= \begin{bmatrix} 0 & -30 & 18 \\ 15 & 65 & -24 \end{bmatrix}
 \end{aligned}$$

$$\text{h. } D \cdot C = \begin{bmatrix} 1 \times 4 + (-2) \times (-3) + 3 \times 5 & 1 \times 6 + (-2) \times 8 + 3 \times (-2) \\ (-3) \times 4 + 6 \times (-3) + (-4) \times 5 & (-3) \times 6 + 6 \times 8 + (-4) \times (-2) \end{bmatrix} = \begin{bmatrix} 25 & -16 \\ -50 & 38 \end{bmatrix}$$

3.

$$\begin{aligned}
 \text{a. } \det A &= \begin{vmatrix} 1 & -1 & -4 & 1 & -1 \\ 5 & 7 & -2 & 5 & 7 \\ 3 & -5 & 6 & 3 & -5 \end{vmatrix} \\
 &= 1 \times 7 \times 6 + (-1) \times (-2) \times 3 + (-4) \times 5 \times (-5) - [3 \times 7 \times (-4) + (-5) \times (-2) \times 1 + 6 \times 5 \times (-1)] \\
 &= 42 + 6 + 100 - (-84) - 10 - (-30) = 252
 \end{aligned}$$

$$\begin{aligned}
 \text{b. } \det B &= \begin{vmatrix} 5 & 9 & -3 & 5 & 9 \\ -2 & 8 & 2 & -2 & 8 \\ 7 & -4 & 6 & 7 & -4 \end{vmatrix} \\
 &= 5 \times 8 \times 6 + 9 \times 2 \times 7 + (-3) \times (-2) \times (-4) - [7 \times 8 \times (-3) + (-4) \times 2 \times 5 + 6 \times (-2) \times 9] \\
 &= 240 + 126 - 24 - (-168) + 40 - (-108) = 658
 \end{aligned}$$

c. $\det C$ does not exist; matrix must be square

d. $\det D$ does not exist; matrix must be square

e. $\det(A \cdot B) = \det A \cdot \det B$ and from parts (a) and (b), $\det(A \cdot B) = 252 \times 658 = 165816$

f. $\det(A \cdot C)$ does not exist because $\det C$ does not exist

Chapter 2 Intermediate Algebra

4.

$$\begin{aligned}\Delta &= \begin{vmatrix} 1 & -2 & 1 & 1 & -2 \\ -2 & 3 & 1 & -2 & 3 \\ 3 & 4 & -5 & 3 & 4 \end{vmatrix} \\ &= 1 \times 3 \times (-5) + (-2) \times 1 \times 3 + 1 \times (-2) \times 4 - [3 \times 3 \times 1 + 4 \times 1 \times 1 + (-5) \times (-2) \times (-2)] \\ &= -15 - 6 - 8 - 9 - 4 + 20 = -22\end{aligned}$$

$$\begin{aligned}D_1 &= \begin{vmatrix} -4 & -2 & 1 & 4 & -2 \\ 9 & 3 & 1 & 9 & 3 \\ 0 & 4 & -5 & 0 & 4 \end{vmatrix} \\ &= -4 \times 3 \times (-5) + (-2) \times 1 \times 0 + 1 \times 9 \times 4 - [0 \times 3 \times 1 + 4 \times 1 \times 4 + (-5) \times 9 \times (-2)] \\ &= 60 + 0 + 36 - 0 + 16 - 90 = 22\end{aligned}$$

$$\begin{aligned}D_2 &= \begin{vmatrix} 1 & -4 & 1 & 1 & -4 \\ -2 & 9 & 1 & -2 & 9 \\ 3 & 0 & -5 & 3 & 0 \end{vmatrix} \\ &= 1 \times 9 \times (-5) + (-4) \times 1 \times 3 + 1 \times (-2) \times 0 - [3 \times 9 \times 1 + 0 \times 1 \times 1 + (-5) \times (-2) \times (-4)] \\ &= -45 - 12 - 0 - 27 - 0 + 40 = -44\end{aligned}$$

$$\begin{aligned}D_3 &= \begin{vmatrix} 1 & -2 & -4 & 1 & -2 \\ -2 & 3 & 9 & -2 & 3 \\ 3 & 4 & 0 & 3 & 4 \end{vmatrix} \\ &= 1 \times 3 \times 0 + (-2) \times 9 \times 3 + (-4) \times (-2) \times 4 - [3 \times 3 \times (-4) + 4 \times 9 \times 1 + 0 \times (-2) \times (-2)] \\ &= 0 - 54 + 32 + 36 - 36 - 0 = -22\end{aligned}$$

$$x_1 = \frac{D_1}{\Delta} = \frac{22}{-22} = -1 \quad x_2 = \frac{D_2}{\Delta} = \frac{-44}{-22} = 2 \quad x_3 = \frac{D_3}{\Delta} = \frac{-22}{-22} = 1$$

5.

$$x_1 - 2x_2 + x_3 = -4 \quad (1)$$

$$-2x_1 + 3x_2 + x_3 = 9 \quad (2)$$

$$3x_1 + 4x_2 - 5x_3 = 0 \quad (3)$$

Multiplication of (1) by 2 yields

$$2x_1 - 4x_2 + 2x_3 = -8 \quad (4)$$

Addition of (2) and (4) yields

$$-x_2 + 3x_3 = 1 \quad (5)$$

Multiplication of (1) by -3 yields

$$-3x_1 + 6x_2 - 3x_3 = 12 \quad (6)$$

Addition of (3) and (6) yields

$$10x_2 - 8x_3 = 12 \quad (7)$$

Multiplication of (5) by 10 yields $-10x_2 + 30x_3 = 10$ (8)

Addition of (7) and (8) yields $22x_3 = 22$ (9)

or $x_3 = 1$ (10)

Substitution of (10) into (7) yields $10x_2 - 8 = 12$ (11)

or $x_2 = 2$ (12)

and substitution of (10) and (12) into (1) yields

$$x_1 - 4 + 1 = -4 \quad (13)$$

or $x_1 = -1$ (14)

6.

$$\Delta = [-1 \ 2 \ -3 \ 5; 1 \ 3 \ 2 \ -1; 3 \ -3 \ 2 \ 4; 4 \ 2 \ 5 \ 1];$$

$$D1 = [14 \ 2 \ -3 \ 5; 9 \ 3 \ 2 \ -1; 19 \ -3 \ 2 \ 4; 27 \ 2 \ 5 \ 1];$$

$$D2 = [-1 \ 14 \ -3 \ 5; 1 \ 9 \ 2 \ -1; 3 \ 19 \ 2 \ 4; 4 \ 27 \ 5 \ 1];$$

$$D3 = [-1 \ 2 \ 14 \ 5; 1 \ 3 \ 9 \ -1; 3 \ -3 \ 19 \ 4; 4 \ 2 \ 27 \ 1];$$

$$D4 = [-1 \ 2 \ -3 \ 14; 1 \ 3 \ 2 \ 9; 3 \ -3 \ 2 \ 19; 4 \ 2 \ 5 \ 27];$$

$$x1 = \det(D1)/\det(\Delta), \quad x2 = \det(D2)/\det(\Delta), \dots$$

$$x3 = \det(D3)/\det(\Delta), \quad x4 = \det(D4)/\det(\Delta)$$

$$x1=1 \quad x2=2 \quad x3=3 \quad x4=4$$

7.

$$\begin{aligned} \det A &= \begin{vmatrix} 1 & 3 & 4 & 1 & 3 \\ 3 & 1 & -2 & 3 & 1 \\ 2 & 3 & 5 & 2 & 3 \end{vmatrix} \\ &= 1 \times 1 \times 5 + 3 \times (-2) \times 2 + 4 \times 3 \times 3 - [2 \times 1 \times 4 + 3 \times (-2) \times 1 + 5 \times 3 \times 3] \\ &= 5 - 12 + 36 - 8 + 6 - 45 = -18 \end{aligned}$$

$$\text{adj} A = \begin{bmatrix} 11 & -3 & -10 \\ -19 & -3 & 14 \\ 7 & 3 & -8 \end{bmatrix}$$

$$A^{-1} = \frac{1}{\det A} \cdot \text{adj} A = \frac{1}{-18} \cdot \begin{bmatrix} 11 & -3 & -10 \\ -19 & -3 & 14 \\ 7 & 3 & -8 \end{bmatrix} = \begin{bmatrix} -11/18 & 3/18 & 10/18 \\ 19/18 & 3/18 & -14/18 \\ -7/18 & -3/18 & 8/18 \end{bmatrix}$$

Chapter 2 Intermediate Algebra

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -11/18 & 3/18 & 10/18 \\ 19/18 & 3/18 & -14/18 \\ -7/18 & -3/18 & 8/18 \end{bmatrix} \begin{bmatrix} -3 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 33/18 - 6/18 + 0 \\ -57/18 - 6/18 + 0 \\ 21/18 + 6/18 + 0 \end{bmatrix} = \begin{bmatrix} 27/18 \\ -63/18 \\ 27/18 \end{bmatrix} = \begin{bmatrix} 1.50 \\ -3.50 \\ 1.50 \end{bmatrix}$$

1	Spreadsheet for Matrix Inversion and							
2	Matrix Multiplication - Exercise 7							
3								
4		1.00	3.00	4.00			-3.00	
5	A=	3.00	1.00	-2.00		B=	-2.00	
6		2.00	3.00	5.00			0.00	
7								
8		-0.61	0.17	0.56			1.50	
9	A ⁻¹	1.06	0.17	-0.78		X=	-3.50	
10		-0.39	-0.17	0.44			1.50	

8.

	A	B	C	D	E	F	G	H	
1	Spreadsheet for Matrix Inversion and								
2	Matrix Multiplication - Exercise 8								
3									
4		2.00	4.00	3.00	-2.00			1.00	
5	A=	2.00	-4.00	1.00	3.00		B=	10.00	
6		-1.00	3.00	-4.00	2.00			-14.00	
7		2.00	-2.00	2.00	1.00			7.00	
8									
9		-1.58	-4.08	1.17	6.75			-11.50	
10	A ⁻¹	0.58	1.08	-0.17	-1.75		X=	1.50	
11		1.50	3.50	-1.00	-5.50			12.00	
12		1.33	3.33	-0.67	-5.00			9.00	

A=[2 4 3 -2; 2 -4 1 3; -1 3 -4 2; 2 -2 2 1];

B=[1 10 -14 7]; AB

ans =

-11.5000

1.5000

12.0000

9.0000

Chapter 3

Fundamentals of Geometry

This chapter discusses the basic geometric figures. It is intended for readers who need to learn the basics of geometry. Readers with a strong mathematical background may skip this chapter. Others will find it useful, as well as a convenient source for review.

3.1 Introduction

Geometry is the mathematics of the properties, measurement, and relationships of points, lines, angles, surfaces, and solids. The science of geometry also includes *analytic geometry*, *descriptive geometry*, *fractal geometry*, *non-Euclidean geometry*^{*}, and spaces with four or more dimensions. It is beyond the scope of this text to discuss the different types of geometry in detail; we will only present the most common figures and their properties such as the number of sides, angles, perimeters, and areas. In this text, we will only be concerned with the so-called *Euclidean Geometry*[†].

3.2 Plane Geometry Figures

A *triangle* is a plane figure that is formed by connecting three points, not all in a straight line, by straight line segments. It is a special case of a *polygon* which is defined as a closed plane figure bounded by three or more line segments. The general form of a triangle is shown in Figure 3.1 where the capital letters A, B, and C denote the angles, and the lower case letters denote the three sides of the triangle.

* In **analytic geometry** straight lines, curves, and geometric figures are represented by algebraic expressions. Any point in a plane may be located by specifying the distance of the point from each of a pair of perpendicular axes. Points in three-dimensional space can be similarly located with respect to three axes. A straight line can always be represented by an equation in the form $ax + by + c = 0$.

The science of making accurate, two-dimensional drawings of three-dimensional geometrical forms is called **descriptive geometry**. The usual technique is by means of orthographic projection, in which the object to be drawn is referred to one or more imaginary planes that are at right angles to one another.

A **fractal** is a geometric pattern that is repeated at ever smaller scales to produce irregular shapes and surfaces that cannot be represented by classical geometry. Fractals are used especially in computer modeling of irregular patterns and structures in nature.

† Euclidean Geometry is based on one of Euclid's postulates which states that through a point outside a line it is possible to draw only one line parallel to that line. For many centuries mathematicians believed that this postulate could be proved on the basis of Euclid's other postulates, but all efforts to discover such a proof were unsuccessful. In the first part of the 19th century the German mathematician Carl Friedrich Gauss, the Russian mathematician Nikolay Ivanovich Lobachevsky, and the Hungarian mathematician János Bolyai independently demonstrated a consistent system of geometry in which Euclid's postulate was replaced by one stating that an infinite number of parallels to a particular line could be drawn. About 1860 the German mathematician Georg Friedrich Bernhard Riemann showed that a geometry with no parallel lines was equally possible. For comparatively small distances, Euclidean geometry and the non-Euclidean geometries are about the same. However, in dealing with astronomical distances and relativity, non-Euclidean geometries give a more precise description of the observed facts than Euclidean geometry.

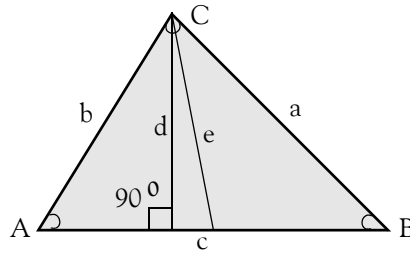


Figure 3.1. General Form of a Triangle

It is customary to show the sides of a triangle opposite to the angles. Thus, side a is opposite to angle A , side b is opposite to angle B and so on.

We will denote the angles with the symbol \angle . Thus, $\angle A$ will mean angle A .

The *perimeter* of a polygon is defined as the closed curve that bounds the plane area of the polygon. In other words, it is the entire length of the polygon's boundary.

The *area* of a polygon is the space inside the boundary of the polygon, and it is expressed in square units. In Figure 3.1, the shaded portion represents the area of this triangle.

Note 3.1

The area of some basic polygons can be computed with simple formulas. However, the formulas for the area of most polygons involve trigonometric functions. We will discuss the most common on the next chapter. It is also possible to compute the areas of irregular polygons with a spreadsheet, as we will see later in this chapter.

A *perpendicular* line is one that intersects another line to form a 90° angle. Thus, a vertical line that intersects a horizontal line, is a perpendicular to the horizontal since it forms a 90° angle with it.

For the triangle of Figure 3.1, line d is a perpendicular line drawn from $\angle C$ down to line c forming 90° angle as shown.

In geometry, the line that joins a vertex of a triangle to the midpoint of the opposite side is called the *median*. Thus in Figure 3.1, line e is a median from $\angle C$ to the midpoint of line c .

Property 1

The sum of the three angles of any triangle is 180° .

Example 3.1

In Figure 3.1 above, $\angle A = 53^\circ$ and $\angle B = 48^\circ$. Find $\angle C$.

Solution:

From Property 1,

$$\angle A + \angle B + \angle C = 180^\circ$$

or

$$\angle C = 180^\circ - (\angle A + \angle B) \quad (3.1)$$

and since

$$\angle A + \angle B = 53^\circ + 48^\circ = 101^\circ$$

by substitution into (3.1),

$$\angle C = 180^\circ - 101^\circ = 79^\circ$$

Example 3.2

In Figure 3.1, $a = 45$ cm, $b = 37$ cm, and $c = 57$ cm. Find the perimeter of the triangle.

Solution:

The perimeter is the sum of the lengths of the three sides of the triangle. Therefore,

$$\text{perimeter} = a + b + c = 45 + 37 + 57 = 139 \text{ cm}$$

The right, the isosceles, and the equilateral triangles are special cases of the general triangle of Figure 3.1. We will discuss each separately. Formulas involving trigonometric functions will be given on the next chapter.

A triangle containing an angle of 90° is a *right triangle*. Thus, Figure 3.2 is a right triangle.

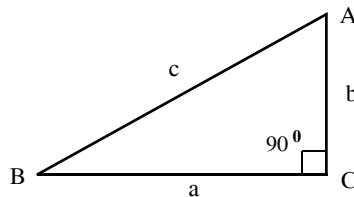


Figure 3.2. Right Triangle

Property 2

If, in a right triangle, $\angle C = 90^\circ$, then $\angle A + \angle B = 90^\circ$. This is a consequence of the fact that in a triangle, the sum of the three angles is 180° .

In a right triangle, the side opposite to the right angle is called the *hypotenuse* of the right triangle. Thus, side c in Figure 3.2, is the hypotenuse of that triangle.

Property 3

If c is the hypotenuse of a right triangle, then

Chapter 3 Fundamentals of Geometry

$$\boxed{c^2 = a^2 + b^2} \quad (3.2)$$

Relation (3.2) is known as the *Pythagorean theorem*.

Example 3.3

In Figure 3.2 above, $c = 5$ cm and $b = 3$ cm. Find the length of the line segment a .

Solution:

We express (3.2) as

$$a^2 = c^2 - b^2$$

or

$$a = \sqrt{c^2 - b^2}$$

Then, with the given values

$$a = \sqrt{5^2 - 3^2} = \sqrt{25 - 9} = \sqrt{16} = 4 \text{ in.}^*$$

Property 4

If c is the hypotenuse of a right triangle, the area of this triangle is found from the relation

$$\boxed{\text{Area} = \frac{1}{2} \cdot a \cdot b} \quad (3.3)$$

Area of Right Triangle

Example 3.4

In Figure 3.2, $a = 4$ cm. and $b = 3$ cm. Find the area of this triangle.

Solution:

From (3.3),

$$\text{Area} = \frac{1}{2}ab = \frac{1}{2} \times 4 \times 3 = 6 \text{ cm}^2.^\dagger$$

* As we know, $\sqrt{16} = \pm 4$ but we reject the negative value since it is unrealistic for this example.

† In area calculations, besides multiplication of the numbers, we must multiply the units also. Thus, for this example, inches \times inches = square inches. Otherwise, the result is meaningless.

A triangle that has two equal sides is called *isosceles* triangle. Figure 3.3 shows an isosceles triangle where sides a and b are equal. Consequently, $\angle A = \angle B$.

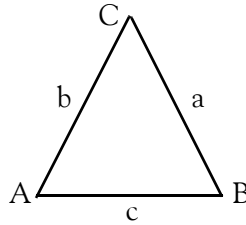


Figure 3.3. Isosceles Triangle

Example 3.5

In Figure 3.3 above, $a = 5$ cm, $c = 4$ cm and $\angle A = 70^\circ$. Compute:

- a. the perimeter
- b. $\angle C$
- c. the area

Solution:

a.
$$\text{Perimeter} = a + b + c$$
 and since this is an isosceles triangle, $a = b$
 Therefore,
$$\text{perimeter} = 2a + c = 2 \times 5 + 4 = 14 \text{ cm}$$

b. By Property 1,
$$\angle A + \angle B + \angle C = 180^\circ$$
 or
$$\angle C = 180^\circ - (\angle A + \angle B)$$
 Since this is an isosceles triangle,
$$\angle B = \angle A = 70^\circ$$

Therefore,
$$\angle C = 180^\circ - (70^\circ + 70^\circ) = 180^\circ - 140^\circ = 40^\circ$$

c. No formula was given to compute the area of an isosceles triangle; however, we can use the formula of the right triangle after we split the isosceles triangle into two equal right triangles. We do this by drawing a perpendicular line from $\angle C$ down to line c forming a 90° angle as shown in Figure 3.4. We denote this line as h (for height).

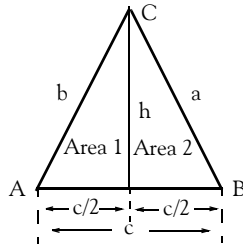


Figure 3.4. Isosceles Triangle for Example 3.5

Since the perpendicular line h divides the isosceles triangle into two equal right triangles, it follows that $\text{Area 1} = \text{Area 2}$. Therefore, we only need to compute Area 1 and multiply it by 2 to obtain the total area. Thus,

$$\text{Area 1} = \frac{1}{2} \times \frac{c}{2} \times h \quad (3.4)$$

where $c/2 = 4/2 = 2$ cm. Next, we need to find the height h . This is found from the Pythagorean theorem stated in Property 3. Then, for the right triangle of the left side of Figure 3.4 above,

$$b^2 = (c/2)^2 + h^2$$

or

$$h^2 = b^2 - (c/2)^2$$

By substitution of the given values,

$$h^2 = 5^2 - (4/2)^2 = 25 - 4 = 21$$

or

$$h = \sqrt{21} = 4.5826$$

Finally, by substitution into (3.4), and multiplying by 2 to obtain the total area, we get

$$\text{Total Area} = 2 \times \text{Area 1} = 2 \times \frac{1}{2} \times \frac{c}{2} \times h = 2 \times 4.5826 = 9.1652 \text{ in}^2$$

A triangle that has all sides equal is called *equilateral triangle*. Figure 3.5 shows an equilateral triangle where sides a , b , and c are equal. Consequently, $\angle A$, $\angle B$, and $\angle C$ are equal.

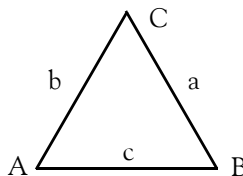


Figure 3.5. Equilateral Triangle

For an equilateral triangle,

$$\angle A = \angle B = \angle C = 60^\circ \quad (3.5)$$

$$\text{Area} = \frac{1}{4}a^2\sqrt{3} \quad (3.6)$$

Congruent triangles are those that are identical in shape and size. If two triangles are congruent, the following statements are true:

- I. When two triangles are congruent, the six corresponding pairs of angles and sides are also congruent.
- II If all three pairs of corresponding sides in two triangles are the same as shown in Figure 3.6, the triangles are congruent.

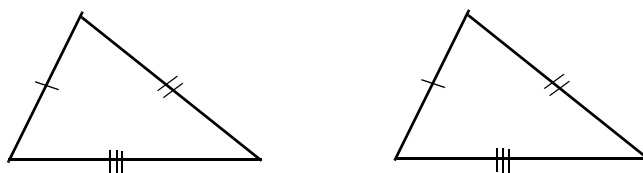


Figure 3.6. The three sides of one triangle are congruent to the three sides of the other triangle

- III If two angles and the included side of one triangle are congruent to two angles and the included side of another triangle as shown in Figure 3.7, the triangles are congruent.

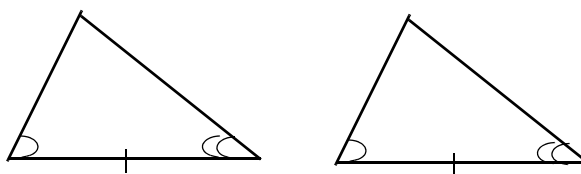


Figure 3.7. Two angles and the included side of one triangle are congruent to two angles and the included side of the other triangle

- IV If two sides and the included angle of one triangle are congruent to two sides and the included angle of another triangle as shown in Figure 3.8, the triangles are congruent.

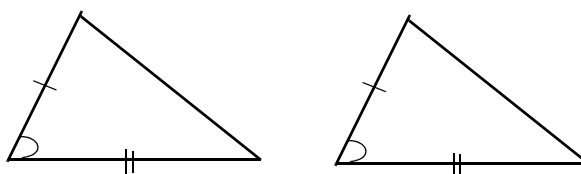


Figure 3.8. Two sides and the included angle of one triangle are congruent to two sides and the included angle of the other triangle

- V If two angles and a non-included side of one triangle are congruent to the two angles and the corresponding non-included side of another triangle as shown in Figure 3.9, the triangles are congruent.

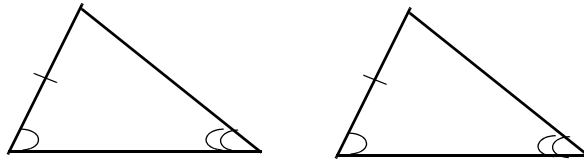


Figure 3.9. Two angles and a non-included side of one triangle are congruent to the two angles and the non-included side of the other triangle

VI If the hypotenuse and the leg of one right triangle are congruent to the corresponding parts of another triangle as shown in Figure 3.10, the two triangles are congruent.

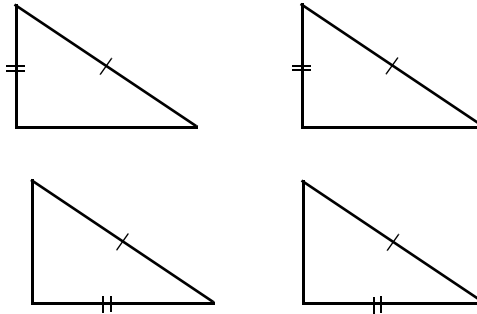


Figure 3.10. The hypotenuse and the leg of one right triangle are congruent to the corresponding parts of another right triangle

VII Two triangles with two sides and a non-included angle may or may not be congruent.

VIII Every triangle can be partitioned into four congruent triangles by connecting the middle points of its three sides as shown in Figure 3.11.

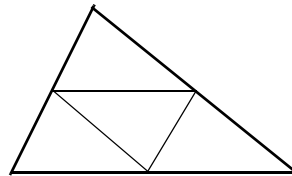


Figure 3.11. Triangle partitioned into four congruent triangles

IX A triangle that can be partitioned into n congruent parts, it can also be partitioned into $4 \times n$, $4^2 \times n$, $4^3 \times n$, ... congruent parts.

X Any isosceles triangle can be partitioned into two congruent triangles by one of its heights.

XI Any equilateral triangle can be partitioned into three congruent triangles by segments connecting its center to its vertices.

XII A right triangle whose sides forming the right angle have a ratio 2 to 1 can be partitioned into five congruent triangles as shown on Figure 3.12.

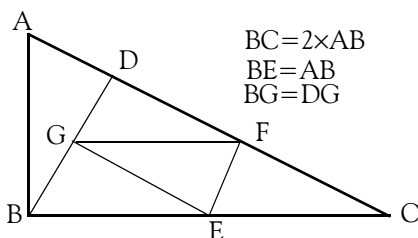


Figure 3.12. Right triangle partitioned into five congruent triangles

In Figure 3.12, line BD is drawn to form right angles with line AC and line EF is drawn to be parallel to line BD. From points E and F lines are drawn to point G which is the middle of line BD.

Figure 3.13 shows two parallel lines x and y intersected by a transversal line z.

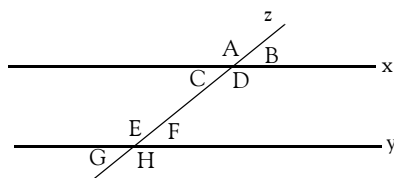


Figure 3.13. Two parallel lines intersected by a transversal

With reference to the lines shown on Figure 3.13:

1. When the transversal line z crosses a line x or line y or both, two sets of opposite angles are formed. Then, using the symbol \angle to denote an angle, and the symbol \cong to indicate congruence, the following relations are true:

$$\begin{aligned} \angle A &\cong \angle D & \angle B &\cong \angle C \\ \angle E &\cong \angle H & \angle F &\cong \angle G \end{aligned} \tag{3.7}$$

2. Corresponding angles on the same side of the transversal are congruent. Thus, on the left side of the transversal we have

$$\angle A \cong \angle E \quad \angle C \cong \angle G \tag{3.8}$$

and on the right side of the transversal we have

$$\angle B \cong \angle F \quad \angle D \cong \angle H \tag{3.9}$$

3. The angles between the two parallel lines x and y are referred to as the *interior* (or inside) angles. Thus, in Figure 3.1 $\angle C$, $\angle D$, $\angle E$, and $\angle F$ are interior angles. Then,

$$\angle C \cong \angle F \quad \angle D \cong \angle E \tag{3.10}$$

4. The angles above and below the two parallel lines x and y are referred to as the *exterior* (or outside) angles. Thus, in Figure 3.1 $\angle A$, $\angle B$, $\angle G$, and $\angle H$ are exterior angles. Then,

Chapter 3 Fundamentals of Geometry

$$\angle A \cong \angle G \quad \angle B \cong \angle H \quad (3.11)$$

Similar triangles have corresponding angles that are congruent (equal) while the lengths of the corresponding sides are in proportion. Thus, the triangles shown on Figure 3.14 are similar if

$$\angle A \cong \angle D \quad \text{and} \quad \angle C \cong \angle F \quad (3.12)$$

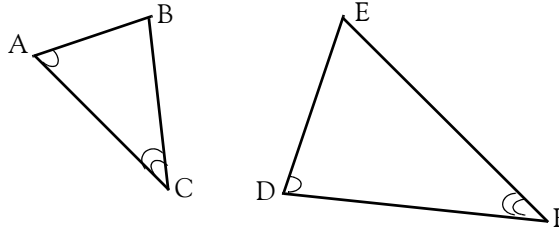


Figure 3.14. Similar triangles

If (3.11) holds, then

$$\frac{AC}{DE} = \frac{BC}{EF} = \frac{AB}{DF} \quad (3.13)$$

We should also remember that:

1. All congruent triangles are similar, but not all similar triangles are congruent.
2. If the angles in two triangles are equal and the corresponding sides are the same size, the triangles are congruent.
3. If the angles in two triangles are equal the triangles are similar.
4. If the angles in two triangles are not equal the triangles are neither similar nor congruent.

A plane figure with four sides and four angles is called *quadrilateral*. Figure 3.15 shows a general quadrilateral.

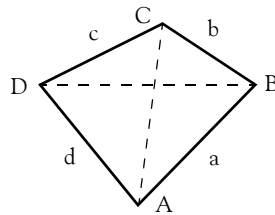


Figure 3.15. General Quadrilateral

The lines joining two non-adjacent vertices of a quadrilateral are called *diagonals*. Thus, in Figure 3.15, the lines AC and BD are the diagonals of that quadrilateral.

Property 5

In any quadrilateral, the sum of the four angles is 360°

Property 6

The diagonals of a quadrilateral with consecutive sides a , b , c , and d are perpendicular if and only if

$$a^2 + c^2 = b^2 + d^2 \quad (3.14)$$

The *square*, *rectangle*, *parallelogram*, *rhombus*, and *trapezoid* are special cases of quadrilaterals. We will discuss each separately.

Property 7

Two lines are said to be in *parallel*, when the distance between them is the same everywhere. In other words, two lines on the same plane that never intersect, are said to be *parallel lines*.*

A four-sided plane figure with opposite sides in parallel, is called *parallelogram*. Figure 3.16 shows a parallelogram.

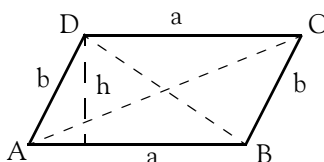


Figure 3.16. Parallelogram

For the parallelogram of Figure 3.16,

$$\begin{aligned} \angle A &= \angle C \\ \angle B &= \angle D \\ \angle A + \angle B &= \angle C + \angle D = 180^\circ \end{aligned} \quad (3.15)$$

$$\boxed{\text{Area} = ah} \quad (3.16)$$

Area of Parallelogram

Note: The perpendicular distance from the center of a regular polygon to any of its sides is called *apothem*.

An equilateral parallelogram is called *rhombus*. It is a special case of a parallelogram where all sides are equal. Figure 3.17 shows a rhombus.

* This statement is consistent with the principles of Euclidean geometry.

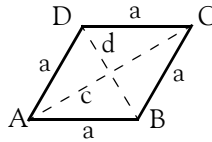


Figure 3.17. Rhombus

For the rhombus of Figure 3.17,

$$c^2 + d^2 = 4a^2 \quad (3.17)$$

$$\boxed{\text{Area} = \frac{1}{2}cd} \quad (3.18)$$

Area of Rhombus

A four-sided plane figure with four right angles, is a *rectangle*. It is a special case of a parallelogram where all angles are equal, that is, 90° each. Figure 3.18 shows a rectangle.

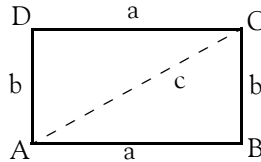


Figure 3.18. Rectangle

For the rectangle of Figure 3.18,

$$\angle A = \angle B = \angle C = \angle D = 90^\circ \quad (3.19)$$

$$c = \sqrt{a^2 + b^2} \quad (3.20)$$

$$\boxed{\text{Area} = ab} \quad (3.21)$$

Area of Rectangle

A plane figure with four equal sides and four equal angles, is a *square*. It is a special case of a rectangle where all sides are equal. Figure 3.19 shows a square.

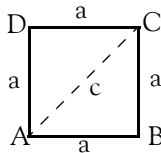


Figure 3.19. Square

For the square of Figure 3.19,

$$\angle A = \angle B = \angle C = \angle D = 90^\circ \quad (3.22)$$

$$c = a\sqrt{2} \tag{3.23}$$

Area = a^2

Area of Square

$$\tag{3.24}$$

A quadrilateral with two unequal parallel sides is a *trapezoid*. Figure 3.20 shows a *trapezoid*.

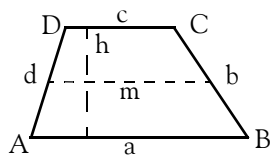


Figure 3.20. Trapezoid

For the trapezoid of Figure 3.20,

$$m = \frac{1}{2}(a + c) \tag{3.25}$$

Area = $\frac{1}{2}(a + c)h = mh$

Area of Trapezoid

$$\tag{3.26}$$

A plane curve that is everywhere equidistant from a given fixed point, referred to as the *center*, is called a *circle*. Figure 3.21 shows a circle where point *o* is the *center* of the circle. The line segment that joins the center of a circle with any point on its circumference is called *radius*, and it is denoted with the letter *r*. A straight line segment passing through the center of a circle, and terminating at the circumference is called *diameter*, and is denoted with the letter *d*. Obviously, $d = 2r$.

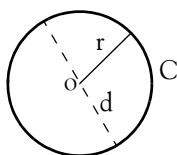


Figure 3.21. Circle

The perimeter of a circle is called *circumference* and it is denoted with the letter *C*.

Note 3.2

The Greek letter π represents the ratio of the circumference of a circle to its diameter. This ratio is an irrational (endless) number, so the decimal places go on infinitely without repeating. It is approximately equal to $22/7$; to five decimal places, is equal to 3.14159. With computers, this value has been figured to more than 100 million decimal places.

For the circle of Figure 3.21,

Chapter 3 Fundamentals of Geometry

$$C = 2\pi r = \pi d \quad (3.27)$$

$$\boxed{\text{Area} = \pi r^2 = \frac{1}{4}\pi d^2} \quad (3.28)$$

Area of Circle

Note 3.3

Locus is defined as the set of all points whose coordinates satisfy a single equation or one or more algebraic conditions.

The locus of points for which the sum of the distances from each point to two fixed points is equal forms an *ellipse**. The two fixed points are called *foci* (plural for focus). Figure 3.22 shows an ellipse.

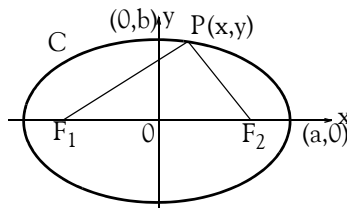


Figure 3.22. Ellipse

For the ellipse of Figure 3.22, F_1 and F_2 are the *foci* and C is the circumference.

In an ellipse,

$$PF_1 + PF_2 = 2a = \text{constant} \quad (3.29)$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (3.30)$$

$$C \approx 2\pi \sqrt{\frac{a^2 + b^2}{2}} \quad (3.31)$$

$$\text{Area} = \pi ab \quad (3.32)$$

The line segments $2a$ and $2b$, are referred to as the *major* and *minor* axes respectively. The foci are always on the major axis; they are located symmetrically from the center O , the intersection of the major and minor axes. The foci locations can be determined with the aid of Figure 3.23 below, where the distance OF_2 , denoted as c , is found from the Pythagorean theorem, that is,

* The ellipse, parabola and hyperbola are topics of analytic geometry and may be skipped. Here, we give a brief description of each because they find wide applications in science and technology. For instance, we hear phrases such as the “elliptic” orbit of a satellite or planet, a “parabolic” reflector antenna of a radar system, that comets move in “hyperbolic” orbits, etc.

$$c = \sqrt{a^2 - b^2} \tag{3.33}$$

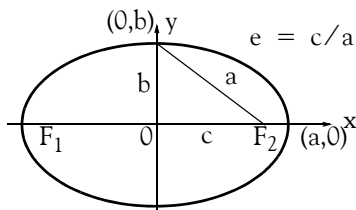


Figure 3.23. Computation of the foci locations

We observe that when the major and minor axes are equal, the ellipse becomes a circle.

Note 3.4

In Figure 3.23, the ratio $e = c/a$ is called the *eccentricity* of the ellipse; it indicates the degree of departure from circularity. In other words, if we keep point a on the major axis fixed and we vary the line segment c over the range $0 \leq c \leq a$, we will obtain ellipses which will vary in shape. An ellipse becomes a circle when $c = 0$, and a straight line when $c = a$.

A plane curve formed by the locus of points equidistant from a fixed line, and a fixed point not on the line, is called *parabola*. Figure 3.24 shows a parabola. The fixed line is called *directrix*, and the fixed point *focus*. The lines PF and PQ at any point on the parabola, are equal.

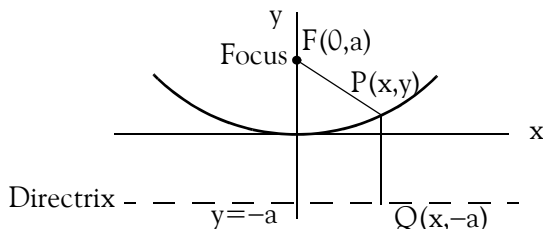


Figure 3.24. Parabola

For the parabola of Figure 3.24,

$$x^2 = 4ay \tag{3.34}$$

The locus of points for which the difference of the distances from two given points, called foci, is a constant, is called *hyperbola*. Figure 3.25 below shows a hyperbola.

For the hyperbola of Figure 3.25,

$$PF_2 - PF_1 = 2a \tag{3.35}$$

$$e = c/a \tag{3.36}$$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{3.37}$$

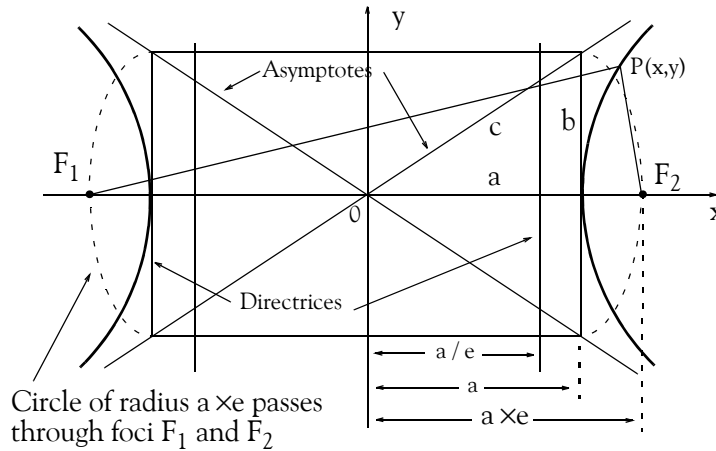


Figure 3.25. Hyperbola

In a hyperbola, the *eccentricity* e , can never be less than 1 .

Note 3.5

An *asymptote* is a line considered a limit to a curve in the sense that the perpendicular distance from a moving point on the curve to the line approaches zero as the point moves an infinite distance from the origin. Two such asymptotes are shown in Figure 3.25.

3.3 Solid Geometry Figures

A solid with six faces, each a parallelogram, and each being parallel to the opposite face, is called *parallelepiped*.

A solid with six faces, each a rectangle and each being parallel to the opposite face, is called *rectangular parallelepiped*. Figure 3.26 shows a rectangular parallelepiped.

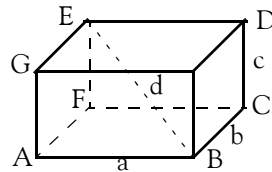


Figure 3.26. Rectangular Parallelepiped

For the rectangular parallelepiped of Figure 3.26, let d = diagonal, T = total surface area, and V = volume. Then,

$$d = \sqrt{a^2 + b^2 + c^2} \tag{3.38}$$

$$T = 2(ab + bc + ca) \tag{3.39}$$

$$V = abc \tag{3.40}$$

A regular solid that has six *congruent** square faces, is called a *cube*. Alternately, a cube is a special case of a rectangular parallelepiped where all faces are squares. Figure 3.27 shows a cube.

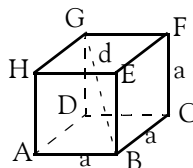


Figure 3.27. Cube

For the cube of Figure 3.27, if we let d = diagonal, T = total surface area, and V = volume, then,

$$d = a\sqrt{3} \tag{3.41}$$

$$T = 6a^2 \tag{3.42}$$

$$V = a^3 \tag{3.43}$$

A solid figure, whose bases or ends have the same size and shape, are parallel to one another, and each of whose sides is a parallelogram, is called *prism*. Figure 3.28 shows triangular, quadrilateral and pentagonal prisms.

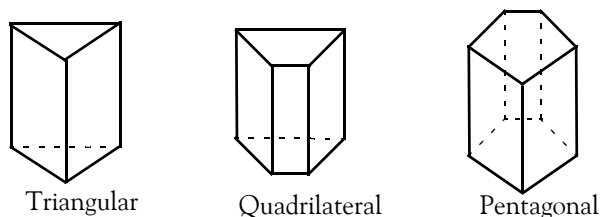


Figure 3.28. Different shapes of prisms

Prisms can assume different shapes; therefore, no general formulas for the total surface and volume are given here. The interested reader may find this information in mathematical tables.

A solid figure with a polygonal base, and triangular faces that meet at a common point, is called *pyramid*. Figure 3.29 below shows a pyramid whose base is a square.

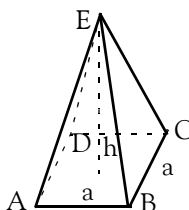


Figure 3.29. Pyramid with square base

* Two or more coplanar (on the same plane) figures, such as triangles, rectangles, etc., are said to be congruent when they coincide exactly when superimposed (placed one over the other).

Chapter 3 Fundamentals of Geometry

For the pyramid of Figure 3.29, the volume V is given by

$$V = \frac{1}{3}(\text{Area of base}) \times \text{height} = \frac{1}{3}a^2h \quad (3.44)$$

The part of a pyramid between two parallel planes cutting the pyramid, especially the section between the base and a plane parallel to the base of a pyramid, is called *frustum of a pyramid*. Figure 3.30 below shows a frustum of a pyramid.

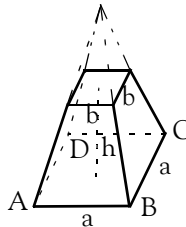


Figure 3.30. *Frustum of a pyramid.*

For the frustum of a pyramid of Figure 3.30, the volume V is

$$V = \frac{1}{3}h(a^2 + b^2 + ab) \quad (3.45)$$

A surface generated by a straight line that moves along a closed curve while always passing through a fixed point, is called a *cone*. The straight line is called the *generatrix*, the fixed point is called the *vertex*, and the closed curve is called the *directrix*. If the generatrix is of infinite length, it generates two conical surfaces on opposite sides of the vertex. If the directrix of the cone is a circle, the cone is referred to as a *circular cone*. Figure 3.31 below shows a right circular cone.

For the cone of Figure 3.31, the volume V is

$$V = \frac{1}{3}\pi r^2h \quad (3.46)$$

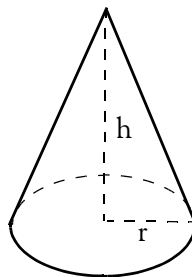


Figure 3.31. *Right circular cone*

The part of a cone between two parallel planes cutting the cone, especially the section between the base and a plane parallel to the base, is called *frustum of a cone*. Figure 3.32 below shows a frustum of a cone.

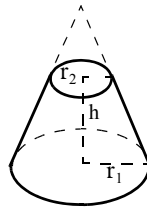


Figure 3.32. Frustum of right circular cone

For the frustum of the cone in Figure 3.32, the volume V is

$$V = \frac{1}{3}\pi h(r_1^2 + r_2^2 + r_1r_2) \quad (3.47)$$

The surface generated by a straight line intersecting and moving along a closed plane curve, the directrix, while remaining parallel to a fixed straight line that is not on or parallel to the plane of the directrix, is called *cylinder*. Figure 3.33 below shows a *right circular cylinder*.

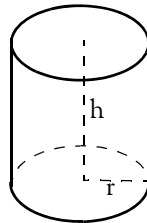


Figure 3.33. Right circular cylinder

For the cylinder of Figure 3.33, the volume V is

$$V = \pi r^2 h \quad (3.48)$$

A three-dimensional surface, all points of which are equidistant from a fixed point is called *sphere*. Figure 3.34 below shows a sphere.

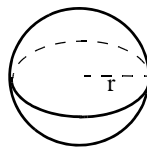


Figure 3.34. Sphere

For the sphere of Figure 3.34, the volume V is

$$V = \frac{4}{3}\pi r^3 \quad (3.49)$$

3.4 Using Spreadsheets to Find Areas of Irregular Polygons

The area of any polygon can be found from the relation

$$\text{Area} = \frac{1}{2}[(x_0 y_1 + x_1 y_2 + x_2 y_3 + \dots + x_{n-1} y_n + x_n y_0) - (x_1 y_0 + x_2 y_1 + x_3 y_2 + \dots + x_n y_{n-1} + x_0 y_n)] \quad (3.50)$$

where x_i and y_i are the coordinates of the vertices (corners) of the polygon whose area is to be found. The formula of 3.50 dictates that we must go around the polygon in a counterclockwise direction. The computations can be made easy with a spreadsheet, such as Excel. The procedure is illustrated with the following example.

Example 3.6

A land developer bought a parcel from a seller who did not know its area. Instead, the seller gave the buyer the sketch of Figure 3.35 below. The distances are in miles. Compute the area of this parcel.

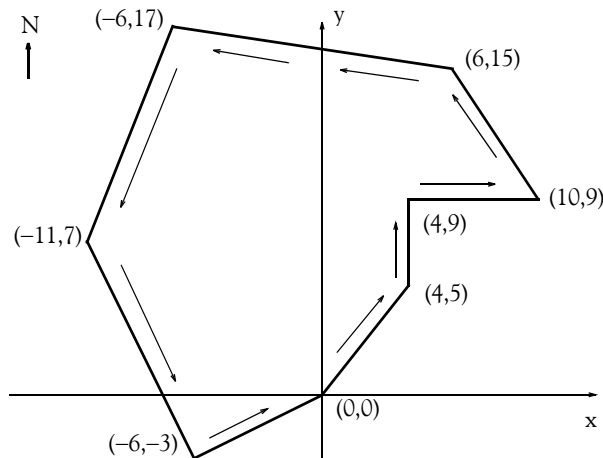


Figure 3.35. Figure for Example 3.6

Solution:

We arbitrarily choose the origin $(0, 0)$ as our starting point, and we go around the polygon in a counterclockwise direction as indicated in Figure 3.35.

The Excel spreadsheet of Figure 3.36 below, computes the area of any polygon with up to 10 vertices.

Area Calculation for Example 3.6				Area =	235
x ₀ =	0	y ₀ =	0		
x ₁ =	4	y ₁ =	5		
x ₂ =	4	y ₂ =	9		
x ₃ =	10	y ₃ =	9		
x ₄ =	6	y ₄ =	15		
x ₅ =	-6	y ₅ =	17		
x ₆ =	-11	y ₆ =	7		
x ₇ =	-6	y ₇ =	-3		
x ₈ =		y ₈ =			
x ₉ =		y ₉ =			
x ₁₀ =		y ₁₀ =			

Figure 3.36. Spreadsheet to compute areas of irregular polygons.

The formula in F3 is

$$=0.5*((B3*D4+B4*D5+B5*D6+B6*D7+B7*D8+B8*D9+B9*D10+B10*D11+B11*D12+B12*D13+B13*D3)-(B4*D3+B5*D4+B6*D5+B7*D6+B8*D7+B9*D8+B10*D9+B11*D10+B12*D11+B13*D12+B3*D13))$$

and this is an application of the general formula of relation (3.50).

The answer is displayed in F2. Thus, the total area is 235 square miles. To verify that the spreadsheet displays the correct result, we can divide the given polygon into triangles and rectangles, as shown in Figure 3.37, find the area of each using the formulas of (3.3), Page 3–4, and (3.21), Page 3–13, and add these areas. The area of each triangle or rectangle, is computed and the values are entered in Table 3.1.

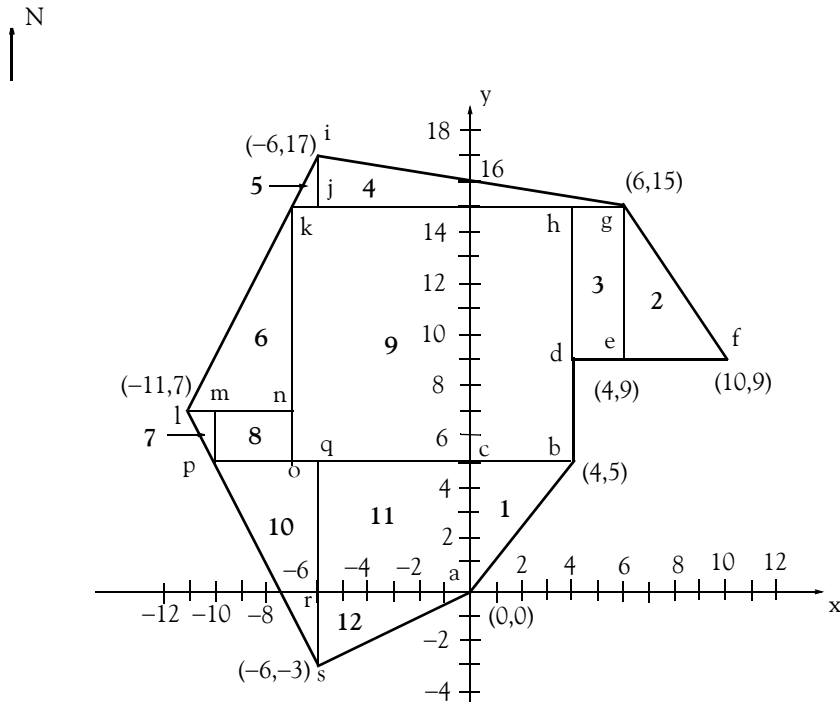


Figure 3.37. Division of Figure 3.35 to rectangles and triangles

TABLE 3.1 Computation of the total area of Figure 3.37

Triangle/Rectangle	Area	Triangle/Rectangle	Area
1 (abc)	10	7 (lmp)	1
2 (efg)	12	8 (mnop)	6
3 (degh)	12	9 (bhko)	110
4 (gij)	12	10 (pqs)	16
5 (ijk)	1	11 (acqr)	30
6 (kln)	16	12 (ars)	9
Total Area (1 through 12) in square miles			235

3.5 Summary

- *Geometry* is the mathematics of the properties, measurement, and relationships of points, lines, angles, surfaces, and solids. The science of geometry also includes analytic geometry, descriptive geometry, fractal geometry, and non-Euclidean geometry. In this text we are only be concerned with the Euclidean Geometry.
- A *polygon* is a closed plane figure bounded by three or more line segments.
- A *triangle* is a plane figure that is formed by connecting three points, not all in a straight line, by straight line segments. It is a special case of a polygon.
- The *perimeter* of a polygon is defined as the closed curve that bounds the plane area of the polygon. In other words, it is the entire length of the polygon's boundary.
- The *area* of a polygon is the space inside the boundary of the polygon, and it is expressed in square units.
- A *perpendicular* line is one that intersects another line to form a 90° angle.
- The line that joins a vertex of a triangle to the midpoint of the opposite side is called the *median*.
- The sum of the three angles of any triangle is 180° .
- A triangle containing an angle of 90° is a *right triangle*. In a right triangle, the side opposite to the right angle is called the *hypotenuse* of the right triangle. If c is the hypotenuse of a right triangle, then

$$c^2 = a^2 + b^2$$

This relation is known as the *Pythagorean theorem*.

- If c is the hypotenuse of a right triangle, the area of this triangle is found from the relation

$$\text{Area} = \frac{1}{2} \cdot a \cdot b$$

- A triangle that has two equal sides is called *isosceles* triangle.
- A triangle that has all sides equal is called *equilateral triangle*. For an equilateral triangle,

$$\angle A = \angle B = \angle C = 60^\circ$$

$$\text{Area} = \frac{1}{4} a^2 \sqrt{3}$$

where A , B , and C are the three angles and a is one of the three equal sides of the triangle.

Chapter 3 Fundamentals of Geometry

- *Congruent triangles* are those that are identical in shape and size. If two triangles are congruent, the following statements are true:
- Two lines are said to be in *parallel*, when the distance between them is the same everywhere. In other words, two lines on the same plane that never intersect, are said to be parallel lines.
- A *transversal line* is a line that intersects two or more parallel lines. Corresponding angles on the same side of the transversal are congruent.
- *Similar triangles* are triangles that have corresponding angles that are congruent (equal) while the lengths of the corresponding sides are in proportion. All congruent triangles are similar, but not all similar triangles are congruent.
- A plane figure with four sides and four angles is called *quadrilateral*. The lines joining two non-adjacent vertices of a quadrilateral are called *diagonals*. In any quadrilateral, the sum of the four angles is 360° . The diagonals of a quadrilateral with consecutive sides a , b , c , and d are perpendicular if and only if

$$a^2 + c^2 = b^2 + d^2$$

- A four-sided plane figure with opposite sides in parallel, is called *parallelogram*.
- The perpendicular distance from the center of a regular polygon to any of its sides is called *apothem*.
- An equilateral parallelogram is called *rhombus*. It is a special case of a parallelogram where all sides are equal.
- A four-sided plane figure with four right angles, is a *rectangle*. It is a special case of a parallelogram where all angles are equal, that is, 90° each.
- A plane figure with four equal sides and four equal angles, is a *square*. It is a special case of a rectangle where all sides are equal.
- A quadrilateral with two unequal parallel sides is a *trapezoid*. Figure 3.20 shows a *trapezoid*.
- A plane curve that is everywhere equidistant from a given fixed point, referred to as the *center*, is called a *circle*. The line segment that joins the center of a circle with any point on its circumference is called *radius*, and it is denoted with the letter r . A straight line segment passing through the center of a circle, and terminating at the circumference is called *diameter*, and is denoted with the letter d . Obviously, $d = 2r$. The perimeter of a circle is called *circumference*.
- The Greek letter π represents the ratio of the circumference of a circle to its diameter. This ratio is an irrational (endless) number, so the decimal places go on infinitely without repeating. It is approximately equal to $22/7$; to five decimal places, is equal to 3.14159.

- *Locus* is defined as the set of all points whose coordinates satisfy a single equation or one or more algebraic conditions.
- The locus of points for which the sum of the distances from each point to two fixed points is equal forms an *ellipse*. The two fixed points are called *foci* (plural for focus). The foci are always on the *major* (longer) axis; they are located symmetrically from the center O that is the intersection of the major and minor (shorter) axes. When the major and minor axes are equal, the ellipse becomes a circle. The *eccentricity* of an ellipse indicates the degree of departure from circularity.
- A plane curve formed by the locus of points equidistant from a fixed line, and a fixed point not on the line, is called *parabola*. The fixed line is called *directrix*, and the fixed point *focus*.
- The locus of points for which the difference of the distances from two given points, called foci, is a constant, is called *hyperbola*. In a hyperbola, the *eccentricity* e , can never be less than 1.
- An *asymptote* is a line considered a limit to a curve in the sense that the perpendicular distance from a moving point on the curve to the line approaches zero as the point moves an infinite distance from the origin.
- A solid with six faces, each a parallelogram, and each being parallel to the opposite face, is called *parallelepiped*. A solid with six faces, each a rectangle and each being parallel to the opposite face, is called *rectangular parallelepiped*.
- A regular solid that has six congruent square faces, is called a *cube*. Alternately, a cube is a special case of a rectangular parallelepiped where all faces are squares.
- A solid figure, whose bases or ends have the same size and shape, are parallel to one another, and each of whose sides is a parallelogram, is called *prism*.
- A solid figure with a polygonal base, and triangular faces that meet at a common point, is called *pyramid*. The part of a pyramid between two parallel planes cutting the pyramid, especially the section between the base and a plane parallel to the base of a pyramid, is called *frustum of a pyramid*.
- A surface generated by a straight line that moves along a closed curve while always passing through a fixed point, is called a *cone*. The straight line is called the *generatrix*, the fixed point is called the *vertex*, and the closed curve is called the *directrix*. If the generatrix is of infinite length, it generates two conical surfaces on opposite sides of the vertex. If the directrix of the cone is a circle, the cone is referred to as a *circular cone*. The part of a cone between two parallel planes cutting the cone, especially the section between the base and a plane parallel to the base, is called *frustum of a cone*.
- The surface generated by a straight line intersecting and moving along a closed plane curve, the directrix, while remaining parallel to a fixed straight line that is not on or parallel to the plane of the directrix, is called *cylinder*.

Chapter 3 Fundamentals of Geometry

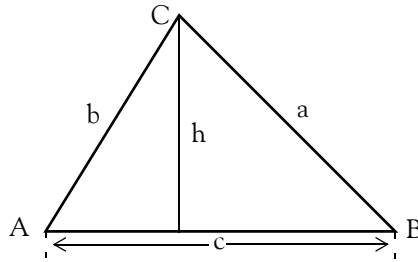
- A three-dimensional surface, all points of which are equidistant from a fixed point is called *sphere*.
- The area of any polygon can be found from the relation

$$\text{Area} = \frac{1}{2}[(x_0 y_1 + x_1 y_2 + x_2 y_3 + \dots + x_{n-1} y_n + x_n y_0) \\ -(x_1 y_0 + x_2 y_1 + x_3 y_2 + \dots + x_n y_{n-1} + x_0 y_n)]$$

where x_i and y_i are the coordinates of the vertices (corners) of the polygon whose area is to be found. This relation dictates that we must go around the polygon in a counterclockwise direction. The computations can be made easy with a spreadsheet, such as Excel.

3.6 Exercises

1. Compute the area of the general triangle shown below in terms of the side c and height h .



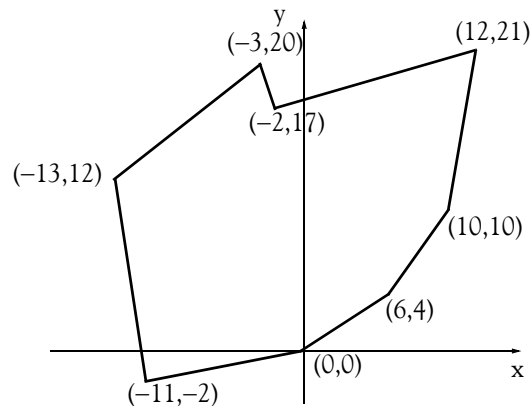
2. For an equilateral triangle below, prove that:

a.
$$\angle A = \angle B = \angle C = 60^\circ$$

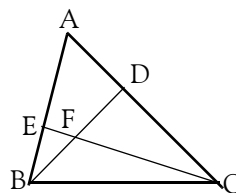
b.
$$\text{Area} = \frac{\sqrt{3}}{4} a^2$$

where a is any of the three equal sides.

3. Compute the area of the irregular polygon shown below with the relation (3.50), Page 3–22.



4. For the triangle shown below, $BD = DC$, $\angle CDB = 90^\circ$, and $DF = AD$. Prove that the triangles ADB and CDF are congruent.

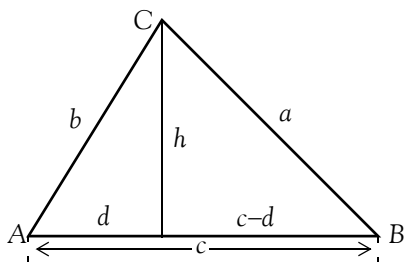


Chapter 3 Fundamentals of Geometry

5. Prove the Pythagorean theorem.

3.7 Solutions to End-of-Chapter Exercises

1.



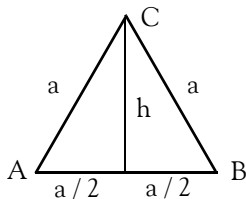
$$\text{Area} = \frac{1}{2} \cdot d \cdot h + \frac{1}{2} \cdot (c-d) \cdot h = \frac{1}{2} \cdot d \cdot h + \frac{1}{2} \cdot c \cdot h - \frac{1}{2} \cdot d \cdot h = \frac{1}{2} \cdot c \cdot h$$

2.

a. For any triangle $\angle A + \angle B + \angle C = 180^\circ$, and since the sides a , b , and c are equal, it follows that the angles opposite to the sides are also equal. Therefore,

$$\angle A = \angle B = \angle C = 60^\circ$$

b.



$$\text{Area} = \frac{1}{2} \cdot \frac{a}{2} \cdot h + \frac{1}{2} \cdot \frac{a}{2} \cdot h = \frac{a}{2} \cdot h \quad (1)$$

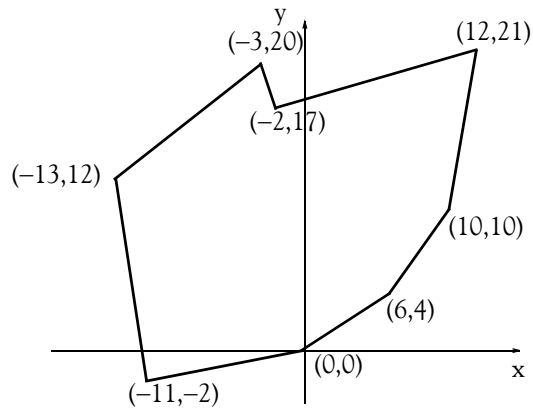
$$h^2 = a^2 - \left(\frac{a}{2}\right)^2 = a^2 - \frac{a^2}{4} = \frac{3}{4} \cdot a^2$$

$$h = \sqrt{\frac{3}{4}} \cdot a = \frac{\sqrt{3}}{2} \cdot a$$

and by substitution into (1)

$$\text{Area} = \frac{\sqrt{3}}{4} a^2$$

3.



$$\text{Area} = \frac{1}{2} [(x_0 y_1 + x_1 y_2 + x_2 y_3 + \dots + x_{n-1} y_n + x_n y_0) - (x_1 y_0 + x_2 y_1 + x_3 y_2 + \dots + x_n y_{n-1} + x_0 y_n)]$$

We arbitrarily choose the origin (0, 0) as our starting point, and we go around the polygon in a counterclockwise direction as in Example 3.6.

The formula in F2 is

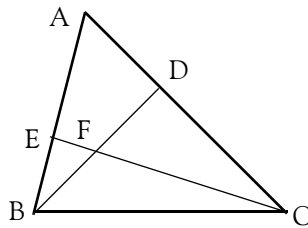
$$\begin{aligned} &= 0.5 * ((B3 * D4 + B4 * D5 + B5 * D6 + B6 * D7 + B7 * D8 + B8 * D9 + B9 * D10 + B10 * D11 \\ &+ B11 * D12 + B12 * D13 + B13 * D3) \\ &- (B4 * D3 + B5 * D4 + B6 * D5 + B7 * D6 + B8 * D7 + B9 * D8 + B10 * D9 + B11 * D10 + B12 * D11 \\ &+ B13 * D12 + B3 * D13)) \end{aligned}$$

and this represents the general formula for the area above.

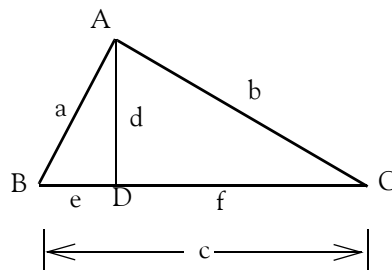
The answer is displayed in F2 in the Excel spreadsheet below. Thus, the total area is 374.5 square units.

	A	B	C	D	E	F	
1	Area Calculation for Exercise 3						
2					Area =	374.5	
3	$x_0 =$	0	$y_0 =$	0			
4	$x_1 =$	6	$y_1 =$	4			
5	$x_2 =$	10	$y_2 =$	10			
6	$x_3 =$	12	$y_3 =$	21			
7	$x_4 =$	-2	$y_4 =$	17			
8	$x_5 =$	-3	$y_5 =$	20			
9	$x_6 =$	-13	$y_6 =$	12			
10	$x_7 =$	-11	$y_7 =$	-2			

4. Since $BD = DC$, $\angle CDB = 90^\circ$, and $DF = AD$. by the Pythagorean theorem the hypotenuses AB and CF of triangles ADB and DFC are equal also. Therefore, these triangles are congruent.



5. Consider the right triangle shown below where $\angle A = 90^\circ$, BC is the hypotenuse, and AD is perpendicular to BC .



We observe that triangles ABD and ABC are similar since both are right triangles and $\angle B$ is common in both triangles. Likewise, triangles ADC and ABC are similar. Since triangles ABD and ABC are similar, it follows that

Chapter 3 Fundamentals of Geometry

$$\frac{c}{a} = \frac{a}{e} \text{ or } ce = a^2 \quad (1)$$

Also since triangles ADC and ABC are similar, it follows that

$$\frac{c}{b} = \frac{b}{f} \text{ or } cf = b^2 \quad (2)$$

Addition of (1) with (2) above yields

$$ce + cf = a^2 + b^2$$

or

$$c(e + f) = a^2 + b^2$$

and since

$$e + f = c$$

we obtain the Pythagorean theorem relation

$$c^2 = a^2 + b^2$$

Chapter 4

Fundamentals of Plane Trigonometry

This chapter discusses the basics of trigonometry. It is intended for readers who need to know the basics of plane trigonometry*. Readers with a strong mathematical background may skip this chapter. Others will find it useful, as well as a convenient source for review.

4.1 Introduction

Trigonometry is the branch of mathematics that is concerned with the relationships between the sides and the angles of triangles. We will now define angle and radian formally.

An *angle* is an angular unit of measure with *vertex* (the point at which the sides of an angle intersect) at the center of a circle, and with sides that subtend (cut off) part of the circumference. If the subtended arc is equal to one-fourth of the total circumference, the angular unit is a *right angle*. If the arc equals half the circumference, the unit is a *straight angle* (an angle of 180°). If the arc equals $1/360$ of the circumference, the angular unit is one *degree*. An angle between 0° and less than 90° is called an *acute* angle, while an angle between 90° and less than 180° is called an *obtuse* angle.

Each degree is subdivided into 60 equal parts called *minutes*, and each minute is subdivided into 60 equal parts called *seconds*. The symbol for degree is $^\circ$; for minutes, $'$; and for seconds, $''$. In trigonometry, it is customary to denote angles with the Greek letters θ (theta) or ϕ (phi).

Another unit of angular measure is the *radian*; it is illustrated with the circle of Figure 4.1.

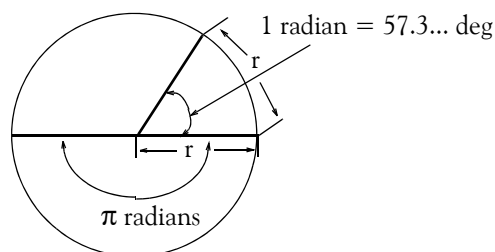


Figure 4.1. Definition of the radian

* Another branch of trigonometry, is the so-called **spherical trigonometry** which deals with triangles that are sections of the surface of a sphere. Spherical trigonometry is used extensively in navigation and space technology. It will not be discussed in this text.

Chapter 4 Fundamentals of Plane Trigonometry

As shown in Figure 4.1, the radian is a circular angle subtended by an arc equal in length to the radius of the circle whose radius is r units in length. The circumference of a circle is $2\pi r$ units; therefore, there are 2π or 6.283... radians in 360° . Also, $\pi = 3.14159...$ radians corresponds to 180° .

4.2 Trigonometric Functions

The trigonometric functions vary with the size of an angle. The six commonly used trigonometric functions can be defined in terms of one of the two acute angles in a right triangle. In addition to the *hypotenuse*, a right triangle has two sides, one adjacent to one of the acute angles, the other opposite as shown in Figure 4.2. The sine, cosine, tangent, cotangent, secant, and cosecant functions are defined in (4.1) through (4.6).

For any angle, the numerical values of the trigonometric functions can be approximated by drawing the angle, measuring, and then calculating the ratios. Actually, most numerical values are found by using formulas, called *trigonometric identities*. If the numerical values of various angles are shown on a graph, such as the sinusoidal graph in Chapter 1, it is clear that the functions are periodic.* A particular angle produces a value equal to the values of many other angles. A single angle is chosen as the basic, or *principal*, angle for that value.

4.3 Trigonometric Functions of an Acute Angle

For the right triangle of Figure 4.2, the trigonometric functions are defined as

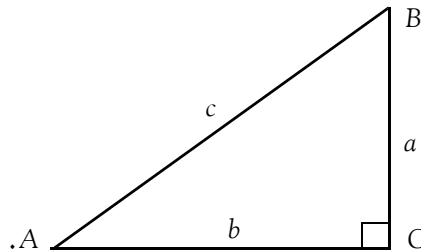


Figure 4.2. Right triangle for defining the trigonometric functions

$$\text{Sine of Angle } A = \sin A = \frac{a}{c} \quad (4.1)$$

$$\text{Cosine of Angle } A = \cos A = \frac{b}{c} \quad (4.2)$$

* A periodic function is repeated at regular intervals. The least interval in the range of the independent variable of a periodic function of a real variable in which all possible values of the dependent variable are assumed, is called the **period**.

$$\text{Tangent of Angle } A = \tan A = \frac{a}{b} \quad (4.3)$$

$$\text{Cotangent of Angle } A = \cot A = \frac{b}{a} \quad (4.4)$$

$$\text{Secant of Angle } A = \sec A = \frac{c}{b} \quad (4.5)$$

$$\text{Cosecant of Angle } A = \csc A = \frac{c}{a} \quad (4.6)$$

From (4.2) and (4.5) we observe that

$$\sec A = \frac{1}{\cos A} \quad (4.7)$$

Also, from (4.1) and (4.6)

$$\csc A = \frac{1}{\sin A} \quad (4.8)$$

and from (4.3) and (4.4)

$$\cot A = \frac{1}{\tan A} \quad (4.9)$$

Finally, from (4.1), (4.2) and (4.3)

$$\tan A = \frac{\sin A}{\cos A} \quad (4.10)$$

4.4 Trigonometric Functions of an Any Angle

Let us consider the rectangular Cartesian coordinate system of Figure 4.3, where the axes x and y divide the plane into four *quadrants*. A quadrant is defined as any of the four areas into which a plane is divided by the reference set of axes in a Cartesian coordinate system, designated *first*, *second*, *third*, and *fourth*, counting counterclockwise from the area in which both coordinates are positive.

Figure 4.4 shows how an angle is generated in the first quadrant. The angle is formed by the ray OP_0 (not shown) which is assumed to be in its initial position, that is, along the positive side of the x -axis. It is rotated counterclockwise about its origin O onto ray OP , and forms a positive angle $^*\theta_1$. This angle is greater than zero but less than 90° .

* A negative angle is generated when the ray is rotated clockwise.

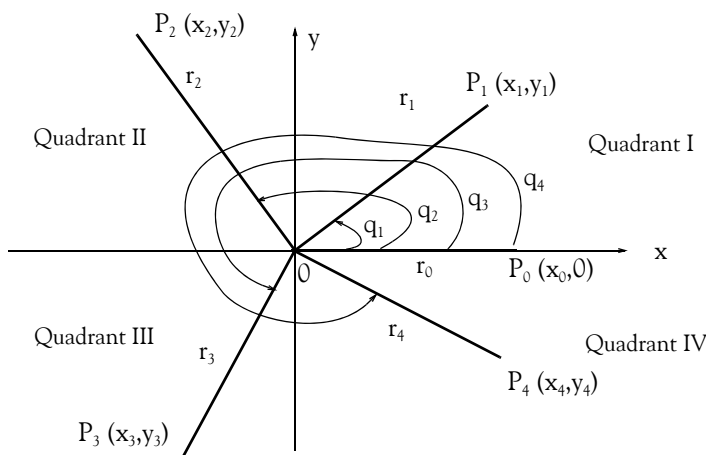


Figure 4.3. Quadrants in Cartesian Coordinates

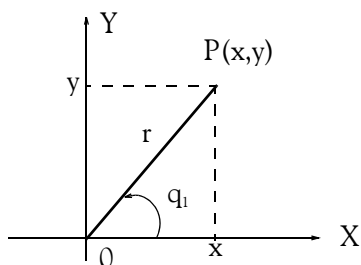


Figure 4.4. Definition of trigonometric functions of angles in first quadrant

Denoting the distance OP as r^* we obtain:

$$\sin \theta_1 = \frac{y}{r} \quad (4.11)$$

$$\cos \theta_1 = \frac{x}{r} \quad (4.12)$$

$$\tan \theta_1 = \frac{y}{x} \quad (4.13)$$

$$\cot \theta_1 = \frac{x}{y} \quad (4.14)$$

$$\sec \theta_1 = \frac{r}{x} \quad (4.15)$$

$$\csc \theta_1 = \frac{r}{y} \quad (4.16)$$

* Here, we are only concerned with the magnitude, i.e., absolute value of r . However, the line segments x and y must be defined as positive or negative. Thus, in (4.11) through (4.16), the values of x and y are both positive since x lies to the right of the reference point O and y lies above it.

Now, let us consider the rectangular Cartesian coordinate system of Figure 4.5 where the ray r lies in the second quadrant, and the angle θ_2 is greater than 90° but less than 180° . Here, x is negative whereas y is positive.

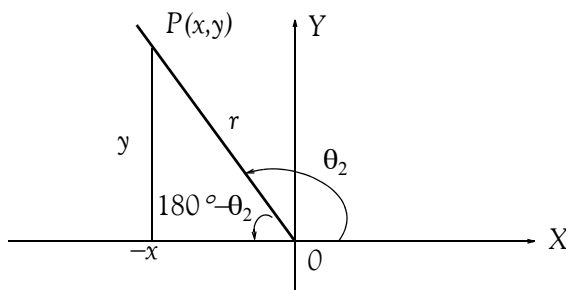


Figure 4.5. Definition of trigonometric functions of angles in second quadrant.

Using the trigonometric reduction formulas (4.55) through (4.62), page 4–8, we obtain

$$\sin(180^\circ - \theta_2) = \sin \theta_2 = \frac{y}{r} \quad (4.17)$$

$$\cos(180^\circ - \theta_2) = -\cos \theta_2 = \frac{-x}{r} \quad (4.18)$$

$$\tan(180^\circ - \theta_2) = -\tan \theta_2 = \frac{y}{-x} \quad (4.19)$$

$$\cot(180^\circ - \theta_2) = \frac{1}{\tan(180^\circ - \theta_2)} = \frac{1}{-\tan \theta_2} = -\cot \theta_2 = \frac{-x}{y} \quad (4.20)$$

$$\sec(180^\circ - \theta_2) = \frac{1}{\cos(180^\circ - \theta_2)} = \frac{1}{-\cos \theta_2} = -\sec \theta_2 = \frac{r}{-x} \quad (4.21)$$

$$\csc(180^\circ - \theta_2) = \frac{1}{\sin(180^\circ - \theta_2)} = \frac{1}{\sin \theta_2} = \csc \theta_2 = \frac{r}{y} \quad (4.22)$$

and we observe that the cosine, tangent, cotangent and secant functions are negative in the second quadrant.

We could follow the same procedure to define the trigonometric functions in the third and fourth quadrants. Instead, we summarize their signs in Table 4.1.

Chapter 4 Fundamentals of Plane Trigonometry

TABLE 4.1 Signs of the trigonometric functions in different quadrants.

Quadrant	Sine	Cosine	Tangent	Cotangent	Secant	Cosecant
I	+	+	+	+	+	+
II	+	-	-	-	-	+
III	-	-	+	+	-	-
IV	-	+	-	-	+	-

From Figures 4.4, Page 4-4, and 4.5, Page 4-5, we observe that the ray r is the hypotenuse of the right triangles formed by it and the line segments x and y . Therefore, x and y can never be greater in length than r . Accordingly, the sine and the cosine functions can never be greater than 1, that is, they can only vary from 0 to ± 1 inclusive. Table 4.2 shows the limits, that is, lowest and highest values which the trigonometric functions can achieve.

TABLE 4.2 Limits of the trigonometric functions

Quadrant	Sine	Cosine	Tangent	Cotangent	Secant	Cosecant
I	0 to +1	+1 to 0	0 to $+\infty$	$+\infty$ to 0	+1 to $+\infty$	$+\infty$ to +1
II	+1 to 0	0 to -1	$-\infty$ to 0	0 to $-\infty$	$-\infty$ to -1	+1 to $+\infty$
III	0 to -1	-1 to 0	0 to $+\infty$	$+\infty$ to 0	-1 to $-\infty$	$-\infty$ to -1
IV	-1 to 0	0 to +1	$-\infty$ to 0	0 to $-\infty$	$+\infty$ to +1	-1 to $-\infty$

4.5 Fundamental Relations and Identities

The formulas of (4.27) through (4.54), pages 7 through 9 show the sines and cosines of common angles. From these, we can derive the angles of the other trigonometric functions using the relations

$$\tan\theta = \frac{\sin\theta}{\cos\theta} \quad (4.23)$$

$$\cot\theta = \frac{\cos\theta}{\sin\theta} = \frac{1}{\tan\theta} \quad (4.24)$$

$$\sec\theta = \frac{1}{\cos\theta} \quad (4.25)$$

$$\csc\theta = \frac{1}{\sin\theta} \quad (4.26)$$

The formulas below give the values of special angles in both degrees and radians.

$$\cos 0^\circ = \cos 360^\circ = \cos 2\pi = 1 \quad (4.27)$$

$$\cos 30^\circ = \cos \frac{\pi}{6} = \frac{\sqrt{3}}{2} = 0.866 \quad (4.28)$$

$$\cos 45^\circ = \cos \frac{\pi}{4} = \frac{\sqrt{2}}{2} = 0.707 \quad (4.29)$$

$$\cos 60^\circ = \cos \frac{\pi}{3} = \frac{1}{2} = 0.5 \quad (4.30)$$

$$\cos 90^\circ = \cos \frac{\pi}{2} = 0 \quad (4.31)$$

$$\cos 120^\circ = \cos \frac{2\pi}{3} = \frac{-1}{2} = -0.5 \quad (4.32)$$

$$\cos 150^\circ = \cos \frac{5\pi}{6} = \frac{-\sqrt{3}}{2} = -0.866 \quad (4.33)$$

$$\cos 180^\circ = \cos \pi = -1 \quad (4.34)$$

$$\cos 210^\circ = \cos \frac{7\pi}{6} = \frac{-\sqrt{3}}{2} = -0.866 \quad (4.35)$$

$$\cos 225^\circ = \cos \frac{5\pi}{4} = \frac{-\sqrt{2}}{2} = -0.707 \quad (4.36)$$

$$\cos 240^\circ = \cos \frac{4\pi}{3} = \frac{-1}{2} = -0.5 \quad (4.37)$$

$$\cos 270^\circ = \cos \frac{3\pi}{2} = 0 \quad (4.38)$$

$$\cos 300^\circ = \cos \frac{5\pi}{3} = 0.5 \quad (4.39)$$

$$\cos 330^\circ = \cos \frac{11\pi}{6} = 0.866 \quad (4.40)$$

$$\sin 0^\circ = \sin 360^\circ = \sin 2\pi = 0 \quad (4.41)$$

$$\sin 30^\circ = \sin \frac{\pi}{6} = \frac{1}{2} = 0.5 \quad (4.42)$$

$$\sin 45^\circ = \sin \frac{\pi}{4} = \frac{\sqrt{2}}{2} = 0.707 \quad (4.43)$$

$$\sin 60^\circ = \sin \frac{\pi}{3} = \frac{\sqrt{3}}{2} = 0.866 \quad (4.44)$$

$$\sin 90^\circ = \sin \frac{\pi}{2} = 1 \quad (4.45)$$

Chapter 4 Fundamentals of Plane Trigonometry

$$\sin 120^\circ = \sin \frac{2\pi}{3} = \frac{\sqrt{3}}{2} = 0.866 \quad (4.46)$$

$$\sin 150^\circ = \sin \frac{5\pi}{6} = \frac{1}{2} = 0.5 \quad (4.47)$$

$$\sin 180^\circ = \sin \pi = 0 \quad (4.48)$$

$$\sin 210^\circ = \sin \frac{7\pi}{6} = \frac{-1}{2} = -0.5 \quad (4.49)$$

$$\sin 225^\circ = \sin \frac{5\pi}{4} = \frac{-\sqrt{2}}{2} = -0.707 \quad (4.50)$$

$$\sin 240^\circ = \sin \frac{4\pi}{3} = \frac{-\sqrt{3}}{2} = -0.866 \quad (4.51)$$

$$\sin 270^\circ = \sin \frac{3\pi}{2} = -1 \quad (4.52)$$

$$\sin 300^\circ = \sin \frac{5\pi}{3} = \frac{-\sqrt{3}}{2} = -0.866 \quad (4.53)$$

$$\sin 330^\circ = \sin \frac{11\pi}{6} = \frac{-1}{2} = -0.5 \quad (4.54)$$

Relations (4.55) through (4.62) below are known as *trigonometric reduction formulas*.

$$\cos(-\theta) = \cos \theta \quad (4.55)$$

$$\cos(90^\circ + \theta) = -\sin \theta \quad (4.56)$$

$$\cos(180^\circ - \theta) = -\cos \theta \quad (4.57)$$

$$\sin(-\theta) = -\sin \theta \quad (4.58)$$

$$\sin(90^\circ + \theta) = \cos \theta \quad (4.59)$$

$$\sin(180^\circ - \theta) = \sin \theta \quad (4.60)$$

$$\tan(90^\circ + \theta) = -\cot \theta \quad (4.61)$$

$$\tan(180^\circ - \theta) = -\tan \theta \quad (4.62)$$

Relations (4.63) through (4.68) below are known as *angle-sum, angle-difference relations*.

$$\cos(\theta + \phi) = \cos \theta \cos \phi - \sin \theta \sin \phi \quad (4.63)$$

$$\cos(\theta - \phi) = \cos \theta \cos \phi + \sin \theta \sin \phi \quad (4.64)$$

$$\sin(\theta + \phi) = \sin \theta \cos \phi + \cos \theta \sin \phi \quad (4.65)$$

$$\sin(\theta - \phi) = \sin\theta \cos\phi - \cos\theta \sin\phi \quad (4.66)$$

$$\tan(\theta + \phi) = \frac{\tan\theta + \tan\phi}{1 - \tan\theta \tan\phi} \quad (4.67)$$

$$\tan(\theta - \phi) = \frac{\tan\theta - \tan\phi}{1 + \tan\theta \tan\phi} \quad (4.68)$$

The trigonometric functions are also available in Excel. We will use them with the examples that follow.

Example 4.1

If $\theta = 60^\circ$ and $\phi = 45^\circ$, compute:

a. $\cos 105^\circ$ b. $\sin 15^\circ$ c. $\tan 105^\circ$

Solution:

a. From (4.29), (4.30), (4.43), and (4.44), Page 4–7, we obtain $\cos 45^\circ = \sqrt{2}/2$, $\cos 60^\circ = 1/2$, $\sin 45^\circ = (\sqrt{2})/2$, and $\sin 60^\circ = \sqrt{3}/2$. Then, with these relations and (4.63), Page 4–8, we obtain

$$\begin{aligned} \cos(\theta + \phi) &= \cos\theta \cos\phi - \sin\theta \sin\phi \\ \cos(60^\circ + 45^\circ) &= \cos 60^\circ \cdot \cos 45^\circ - \sin 60^\circ \cdot \sin 45^\circ \\ \cos(105^\circ) &= \frac{1}{2} \cdot \frac{\sqrt{2}}{2} - \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{4} - \frac{\sqrt{6}}{4} = \frac{1}{4} \cdot (\sqrt{2} - \sqrt{6}) \\ &= \frac{1}{4} \cdot (1.4142 - 2.4495) = \frac{1}{4} \cdot (-1.0353) = -0.2588 \end{aligned}$$

Check with Excel:

Excel requires that the angle be expressed in radians; therefore, we convert 105° to radians using the relation

$$\text{rad} = \text{deg} \cdot \frac{\pi}{180^\circ} \quad (4.69)$$

For this example, $105^\circ \cdot \frac{\pi}{180^\circ} = 1.8326$ radians. Then, in any cell of the spreadsheet we type the formula

=COS(1.8326) and Excel displays -0.2588 . This is the same answer we found using the relation of (4.63).

Chapter 4 Fundamentals of Plane Trigonometry

b. From (4.29), (4.30), (4.43), and (4.44), Page 4–7, we obtain $\cos 45^\circ = \sqrt{2}/2$, $\cos 60^\circ = 1/2$, $\sin 45^\circ = (\sqrt{2})/2$, and $\sin 60^\circ = \sqrt{3}/2$. Then, with these relations and (4.66), Page 4–9, we obtain

$$\begin{aligned}\sin(\theta - \phi) &= \sin\theta \cos\phi - \cos\theta \sin\phi \\ \sin(60^\circ - 45^\circ) &= \sin 60^\circ \cdot \cos 45^\circ - \cos 60^\circ \cdot \sin 45^\circ \\ \sin(15^\circ) &= \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{2}}{2} - \frac{1}{2} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{6}}{4} - \frac{\sqrt{2}}{4} = \frac{1}{4} \cdot (\sqrt{6} - \sqrt{2}) \\ &= \frac{1}{4} \cdot (2.4495 - 1.4142) = \frac{1}{4} \cdot (1.0353) = 0.2588\end{aligned}$$

Check with Excel:

For this example, $15^\circ \cdot \frac{\pi}{180^\circ} = 0.2618$ rad.. Then, in any cell of the spreadsheet we type the formula

=SIN(0.2618) and Excel displays 0.2588. This is the same answer we found with the trigonometric relation of (4.66).

We observe that the answer in part (b) is the same as in part (a), except that it is of opposite sign. This is not just a coincidence; it can be verified with (4.56), Page 4–8, which can be expressed also as $\sin\theta = -\cos(90^\circ + \theta)$. Then,

$$\sin 15^\circ = -\cos(90^\circ + 15^\circ) = -\cos(105^\circ)$$

c. From (4.23), Page 4–6,

$$\tan\theta = \frac{\sin\theta}{\cos\theta}$$

Then,

$$\tan 105^\circ = \frac{\sin 105^\circ}{\cos 105^\circ}$$

To find the value of $\sin 105^\circ$, we use (4.65), Page 4–8. Thus,

$$\begin{aligned}\sin(\theta + \phi) &= \sin\theta \cos\phi + \cos\theta \sin\phi \\ \sin(60^\circ + 45^\circ) &= \sin 60^\circ \cdot \cos 45^\circ + \cos 60^\circ \cdot \sin 45^\circ \\ \sin(105^\circ) &= \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{2}}{2} + \frac{1}{2} \cdot \frac{\sqrt{2}}{2} = \frac{1}{4}(\sqrt{6} + \sqrt{2}) \\ &= \frac{1}{4} \cdot (2.4495 + 1.4142) = \frac{1}{4} \cdot (3.8637) = 0.9659\end{aligned}$$

We already know the value of $\cos 105^\circ$ from part (a). Therefore,

$$\tan 105^\circ = \frac{\sin 105^\circ}{\cos 105^\circ} = \frac{0.9659}{-0.2588} = -3.7322 \quad (4.70)$$

Let us check this answer with Excel. As before, we first need to convert 105° to radians. This conversion yields $105^\circ \cdot \frac{\pi}{180^\circ} = 1.8326$ rad. Then, in any cell of the spreadsheet we type the formula

=TAN(1.8326) and Excel displays -3.732 . This is the same answer we found in (4.70).

Relations (4.71) through (4.76) below are known as *fundamental trigonometric identities*.

$$\cos^2\theta + \sin^2\theta = 1 \quad (4.71)$$

$$\cos 2\theta = \cos^2\theta - \sin^2\theta \quad (4.72)$$

$$\sin 2\theta = 2\sin\theta\cos\theta \quad (4.73)$$

$$\tan 2\theta = \frac{2\tan\theta}{1 - \tan^2\theta} \quad (4.74)$$

$$\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta) \quad (4.75)$$

$$\sin^2\theta = \frac{1}{2}(1 - \cos 2\theta) \quad (4.76)$$

Relations (4.77) through (4.80) below are known as *function-product relations*.

$$\cos\theta\cos\phi = \frac{1}{2}\cos(\theta + \phi) + \frac{1}{2}\cos(\theta - \phi) \quad (4.77)$$

$$\cos\theta\sin\phi = \frac{1}{2}\sin(\theta + \phi) - \frac{1}{2}\sin(\theta - \phi) \quad (4.78)$$

$$\sin\theta\cos\phi = \frac{1}{2}\sin(\theta + \phi) + \frac{1}{2}\sin(\theta - \phi) \quad (4.79)$$

$$\sin\theta\sin\phi = \frac{1}{2}\cos(\theta - \phi) - \frac{1}{2}\cos(\theta + \phi) \quad (4.80)$$

4.6 Triangle Formulas

Let Figure 4.6 be any triangle.

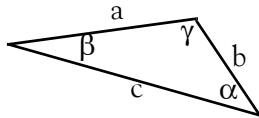


Figure 4.6. Triangle for definition of the laws of sines, cosines, and tangents.

Then,

By the *law of sines*:

$$\frac{a}{\sin \alpha} = \frac{b}{\sin \beta} = \frac{c}{\sin \gamma} \quad (4.81)$$

By the *law of cosines*:

$$a^2 = b^2 + c^2 - 2bc \cos \alpha \quad (4.82)$$

$$b^2 = a^2 + c^2 - 2ac \cos \beta \quad (4.83)$$

$$c^2 = a^2 + b^2 - 2ab \cos \gamma \quad (4.84)$$

By the *law of tangents*:

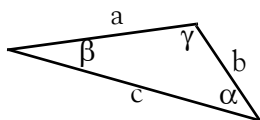
$$\frac{a-b}{a+b} = \frac{\tan \frac{1}{2}(\alpha - \beta)}{\tan \frac{1}{2}(\alpha + \beta)} \quad \frac{b-c}{b+c} = \frac{\tan \frac{1}{2}(\beta - \gamma)}{\tan \frac{1}{2}(\beta + \gamma)} \quad \frac{c-a}{c+a} = \frac{\tan \frac{1}{2}(\gamma - \alpha)}{\tan \frac{1}{2}(\gamma + \alpha)} \quad (4.85)$$

Example 4.2

For the triangle of Figure 4.7 below, find the length of side a using:

a. the law of sines

b. the law of cosines



$$b = 3 \text{ cm} \quad c = 7 \text{ cm} \\ \alpha = 45^\circ \quad \beta = 25^\circ$$

Figure 4.7. Triangle for Example 4.2

Solution:

a. By the law of sines,

$$\frac{a}{\sin \alpha} = \frac{b}{\sin \beta}$$

or

$$\frac{a}{\sin 45^\circ} = \frac{3}{\sin 25^\circ}$$

From (4.43), Page 4–7, $\sin 45^\circ = 0.707$.

The sine of 25° is not among those listed in the formulas given in (4.27) through (4.54); we will find it with Excel. Thus, with `=SIN(25*PI()/180)`, we obtain $\sin 25^\circ = 0.422$. Then,

$$\frac{a}{0.707} = \frac{3}{0.422}$$

or

$$a = \frac{3}{0.422} \times 0.707 = 5.26 \text{ cm}$$

b. By the law of cosines,

$$a^2 = b^2 + c^2 - 2bc \cos \alpha$$

or

$$a^2 = 3^2 + 7^2 - 2 \times 3 \times 7 \cos 45^\circ = 58 - 42 \times 0.707 = 28.306$$

Therefore,

$$a = 5.32$$

The small difference between the answers in (a) and (b) is due to the rounding of numbers.

There are many other trigonometric relations that are not given in this text. These can be found in trigonometry textbooks.

4.7 Inverse Trigonometric Functions

The notation $\cos^{-1}y$ or *arccos* y is used to denote an angle whose cosine is y . Thus, if $y = \cos x$, then $x = \cos^{-1}y$. Similarly, if $w = \sin v$, then $v = \sin^{-1}w$, and if $z = \tan u$, then $u = \tan^{-1}z$. These are called *Inverse Trigonometric Functions*.

Example 4.3

Find the angle θ if $\cos^{-1}0.5 = \theta$.

Solution:

Here, we want to find the angle θ given that its cosine is 0.5. From (4.30), Page 4–7, $\cos 60^\circ = 0.5$. Therefore, $\theta = 60^\circ$.

4.8 Area of Polygons in Terms of Trigonometric Functions

We can use the trigonometric functions to compute areas of triangles, quadrilaterals, and other geometric figures. We present just a few.

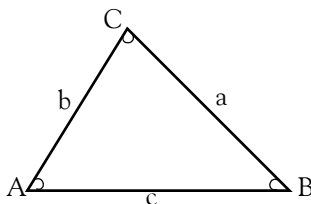


Figure 4.8. General triangle

The area of the general triangle of Figure 4.8 can be found with the formula

$$\text{Area} = \frac{1}{2}ab\sin C = \frac{c^2 \sin A \sin B}{2 \sin C} \quad (4.86)$$

Similarly, formulas for finding the areas of other polygons such as those of Figures 4.9 and 4.10 below in terms of trigonometric functions, are given below. Many others can be found in mathematical tables reference books.

The area of the general quadrilateral of Figure 4.9 can be found with the formula

$$\text{Area} = \frac{1}{2}ef\sin\theta = \frac{1}{4}(b^2 + d^2 - a^2 - c^2)\tan\theta \quad (4.87)$$

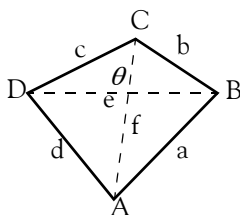


Figure 4.9. General quadrilateral

The area of the parallelogram of Figure 4.10 can be found with the formula

$$\text{Area} = ab\sin A = ab\sin B \quad (4.88)$$

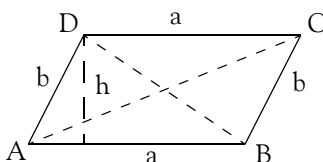


Figure 4.10. Parallelogram

Besides using Excel to find the values of trigonometric functions, one can also use MATLAB. Some of the MATLAB trigonometric functions are listed in Table 4.3 below.

TABLE 4.3 Common MATLAB Trigonometric Functions

acos(x)	Inverse cosine
angle(x)	Four quadrant angle
asin(x)	Inverse sine
atan(x)	Inverse tangent
cos(x)	Cosine
sin(x)	Sine
tan(x)	Tangent

4.9 Summary

- *Trigonometry* is the branch of mathematics that is concerned with the relationships between the sides and the angles of triangles.
- An *angle* is an angular unit of measure with *vertex* (the point at which the sides of an angle intersect) at the center of a circle, and with sides that subtend (cut off) part of the circumference.
- If the subtended arc is equal to $1/360$ of the circumference, the angular unit is one *degree*. Each degree is subdivided into 60 equal parts called *minutes*, and each minute is subdivided into 60 equal parts called *seconds*.
- If the subtended arc is equal to one-fourth of the total circumference, the angular unit is a *right angle*. If the arc equals half the circumference, the unit is a *straight angle* (an angle of 180°).
- An angle between 0° and less than 90° is called an *acute* angle, while an angle between 90° and less than 180° is called an *obtuse* angle.
- The *radian* is a circular angle subtended by an arc equal in length to the radius of the circle whose radius is r units in length.
- The circumference of a circle is $2\pi r$ units; therefore, there are 2π or $6.283\dots$ radians in 360° . Also, $\pi = 3.14159\dots$ radians corresponds to 180° .
- The trigonometric functions vary with the size of an angle. They are defined in terms of one of the two acute angles in a right triangle. The six commonly used trigonometric functions are the sine, cosine, tangent, cotangent, secant, and cosecant.
- The numerical values of the trigonometric functions can be found by using formulas, called *trigonometric identities*.
- A quadrant is defined as any of the four areas into which a plane is divided by the reference set of axes in a Cartesian coordinate system, designated *first*, *second*, *third*, and *fourth*, counting counterclockwise from the area in which both coordinates are positive.
- We can use trigonometric reduction formulas to determine which trigonometric function are positive or negative in each of the four quadrants.
- The sine and the cosine functions can only vary from 0 to ± 1 inclusive.
- We can find the numerical values of trigonometric functions from trigonometric reduction formulas, angle-sum, angle-difference relations, fundamental trigonometric identities, and function-product relations.
- We can find the lengths of the sides and angles of any triangle by the use of the law of sines, law of cosines, and law of tangents.

- The Inverse Trigonometric Functions allow us to find angles when the trigonometric functions are known.
- We can use the trigonometric functions to compute areas of triangles, quadrilaterals, and other geometric figures.
- We can verify the results of our computations with the Excel and/or MATLAB trigonometric functions.

Chapter 4 Fundamentals of Plane Trigonometry

4.10 Exercises

1. If $\theta = 45^\circ$ and $\phi = 30^\circ$, compute:

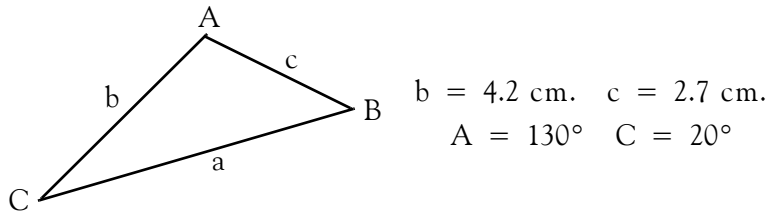
- a. $\cos 15^\circ$ b. $\sin 75^\circ$ c. $\tan 75^\circ$

Verify your answers with Excel.

2. Find the angle θ if $\tan^{-1} 1 = \theta$. Verify your answer with the Excel **ATAN** function.

3. For the triangle below, find the length of side a using:

- a. the law of sines
b. the law of cosines



4.11 Solutions to End-of-Chapter Exercises

1.

- a. From (4.28), $\cos 30^\circ = \sqrt{3}/2$ and from (4.29) $\cos 45^\circ = \sqrt{2}/2$. Also, from (4.42) $\sin 30^\circ = 1/2$ and from (4.43) $\sin 45^\circ = \sqrt{2}/2$. With these relations and (4.64) we obtain

$$\begin{aligned}\cos(\theta - \phi) &= \cos\theta \cos\phi + \sin\theta \sin\phi \\ \cos(45^\circ - 30^\circ) &= \cos 45^\circ \cdot \cos 30^\circ + \sin 45^\circ \cdot \sin 30^\circ \\ \cos(15^\circ) &= \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{3}}{2} + \frac{\sqrt{2}}{2} \cdot \frac{1}{2} = \frac{\sqrt{6}}{4} + \frac{\sqrt{2}}{4} = \frac{1}{4} \cdot (\sqrt{6} + \sqrt{2}) \\ &= \frac{1}{4} \cdot (2.4495 + 1.4142) = \frac{1}{4} \cdot (3.8637) = 0.9659\end{aligned}$$

Check with Excel: =COS(15*PI()/180) returns 0.9659

- b. From (4.28), $\cos 30^\circ = \sqrt{3}/2$ and from (4.29) $\cos 45^\circ = \sqrt{2}/2$. Also, from (4.42) $\sin 30^\circ = 1/2$ and from (4.43) $\sin 45^\circ = \sqrt{2}/2$. With these relations and (4.65) we obtain

$$\begin{aligned}\sin(\theta + \phi) &= \sin\theta \cos\phi + \cos\theta \sin\phi \\ \sin(45^\circ + 30^\circ) &= \sin 45^\circ \cdot \cos 30^\circ + \cos 45^\circ \cdot \sin 30^\circ \\ \cos(15^\circ) &= \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{3}}{2} + \frac{\sqrt{2}}{2} \cdot \frac{1}{2} = \frac{\sqrt{6}}{4} + \frac{\sqrt{2}}{4} = \frac{1}{4} \cdot (\sqrt{6} + \sqrt{2}) \\ &= \frac{1}{4} \cdot (2.4495 + 1.4142) = \frac{1}{4} \cdot (3.8637) = 0.9659\end{aligned}$$

Check with Excel: =SIN(75*PI()/180) returns 0.9659. This is the same value as in part (a) and it is consistent with relation (4.59), i.e., $\sin(90^\circ + \theta) = \cos\theta$ which can be written as $\sin(90^\circ - \theta) = \cos(-\theta) = \cos\theta$. Thus, with $\theta = 15^\circ$, $\sin(90^\circ - 15^\circ) = \sin 75^\circ = \cos 15^\circ$

- c. We use (4.23), that is,

$$\tan\theta = \frac{\sin\theta}{\cos\theta}$$

Then,

$$\tan 75^\circ = \frac{\sin 75^\circ}{\cos 75^\circ}$$

We know the value of $\sin 75^\circ$ from part (b). To find the value of $\cos 75^\circ$ we make use of $\cos 30^\circ = \sqrt{3}/2$, $\cos 45^\circ = \sqrt{2}/2$, $\sin 30^\circ = 1/2$, and $\sin 45^\circ = \sqrt{2}/2$. With these relations and (4.63) we obtain

Chapter 4 Fundamentals of Plane Trigonometry

$$\begin{aligned}\cos(\theta + \phi) &= \cos\theta \cos\phi - \sin\theta \sin\phi \\ \cos(45^\circ + 30^\circ) &= \cos 45^\circ \cdot \cos 30^\circ - \sin 45^\circ \cdot \sin 30^\circ \\ \cos(75^\circ) &= \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{3}}{2} - \frac{\sqrt{2}}{2} \cdot \frac{1}{2} = \frac{\sqrt{6}}{4} + \frac{\sqrt{2}}{4} = \frac{1}{4} \cdot (\sqrt{6} - \sqrt{2}) \\ &= \frac{1}{4} \cdot (2.4495 - 1.4142) = \frac{1}{4} \cdot (1.0353) = 0.2588\end{aligned}$$

and

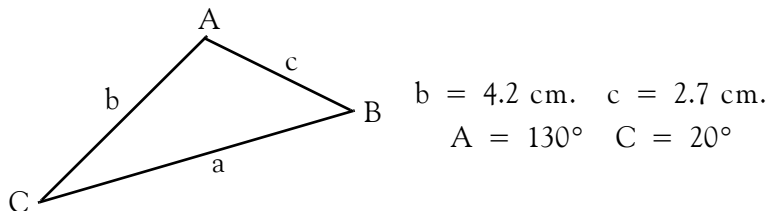
$$\tan 75^\circ = \frac{\sin 75^\circ}{\cos 75^\circ} = \frac{0.9659}{0.2588} = 3.7322$$

Check with Excel: =TAN(75*PI()/180) returns 3.7321

2. We are seeking the value of the angle θ such that $\tan\theta = 1$. Since $\tan\theta = (\sin\theta)/(\cos\theta)$, the ratio $\sin\theta/\cos\theta$ will be equal to 1 when $\sin\theta$ and $\cos\theta$ have the same value. This occurs when $\theta = 45^\circ$ since $\sin 45^\circ = \cos 45^\circ = \sqrt{2}/2$. Therefore, $\tan^{-1}1 = \theta$ implies that $\theta = 45^\circ$.

Check with Excel: =ATAN(1)*180/PI() returns 45

3.



a.

$$\frac{a}{\sin A} = \frac{c}{\sin C}$$

or

$$\frac{a}{\sin 130^\circ} = \frac{2.7}{\sin 20^\circ}$$

We will find $\sin 130^\circ$ and $\sin 20^\circ$ with Excel. Thus, with =SIN(130*PI()/180) we obtain $\sin 130^\circ = 0.766$, and with =SIN(20*PI()/180) we obtain $\sin 20^\circ = 0.342$. Then,

$$\frac{a}{0.766} = \frac{2.7}{0.342}$$

or

$$a = \frac{2.7}{0.342} \times 0.766 = 6.05 \text{ cm}$$

b.

$$a^2 = b^2 + c^2 - 2bc \cos A$$

$$a^2 = 4.2^2 + 2.7^2 - 2 \times 4.2 \times 2.7 \times \cos 130^\circ = 58 - 42 \times 0.707 = 28.306$$

and with Excel =COS(130*PI()/180) we find that $\cos 130^\circ = -0.643$

Therefore,

$$a^2 = 17.64 + 7.29 + 14.58 = 39.51$$

$$a = 6.28$$

Chapter 5

Fundamentals of Calculus

This chapter introduces the basics of differential and integral calculus. It is intended for readers who need an introduction or an accelerated review of this topic. Readers with a strong mathematical background may skip this chapter. Others may find it useful, as well as a convenient source for review.

5.1 Introduction

Calculus is the branch of mathematics that is concerned with concepts such as the rate of change, the slope of a curve at a particular point, and the calculation of an area bounded by curves. Many principles governing physical processes are formulated in terms of rates of change. Calculus is also widely used in the study of statistics and probability.

The fundamental concept of calculus is the theory of limits of functions. A *function* is a defined relationship between two or more variables. One variable, called *dependent variable*, approaches a limit as another variable, called *independent variable*, approaches a number or becomes infinite. In calculus, we are interested in related variables. For instance, the amount of postage required to mail a package is related to its weight. Here, the weight of the package is considered the independent variable, and the amount of postage is considered the dependent variable.

The two branches into which elementary calculus is usually divided, are the *differential calculus*, based on the limits of ratios, and the *integral calculus*, based on the limits of sums.

5.2 Differential Calculus

Suppose that the dependent variable, denoted as y , is a function of the independent variable, denoted as x . This relationship is written as $y = f(x)$. If the variable x changes by a particular amount h , the variable y will also change by a predictable amount k . The ratio of the two amounts of change k/h is called a *difference quotient*. If the rate of change differs over time, this quotient indicates the *average rate of change* of $y = f(x)$ in a particular amount of time.

If the ratio k/h has a limit as h approaches 0, this limit is called the *derivative* of y . The derivative of y may be interpreted as the *slope* of the curve graphed by the equation $y = f(x)$, measured at a particular point. It may also be interpreted as the instantaneous rate of change of y . The process of finding a derivative is called *differentiation*.

If the derivative of y is found for all applicable values of x , a new function is obtained. If $y = f(x)$, the new function, which we call *the first derivative*, is written as y' , or dy/dx , or $df(x)/dx$. If the first derivative y' is itself a function of x , its derivative can also be found; this is called the *second*

Chapter 5 Fundamentals of Calculus

derivative of y and it is written as y'' or d^2y/dx^2 or $d^2f(x)/dx^2$. Third and higher derivatives also exist and they are written similarly.

Example 5.1

Suppose that the distance between City A and City B is 500 miles. Distance is usually denoted with the letter s , so for this example $s = 500$ miles. This distance is fixed; it does not change with time; therefore, we say that distance is independent of time because whether we decide to drive from City A to City B today or tomorrow, we need to drive a distance of 500 miles. What is more important to us, is the time we need to spend driving. This depends on the *velocity* (speed)* which, in turn, depends on traffic conditions, and our driving habits. Thus, if we start driving from City A towards City B at a constant speed of 50 miles per hour, we will spend ten hours driving. As we drive, the distance changes, that is, it is being reduced. For instance, after driving for one hour, we have covered 50 miles and the remaining distance has been reduced 450 miles. Likewise, if our speed during the second hour is 65 miles per hour, the remaining distance is further reduced to 385 miles.

We denote velocity with the letter v and time with t . Then,

$$v = s/t \quad (5.1)$$

that is, velocity is the distance divided per unit of time; in this case, miles divided by hours.

In reality, it is impossible to maintain exactly the same speed per unit of time, that is, when we say that during the first hour we drove at a velocity of 50 miles per hour, this is an *average velocity* since the velocity may vary from 49 to 51, or from 48 to 52 miles per hour. Therefore, the formula of (5.1) is valid for average velocities only.

The *instantaneous velocity*, that is, at any instant, the velocity is expressed as

$$v(t) = \frac{ds}{dt} \quad (5.2)$$

* Velocity is speed with direction assigned to it.

† The development of this expression follows. Let s_1 be the distance at time t_1 and s_2 the distance at time t_2 . Next, let the difference between s_2 and s_1 be denoted with the Greek letter Δ (delta), that is, $\Delta s = s_2 - s_1$ and likewise, let the difference between t_2 and t_1 be denoted as $\Delta t = t_2 - t_1$. Then, $v(t) = \frac{s_2 - s_1}{t_2 - t_1} = \frac{\Delta s}{\Delta t}$. Now, let $\Delta t \rightarrow 0$; this notation says that we let Δt become so small that it approaches zero in the limit but never becomes exactly zero. Obviously, as Δt becomes smaller and smaller, so does Δs . Then, $v(t) = \lim_{\Delta t \rightarrow 0} \frac{\Delta s}{\Delta t} = \frac{ds}{dt}$

where the notation $v(t)$ is used to indicate that the velocity is a function of time, ds denotes a very small change in distance, and dt is a very small change in time. In other words, the velocity at any instance, is the *change* of distance caused by a *change* in time. Mathematically, (5.2) states that instantaneous velocity is the *first derivative* of the distance with respect to time, or that the instantaneous velocity is found by *differentiating* the distance with respect to time.

Instantaneous acceleration, denoted with the letter a , is defined as the first derivative of velocity with respect to time, that is,

$$a(t) = \frac{dv}{dt} \tag{5.3}$$

and since the instantaneous velocity is the first derivative of the distance with respect to time, we say that the instantaneous acceleration is the *second derivative* of the distance with respect to time. Stated mathematically,

$$a(t) = \frac{d^2s}{dt^2} \tag{5.4}$$

5.3 The Derivative of a Function

Let us consider the curve of Figure 5.1 where $y = f(x)$ meaning that y is a function of x . In other words, the values of y depend on the values of x .

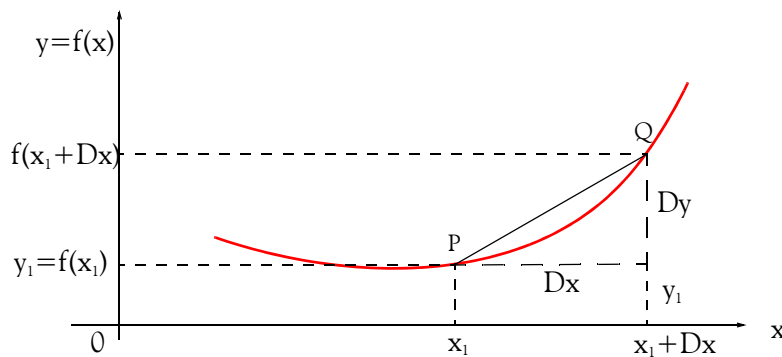


Figure 5.1. Definition of the derivative of a function.

In Figure 5.1, point $P(x_1, y_1)$ is considered to be fixed; that is, x_1 and y_1 are held constant. Let $Q(x_1 + \Delta x, y_1 + \Delta y)$ be another point close to it; then,

$$y_1 + \Delta y = f(x_1 + \Delta x) \tag{5.5}$$

Subtraction of $y_1 = f(x_1)$ from (5.5) yields

Chapter 5 Fundamentals of Calculus

$$\Delta y = f(x_1 + \Delta x) - f(x_1) \quad (5.6)$$

We denote the slope of the straight line PQ as m ; then,

$$m = \frac{\Delta y}{\Delta x} = \frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x} \quad (5.7)$$

Now, suppose we keep x_1 in its fixed position and we let Δx become smaller and smaller approaching zero but never becoming exactly zero. In this case, the slope m approaches a constant value which we call it its *limit*. At this limit, the slope becomes tangent to the curve at point P; we denote this slope as m_{tan} . Stated mathematically,

$$m_{tan} = \lim_{Q \rightarrow P} m = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x} \quad (5.8)$$

In the above discussion, we assumed that the limit as $\Delta x \rightarrow 0$ exists. This is not always the case; this limit may or may not exist. If it does exist, the function is said to have a *derivative*, or to be *differentiable* at x_1 . Therefore, in differential calculus we are concerned with two general problems:

- I. Given a function $f(x)$, find those values of the independent variable x at which $f(x)$ has a derivative.
- II. If $f(x)$ has a derivative, find it.

We express the derivative of (5.8) as

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \quad (5.9)$$

where x is held fixed, and Δx varies approaching zero, but never becoming exactly zero; otherwise the right side of (5.9) would be reduced to the *indeterminate** form $0/0$.

Note 5.1

Occasionally, we may need to use certain algebraic identities. We list a few below; others will be given as needed.

$$(a + b)^2 = a^2 + 2ab + b^2 \quad (5.10)$$

$$(a - b)^2 = a^2 - 2ab + b^2 \quad (5.11)$$

* Indeterminate forms are those that do not lead up to a definite result or ending. Other indeterminate forms are $\infty - \infty$, ∞^0 , 1^∞ , ∞/∞ , and 0^0 .

$$(a^2 - b^2) = (a + b) \cdot (a - b) \quad (5.12)$$

$$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3 \quad (5.13)$$

$$(a - b)^3 = a^3 - 3a^2b + 3ab^2 - b^3 \quad (5.14)$$

Example 5.2

The curve of Figure 5.2 is described as

$$y = f(x) = x^3 - 3x^2 + 4x + 2 \quad (5.15)$$

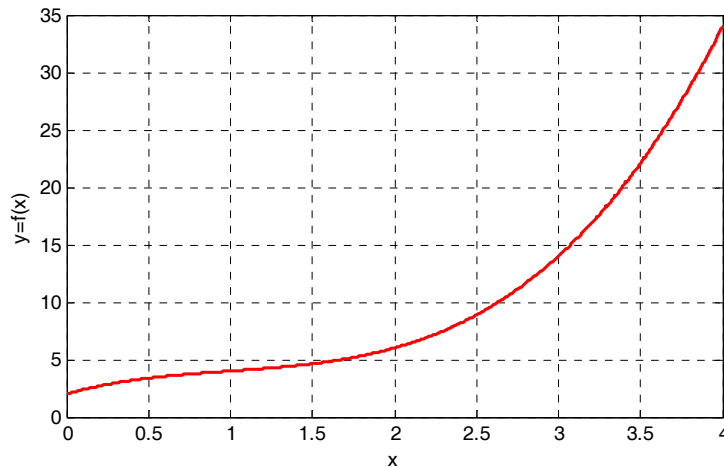


Figure 5.2. Curve for Example 5.2

Compute the derivative of this function using relation (5.9).

Solution:

As a first step we compute $f(x + \Delta x) - f(x)$. Thus, in (5.15), we replace x with $x + \Delta x$; then we subtract $f(x)$ from it as shown below.

$$\begin{aligned} f(x + \Delta x) - f(x) &= (x + \Delta x)^3 - 3(x + \Delta x)^2 + 4(x + \Delta x) + 2 - (x^3 - 3x^2 + 4x + 2) \\ &= x^3 + 3x^2\Delta x + 3x\Delta x^2 + \Delta x^3 - 3(x^2 + 2x\Delta x + \Delta x^2) + 4x + 4\Delta x + 2 - (x^3 - 3x^2 + 4x + 2) \\ &= x^3 - x^3 - 3x^2 + 3x^2 + 4x - 4x + 2 - 2 + 3x^2\Delta x + 3x\Delta x^2 + \Delta x^3 - 6x\Delta x - 3\Delta x^2 + 4\Delta x \\ &= 3x^2\Delta x + 3x\Delta x^2 + \Delta x^3 - 6x\Delta x - 3\Delta x^2 + 4\Delta x \\ &= \Delta x(3x^2 + 3x\Delta x + \Delta x^2 - 6x - 3\Delta x + 4) \end{aligned}$$

By substitution of the last line of the above expression into (5.9), we obtain

Chapter 5 Fundamentals of Calculus

$$\begin{aligned}\frac{dy}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{\Delta x(3x^2 + 3x\Delta x + \Delta x^2 - 6x - 3\Delta x + 4)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} (3x^2 + 3x\Delta x + \Delta x^2 - 6x - 3\Delta x + 4) = 3x^2 - 6x + 4\end{aligned}\tag{5.16}$$

It is important to remember that the division by Δx on the right side term in the second line of (5.16), *must* be performed before we let $\Delta x \rightarrow 0$.

In Example 5.2, we found that the computation of the derivative of a polynomial* function, although not difficult, it involves a tedious process. Fortunately, we can use some formulas that will enable us to compute the derivatives of polynomials very easy, and rather quickly. The rules and formulas for polynomials and other functions of interest are given below without proof. The proofs can be found in calculus textbooks.

1. The derivative[†] of a constant is zero. Thus, if c is any constant,

$$\frac{d}{dx}c = 0\tag{5.17}$$

Example 5.3

It is given that $y = f(x) = 4$. Find its derivative dy/dx .

Solution:

Here, y is a constant; its value does not depend on the independent variable x . Therefore,

$$\frac{dy}{dx} = 0$$

* A polynomial is an algebraic expression consisting of one or more summed terms, each term consisting of a constant multiplier and one or more variables raised to integral powers, that is, integer exponents. For example, $x^2 - 5x + 6$ and $2p^3q + y$ are polynomials.

† Henceforth, the terminology “derivative” will mean the first derivative, that is, if $y = f(x)$, its derivative is $\frac{d}{dx}f(x)$. For higher order derivatives we will use the terminology “second derivative”, “third derivative” and so on.

2. If

$$y = f(x) = cx^n \quad (5.18)$$

where c is any constant, and n is any non-zero positive integer, the derivative of y with respect to x is

$$\frac{dy}{dx} = \frac{d}{dx}cx^n = ncx^{n-1} \quad (5.19)$$

Example 5.4

It is given that $y = f(x) = 5x^4$. Find its derivative dy/dx .

Solution:

For this example, (5.19) is applicable. Then,

$$\frac{dy}{dx} = \frac{d}{dx}(5x^4) = 4 \times 5x^{4-1} = 20x^3$$

3. The derivative of the sum of a finite number of differentiable functions is equal to the sum of their derivatives. That is, if

$$y = u_1 + u_2 + u_3 + \dots + u_n \quad (5.20)$$

then,

$$\frac{dy}{dx} = \frac{d(u_1 + u_2 + u_3 + \dots + u_n)}{dx} = \frac{du_1}{dx} + \frac{du_2}{dx} + \frac{du_3}{dx} + \dots + \frac{du_n}{dx} \quad (5.21)$$

Example 5.5

It is given that $y = f(x) = x^3 - 3x^2 + 4x + 2$. Find its derivative dy/dx .

Solution:

For this example, we apply (5.21). Then,

$$\begin{aligned} \frac{dy}{dx} &= \frac{d}{dx}(x^3 - 3x^2 + 4x + 2) = \frac{d}{dx}(x^3) + \frac{d}{dx}(-3x^2) + \frac{d}{dx}(4x) + \frac{d}{dx}(2) \\ &= 3x^2 - 6x + 4 + 0 = 3x^2 - 6x + 4 \end{aligned}$$

Chapter 5 Fundamentals of Calculus

This is the same answer as in Example 5.2, which was found by direct application of the definition of the derivative.

By our earlier discussion on velocity and acceleration, if $s = f(t)$ represents the position of a moving body at time t , then the first derivative $\frac{ds}{dt}$ yields the velocity v , and the second derivative $\frac{d^2s}{dt^2}$ yields the acceleration a .

Example 5.6

A salesman drives his automobile along a highway where the average speed is approximately 10 miles per hour due to bad weather and heavy traffic. From past experience, he knows that the distance as a function of time is

$$s(t) = t^3 - 6t^2 - 2t + 50 \quad (5.22)$$

for $t > 0$, that is, for positive time, where t is in hours and s is in miles. After driving for some time, the salesman, through a highway sign, is informed that one of the lanes ahead is closed to traffic. Frustrated over this situation, he stops, makes a U-turn, and proceeds in the opposite direction to return to his office.

Find the distance he has traveled before reversing direction and the acceleration at that point.

Solution:

We will use Excel to plot the distance s versus time t , so that we can see how the distance decreases as time increases.

We start with a blank spreadsheet. In Column A of the spreadsheet shown in Figure 5.3, we enter the time t in increments of 0.020 hours. Thus, in A2 we type 0.000 and in A3 we type 0.020. Then, using the *autofill** feature, we fill-in A4:A252, where A252 contains the value 5.000. We use Column B for the distance s . In B2 we type the formula `=A2^3-8*A2^2-2*A2+50`; this represents the expression of (5.22). Then, we copy B2 down to the range B3:B252. Next, using the *Chart Wizard* we obtain the plot shown on Figure 5.3. For brevity, only a partial list of the values of time and distance are shown.

* To use this feature, we highlight cells A2 and A3. We observe that on the lower right corner of A3, there is a small black square; this is called the fill handle. If it does not appear on the spreadsheet, we can make it visible by performing the sequential steps *Tools*>*Options*, select the *Edit* tab, and place a check mark on the *Drag and Drop* setting. Next, we point the mouse to the fill handle and we observe that the mouse pointer appears as a small cross. Then, we click, hold down the mouse button, we drag it to A252 and release the mouse button. We observe that, as we drag the fill handle, a pop-up note shows the cell entry for the last value in the range.

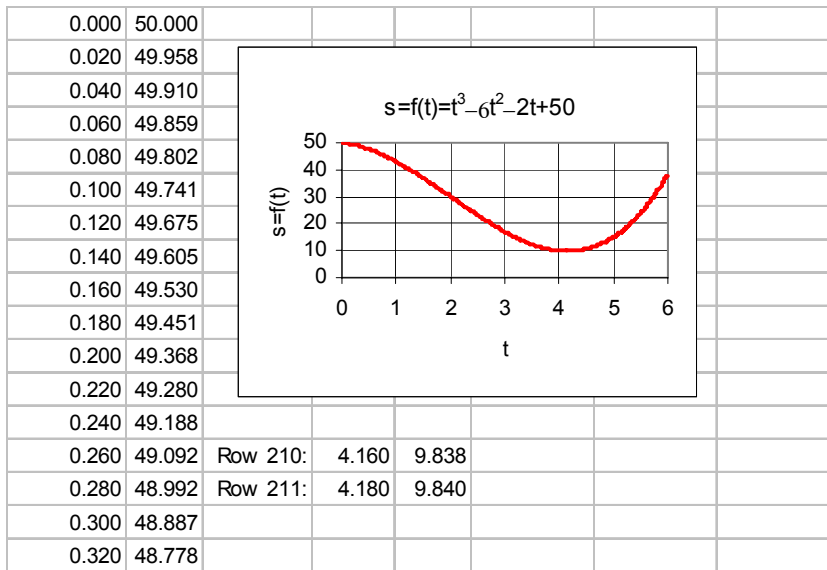


Figure 5.3. Plot for the function of Example 5.6

Rows 210 and 211 indicate a change from decreasing to increasing values in distance s . That is, as time increases, the distance is increasing instead of decreasing. This is also shown on the plot, but we can only obtain an approximation from the plot which shows that the salesman intended to drive a distance of 50 miles, but after driving for approximately $50 - 10 = 40$ miles, he reverses direction.

To find the exact distance that the salesman traveled before reversing direction, we make use of the velocity v and acceleration a relations which, as we saw earlier, are the first and second derivatives of the distance s respectively.

Using our intuition, we conclude that his velocity is zero when he stops to reverse direction, and to reach zero velocity he decelerates*.

Differentiating (5.22) once to find the velocity v , and then, one more time to obtain the acceleration a , we obtain

$$v = \frac{ds}{dt} = 3t^2 - 12t - 2 \tag{5.23}$$

and

$$a = \frac{dv}{dt} = \frac{d^2s}{dt^2} = 6t - 12 \tag{5.24}$$

* Deceleration is negative acceleration, that is, decrease in the rate of change of velocity.

Chapter 5 Fundamentals of Calculus

The velocity v is zero when (5.23) equals zero, that is, when

$$3t^2 - 12t - 2 = 0 \quad (5.25)$$

Using the quadratic formula we obtain

$$t_1 = \frac{-(-12) + \sqrt{12^2 - 4 \times 3 \times (-2)}}{2 \times 3} = \frac{12 + \sqrt{144 + 24}}{6} = \frac{12 + \sqrt{168}}{6} = 4.16 \text{ hours} \quad (5.26)$$

and

$$t_2 = \frac{-(-12) - \sqrt{12^2 - 4 \times 3 \times (-2)}}{2 \times 3} = \frac{12 - \sqrt{144 + 24}}{6} = \frac{12 - \sqrt{168}}{6} = -0.16 \text{ hours} \quad (5.27)$$

Since we are only interested in positive time, that is, values of t greater than zero, we ignore the value of t_2 in (5.27).

The distance traveled before reversing direction is found by substitution of t_1 into (5.22). Thus,

$$\begin{aligned} s &= t^3 - 6t^2 - 2t + 50 = 4.173^3 - 6 \times 4.173^2 - 2 \times 4.173 + 50 \\ &= 72.668 - 104.484 - 8.346 + 50 = 9.838 \end{aligned} \quad (5.28)$$

This result indicates that, after driving for 4.16 hours, the distance s has been reduced from 50 to 9.838 miles. Thus, the distance traveled before reversing direction is $50.000 - 9.838 = 40.162$ miles.

The acceleration (deceleration in this example) a , is found by substitution of t_1 into (5.24). Thus,

$$a = 6t - 12 = 6 \times 4.173 - 12 = 13.038 \frac{\text{miles}}{\text{hr}^2} \quad (5.29)$$

It is important to remember that, whereas the unit of velocity is *miles per hour*, the unit of acceleration is *miles per hour²*.

Derivatives are also very important in business applications. To illustrate, suppose that a company manufactures x units of a certain product per month. The total cost y (in dollars) for the production of those units, is a function of the units produced, that is, $y = f(x)$, and normally, it includes a percentage of the cost of the manufacturing facilities, maintenance, labor, taxes, etc.

Suppose that the cost of producing $x + \Delta x$ units is $y + \Delta y$ dollars. The ratio $\Delta y / \Delta x$ represents the increase in cost per unit increase in output. The limit of this ratio as $\Delta x \rightarrow 0$, is called the *marginal cost*. Stated in other words, if the total monthly cost is y for x units per month, the marginal cost

is the derivative dy/dx which indicates the rate of increase of cost y per unit increase in production from the reference quantity x . The company is also interested in revenue and profits. A discussion on maxima and minima will help us to understand how the company can adjust its production to achieve maximum profit. We will reconsider this case in Example 5.8, after our discussion on maxima and minima in Section 5.4 below.

5.4 Maxima and Minima

Differential calculus is a very powerful tool for solving problems in which we are interested in finding maximum and minimum values. An easy way to determine the approximate values of the maxima or minima of a function, is to plot the function $y = f(x)$ using Excel (or MATLAB) as shown in Example 5.7 that follows.

Example 5.7

Use Excel to plot the function

$$y = f(x) = \frac{1}{3}x^3 - 2x^2 + 3x + 2 \tag{5.30}$$

and give approximate maximum and minimum values.

Solution:

Following the same procedure as in Example 5.6, we construct the plot shown in Figure 5.4.

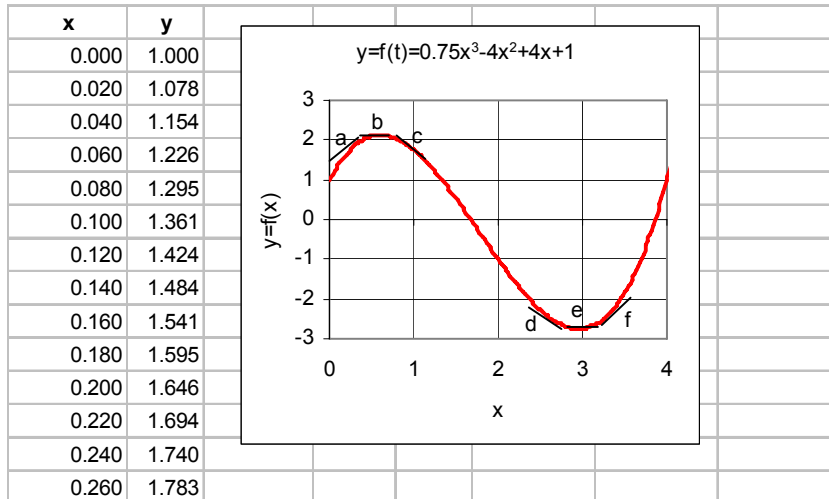


Figure 5.4. Plot for Example 5.7

Points a and c are assumed to be very close to point b. Likewise, d and f are very close to e. At points a and f we see that y is an increasing function of x . This means that y increases as x

Chapter 5 Fundamentals of Calculus

increases. At these points the slope of the tangent to the curve dy/dx is said to be positive. At points b and e, we see that the slope of the tangent to the curve dy/dx is said to be zero since, precisely at these points, y neither increases nor decreases. At points c and d, we see that y is a decreasing function of x ; this means that y decreases as x increases. At these points, the slope of the tangent to the curve dy/dx is said to be negative. Obviously, the maximum and minimum values of y occur when the slope dy/dx is zero. From Figure 5.4, we see that a maximum occurs at point b when the transition is from a positive slope, at point a in this case, to a negative slope at point c. Conversely, a minimum occurs at point d when the transition is from a negative slope at point d, to a positive slope at point f.

It is shown in calculus textbooks that we can determine whether a value represents a maximum or a minimum, by performing the following test.

I. If the second derivative $\frac{d^2y}{dx^2}$ is positive when $\frac{dy}{dx} = 0$, y is a minimum.

II. If the second derivative $\frac{d^2y}{dx^2}$ is negative when $\frac{dy}{dx} = 0$, y is a maximum.

III. If the second derivative $\frac{d^2y}{dx^2}$ is zero when $\frac{dy}{dx} = 0$, the test fails.

Of course, it is highly recommended that we always plot the function $y = f(x)$ to determine the approximate values by visual inspection.

We now return to our discussion on the manufacturing company that produces x units of a particular product, with Example 5.8 that follows.

Example 5.8

The manufacturing company can sell x units of a certain product per month at a price $p = 200 - 0.01x$ dollars per unit. The management of the company has determined that it costs $y = 50x + 20000$ dollars to manufacture the x units.

- How many items should the company produce to maximize the profit?
- How much should each item sell for to achieve this profit?

* Usually, at a higher price the company would sell less and at a lower price it would sell more.

Solution:

The profit is revenue minus cost; therefore,

$$\begin{aligned}
 \text{Profit} &= xp - y \\
 &= x(200 - 0.01x) - (50x + 20000) = 200x - 0.01x^2 - 50x - 20000 \\
 &= 150x - 0.01x^2 - 20000
 \end{aligned}
 \tag{5.31}$$

a. To find the value of x which will maximize the profit, we form the first derivative of (5.31) and we set it equal to zero, that is,

$$\frac{d}{dx} \text{Profit} = \frac{d}{dx}(150x - 0.01x^2 - 20000) = 150 - 0.02x = 0
 \tag{5.32}$$

and solving (5.32) for x we obtain $x = 7500$.

We do not know whether this value of x is a maximum or minimum. This can be determined either by taking the second derivative of (5.31), or equivalently the first derivative of (5.32), or by plotting it. The second derivative yields -0.02 and since this is a negative value, we conclude that the value of $x = 7500$ is a maximum, and therefore, it represents the number of units that will maximize the profit. This is confirmed with the plot of Figure 5.5.

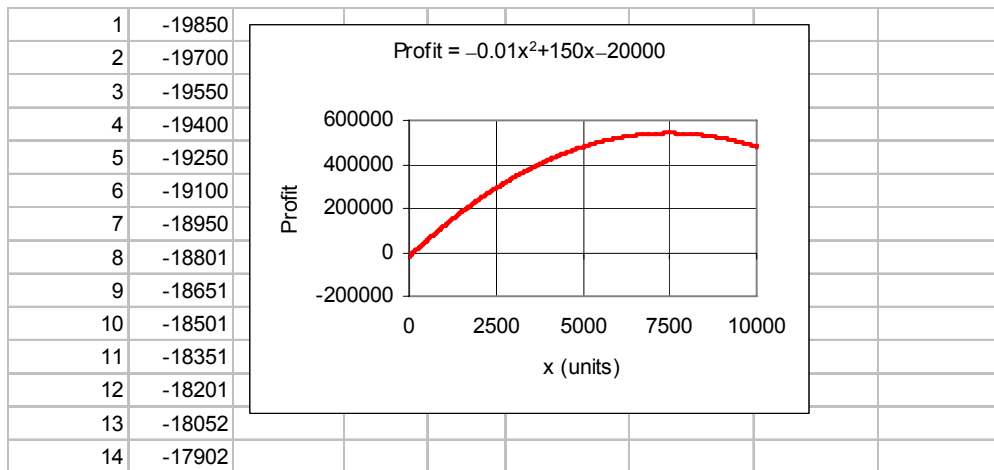


Figure 5.5. Plot for Example 5.8.

b. To achieve the maximum profit each unit must be sold at

$$\text{price} = 200 - 0.01x = 200 - 0.01 \times 7500 = 200 - 75 = \$125 \text{ per unit}$$

We will conclude the section on differential calculus with the formulas below where u , v , and w

Chapter 5 Fundamentals of Calculus

are functions of x , and a , c , and n denote constant values.

$$\frac{d}{dx}(a) = 0 \quad (5.33)$$

$$\frac{d}{dx}(x) = 1 \quad (5.34)$$

$$\frac{d}{dx}(au) = a \frac{du}{dx} \quad (5.35)$$

$$\frac{d}{dx}(u + v - w) = \frac{du}{dx} + \frac{dv}{dx} - \frac{dw}{dx} \quad (5.36)$$

$$\frac{d}{dx}(uv) = u \frac{dv}{dx} + v \frac{du}{dx} \quad (5.37)$$

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2} \quad (5.38)$$

$$\frac{d}{dx}(u^n) = nu^{n-1} \frac{du}{dx} \quad (5.39)$$

$$\frac{d}{dx}(\sqrt{u}) = \frac{1}{2\sqrt{u}} \frac{du}{dx} \quad (5.40)$$

$$\frac{d}{dx}\left(\frac{1}{u}\right) = -\frac{1}{u^2} \frac{du}{dx} \quad (5.41)$$

$$\frac{d}{dx}\left(\frac{1}{u^n}\right) = -\frac{n}{u^{n+1}} \frac{du}{dx} \quad (5.42)$$

$$\frac{d}{dx}(\ln u) = \frac{1}{u} \frac{du}{dx} \quad (5.43)$$

$$\frac{d}{dx}(e^u) = e^u \frac{du}{dx} \quad (5.44)$$

$$\frac{d}{dx}(u^v) = vu^{v-1} \frac{du}{dx} + (\ln u)u^v \frac{dv}{dx} \quad (5.45)$$

$$\frac{d}{dx}(\sin u) = (\cos u) \frac{du}{dx} \quad (5.46)$$

$$\frac{d}{dx}(\cos u) = -(\sin u)\frac{du}{dx} \quad (5.47)$$

$$\frac{d}{dx}(\tan u) = (\sec^2 u)\frac{du}{dx} \quad (5.48)$$

5.5 Integral Calculus

Integral calculus involves the process of finding the function itself when its derivative is known. This is the inverse process to finding the derivative of a function. The process of finding an integral of a function is called *integration*. There are two kinds of integration, *indefinite* and *definite*. We will discuss the indefinite type of integration first.

Suppose we are given the relation

$$\frac{dy}{dx} = 3x^2 \quad (5.49)$$

and we are asked to find y . Relation (5.49) is a simple differential equation, and although we do not know how to solve differential equations, using our knowledge with derivatives we can say that

$$y = x^3 \quad (5.50)$$

because we know that if we take the first derivative of (5.50) we will obtain (5.49). However, we must realize that the derivatives of other functions such as

$$y = x^3 + 2, \quad y = x^3 - 7, \quad y = x^3 + \sqrt{5} \quad (5.51)$$

when differentiated, will also yield (5.49). Therefore, we can say that the solution of the differential equation of (5.49) is, in general,

$$y = x^3 + C \quad (5.52)$$

and the constant C is referred to as an *arbitrary constant of integration*.

Relation (5.49) is called a *simple differential equation*. Differential equations are very useful not only in engineering and in the physical sciences, but also in the business and finance sectors.

Differential equations, and consequently integration, require the ability to guess the answer. Fortunately, there are certain formulas and procedures that we can use to minimize the amount of guesswork. We will present the most common. First, we will introduce the concept of *differentials*.

The differential equation of (5.49) can also be written as

$$dy = 3x^2 dx \quad (5.53)$$

and when expressed in this form, we say that dy is the differential of y in terms of x and dx , and dx is the differential of x .

Note 5.2

The symbol \int denotes integration. The integral $\int f(x)dx$ is an *indefinite integral* and $\int_a^b f(x)dx$ is a *definite integral*.

Note 5.3

When we evaluate indefinite integrals, we *must* always add the constant of integration C .

5.6 Indefinite Integrals

1. The integral of the differential of a function x is x plus an arbitrary constant C . Stated mathematically,

$$\int dx = x + C \quad (5.54)$$

2. A constant, say a , may be written in front of the integral sign. In other words,

$$\int a dx = a \int dx \quad (5.55)$$

3. The integral of the sum of two or more differentials is the sum of their integrals. That is,

$$\int (dx_1 + dx_2 + \dots + dx_n) = \int dx_1 + \int dx_2 + \dots + \int dx_n \quad (5.56)$$

4. If $n \neq -1$, that is, if n is not equal to minus one, the integral of $x^n dx$ is obtained by adding one to the exponent and dividing by the new exponent. In other words,

$$\int x^n dx = \frac{x^{n+1}}{n+1} + C \quad (5.57)$$

We have just introduced the elementary forms of integrals. There are many more; they can be found in mathematical tables listed as *tables of integrals*. Table 5.1 lists the most common differentials and their integral counterparts.

5.7 Definite Integrals

Let us consider the curve $y = f(x)$ shown in Figure 5.6.

TABLE 5.1 Differentials and their integral counterparts

	<i>Differentials</i>	<i>Integrals</i>
1	$du = \frac{du}{dx} dx$	$\int du = u + C$
2	$d(au) = a du$	$\int a dx = a \int dx$
3	$d(x_1 + x_2 + \dots + x_n) = dx_1 + dx_2 + \dots + dx_n$	$\int (dx_1 + dx_2 + \dots + dx_n) = \int dx_1 + \int dx_2 + \dots + \int dx_n$
4	$d(x^n) = nx^{n-1} dx$	$\int x^n dx = \frac{x^{n+1}}{n+1} + C$
5	$d(\ln x) = \frac{dx}{x}$	$\int \frac{dx}{x} = \ln x + C$
6	$d(e^x) = e^x dx$	$\int e^x dx = e^x + C$
7	$d(a^x) = a^x \ln a dx$	$\int a^x dx = \frac{a^x}{\ln a} + C$
8	$d(\sin x) = \cos x dx$	$\int \cos x dx = \sin x + C$
9	$d(\cos x) = -\sin x dx$	$\int \sin x dx = -\cos x + C$
10	$d(\tan x) = \sec^2 x dx$	$\int \sec^2 x dx = \tan x + C$

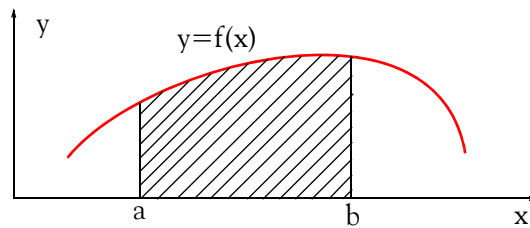


Figure 5.6. Area under $y=f(x)$

The area under the curve $y = f(x)$ between points a and b can be found from the integral

$$\boxed{\int_a^b f(x) dx = F(x) \Big|_a^b = F(b) - F(a)} \quad (5.58)$$

Relation (5.58) is known as *the fundamental theorem of integral calculus*.

Chapter 5 Fundamentals of Calculus

In (5.58), $F(x)$ is the integral of $f(x)dx$, and the meaning of the notation $F(x)|_a^b$ is that we must first replace x with the upper value b , that is, we set $x = b$ to obtain $F(b)$, and from it we subtract the value $F(a)$, which is obtained by setting $x = a$.

The lower value a in (5.58) is called the *lower limit of integration*, and the upper value b is the *upper limit of integration*.

The left side of (5.58) is a *definite integral*. In the evaluation of all definite integrals, the constant of integration C is zero; therefore, it is omitted.

Example 5.9

A curve is described by the equation

$$y = f(x) = 2x^3 - 6x + 3 \quad (5.59)$$

Compute the area under this curve between points $x_1 = 1$ and $x_2 = 3$.

Solution:

Let us use Excel to plot the curve of (5.59). The plot is shown in Figure 5.7.

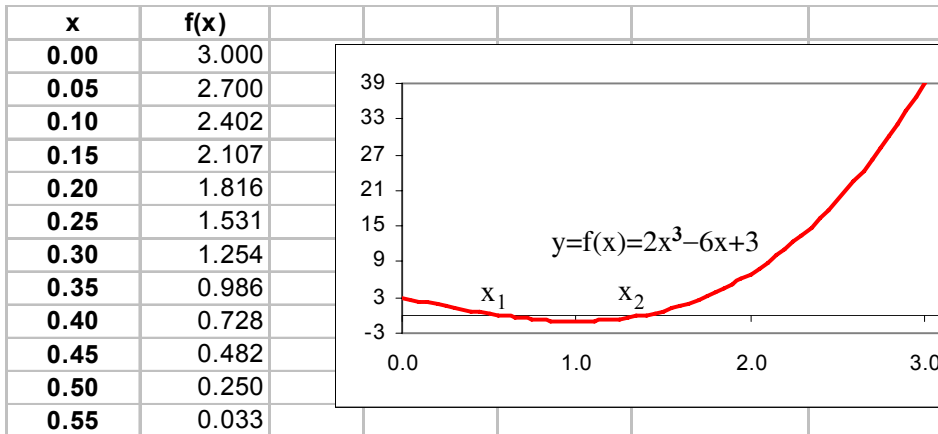


Figure 5.7. Plot for the function of Example 5.7.

To generate the above plot, we entered several values of x in Column A of the spreadsheet, starting with A2, and in B2 we typed the formula $=2*A2^3-6*A2+3$; this was copied down to the same row as the values of x in Column A.

The plot of Figure 5.7 shows that the area from about $x = 0.6$ to about $x = 1.4$ is negative;* it is positive for all other values of $x \geq 0$.

For this example the net area is

$$\begin{aligned} \int_1^3 (2x^3 - 6x + 3)dx &= \left(2\frac{x^4}{4} - 6\frac{x^2}{2} + 3x\right)\Bigg|_1^3 = \left(\frac{1}{2}x^4 - 3x^2 + 3x\right)\Bigg|_1^3 \\ &= \left(\frac{81}{2} - 27 + 9\right) - \left(\frac{1}{2} - 3 + 3\right) = 22.5 - 0.5 = 22 \text{ sq units} \end{aligned}$$

Note 5.4

Since the result of integration in most cases[†] denotes area, we must express the answer in square units such as in^2 , cm^2 and so on.

It is shown in advanced mathematics textbooks that in compound interest computations, where the accrued interest is added to the principal at the end of each period, the amount at the end of a period can be found from the relation

$$\int_{p_1}^{p_2} \frac{1}{p} dp = \int_0^t r dt \tag{5.60}$$

where p_1 is the amount deposited at the beginning of a period, p_2 the amount at the end of the period, r is the interest rate per year, and t is the time, in years, that it takes to increase the amount from p_1 to p_2 .

Example 5.10

Suppose that a bank pays 7.2% interest on deposits of \$10,000 or more, provided that the deposited amount is not withdrawn in less than a year. If \$15,000 is deposited today, in how many years will this amount grow to \$30,000 if no withdrawals are made?

Solution:

* The fact that the curve is negative within an interval of values of x should not be a reason for concern; integration of the given function using (5.58), will yield the net area.

† Another class of integrals, referred to as “line integrals” yield the answer in units other than square units. For instance, in physics, we can use a line integral to find the work or energy in foot–pounds.

Chapter 5 Fundamentals of Calculus

For this example, $p_1 = 15000$, $p_2 = 30000$, $r = 0.072$, and we wish to find t . Then,

$$\int_{15000}^{30000} \frac{1}{p} dp = \int_0^t 0.072 dt \quad (5.61)$$

Integrating both sides of (5.61) and rearranging, we obtain

$$0.072t = \ln(30000 - 15000) = \ln\left(\frac{30000}{15000}\right) = \ln 2$$

Solving for t , we obtain

$$t = \frac{\ln 2}{0.072} = 9.627 \text{ years}$$

Therefore, one must leave his money in a bank or savings and loans institution for approximately 10 years for the original deposit to double if the interest rate is 7.2%.

The equation

$$a \frac{d^2 y}{dx^2} + b \frac{dy}{dx} + cy = 0 \quad (5.62)$$

is called a *second order differential equation*.

Suppose, for example, that an object weighing 10 lb. is hung on a helical spring, and causes the spring to stretch 1 inch. The object is then pulled down 2 more inches and released. The resulting motion is referred to as *simple harmonic motion* and it is described by a second order differential equation such as (5.62).

Of course, higher order differential equations exist; these are discussed in differential equations textbooks.

5.8 Summary

- *Calculus* is the branch of mathematics that is concerned with concepts such as the rate of change, the slope of a curve at a particular point, and the calculation of an area bounded by curves.
- The fundamental concept of calculus is the theory of limits of functions.
- A *function* is a defined relationship between two or more variables. One variable, called *dependent variable*, approaches a limit as another variable, called *independent variable*, approaches a number or becomes infinite.
- The two branches into which elementary calculus is usually divided, are the *differential calculus*, based on the limits of ratios, and the *integral calculus*, based on the limits of sums.
- If the limit as $\Delta x \rightarrow 0$ exists the function is said to have a *derivative*, or to be *differentiable* at x_1 and the derivative is defined as

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

- The derivative of a constant is zero.

- If

$$y = f(x) = cx^n$$

where c is any constant, and n is any non-zero positive integer, the derivative of y with respect to x is

$$\frac{dy}{dx} = \frac{d}{dx} cx^n = ncx^{n-1}$$

- The derivative of the sum of a finite number of differentiable functions is equal to the sum of their derivatives. That is, if

$$y = u_1 + u_2 + u_3 + \dots + u_n$$

then,

$$\frac{dy}{dx} = \frac{d(u_1 + u_2 + u_3 + \dots + u_n)}{dx} = \frac{du_1}{dx} + \frac{du_2}{dx} + \frac{du_3}{dx} + \dots + \frac{du_n}{dx}$$

- If the second derivative d^2y/dx^2 is positive when $dy/dx = 0$, y is a minimum.
- If the second derivative d^2y/dx^2 is negative when $dy/dx = 0$, y is a maximum.
- Integral calculus involves the process of finding the function itself when its derivative is known.

Chapter 5 Fundamentals of Calculus

- There are two kinds of integrals, indefinite and definite.
- When evaluating indefinite integrals, we add the constant C is referred to as an *arbitrary constant of integration*.

- The integral of the differential of a function x is x plus an arbitrary constant C . That is,

$$\int dx = x + C$$

- A constant may be written in front of the integral sign. In other words,

$$\int a dx = a \int dx$$

- The integral of the sum of two or more differentials is the sum of their integrals. That is,

$$\int (dx_1 + dx_2 + \dots + dx_n) = \int dx_1 + \int dx_2 + \dots + \int dx_n$$

- If $n \neq -1$, that is, if n is not equal to minus one, the integral of $x^n dx$ is obtained by adding one to the exponent and dividing by the new exponent. In other words,

$$\int x^n dx = \frac{x^{n+1}}{n+1} + C$$

- The fundamental theorem of integral calculus states that

$$\int_a^b f(x) dx = F(x) \Big|_a^b = F(b) - F(a)$$

- In definite integrals the lower value is called the lower limit of integration, and the upper value b is the upper limit of integration.
- Since the result of integration in most cases denotes area, we must express the answer in square units.
- Higher order differential equations exist; these are discussed in differential equations textbooks.

5.9 Exercises

1. Find the derivatives of the following functions.

$$\text{a. } y = 2x^6 - 3x^2 \quad \text{b. } y = \frac{1}{x} \quad \text{c. } y = e^{-3x} \quad \text{d. } y = \sqrt{4x}$$

2. Evaluate the following indefinite integrals.

$$\text{a. } \int 3x^2 dx \quad \text{b. } \int (5x^3 + 6x^2 - 12) dx \quad \text{c. } \int 3e^{2x} dx$$

3. Evaluate the following definite integrals, that is, find the area under the curve for the lower and upper limits of integration.

$$\text{a. } \int_1^2 (2x^5 - 6x^3 + 3x) dx \quad \text{b. } \int_3^5 (3x^2 + 6x - 8) dx$$

4. Suppose that a bank pays 6% interest on deposits of \$5,000 or more provided that the deposited amount is not withdrawn in less than a year. If \$8,000 is deposited today, in how many years will this amount grow to \$16,000?

5.10 Solutions to End-of-Chapter Exercises

1.

a.
$$y = 2x^6 - 3x^2$$

$$\frac{dy}{dx} = 6 \times 2 \times x^{6-1} - 2 \times 3 \times x^{2-1} = 12x^5 - 6x$$

Check with the MATLAB **diff(f)** function:

```
syms x y % define symbolic variables x and y
```

```
y = 2*x^6-3*x^2; diff(y)
```

```
ans =
```

```
12*x^5-6*x
```

b.
$$y = 1/x = x^{-1}$$

$$\frac{dy}{dx} = -1 \times x^{-1-1} = -x^{-2} = -1/x^2$$

Check with the MATLAB **diff(f)** function:

```
syms x y % define symbolic variables x and y
```

```
y = 1/x; diff(y)
```

```
ans =
```

```
-1/x^2
```

c.
$$y = e^{-3x}$$

$$\frac{dy}{dx} = -3 \times x^{-3x} = -3e^{-3x}$$

Check with the MATLAB **diff(f)** function:

```
syms x y % define symbolic variables x and y
```

```
y=exp(-3*x); diff(y)
```

```
ans =
```

```
-3*exp(-3*x)
```

d.
$$y = \sqrt{4x} = (4x)^{1/2} = 4^{1/2} \times x^{1/2} = 2x^{1/2}$$

$$\frac{dy}{dx} = \frac{1}{2} \times 2 \times x^{1/2-1} = x^{-(1/2)} = 1/x^{1/2} = 1/\sqrt{x}$$

Check with the MATLAB **diff(f)** function:

```
syms x y % define symbolic variables x and y
y=sqrt(4*x); diff(y)
ans =
1/x^(1/2)
```

2.

a.
$$\int 3x^2 dx = 3 \int x^2 dx = 3 \frac{x^{2+1}}{2+1} = x^3 + C$$

Check with the MATLAB **int(f)** function:

```
syms x y % define symbolic variables x and y
y = 3*x^2; int(y)
ans =
x^3
```

b.
$$\begin{aligned} \int (5x^3 + 6x^2 - 12) dx &= 5 \int x^3 dx + 6 \int x^2 dx - 12 \int dx = 5 \frac{x^4}{4} + 6 \frac{x^3}{3} - 12x + C \\ &= (5/4)x^4 + 2x^3 - 12x + C \end{aligned}$$

Check with the MATLAB **int(f)** function:

```
syms x y % define symbolic variables x and y
y=5*x^3+6*x^2-12; int(y)
ans =
5/4 * x^4 + 2 * x^3 - 12 * x
```

c.
$$\int 3e^{2x} dx = 3 \int e^{2x} dx = 3 \frac{e^{2x}}{2} = (3/2)e^{2x} + C$$

Check with the MATLAB **int(f)** function:

```
syms x y % define symbolic variables x and y
y=3*exp(2*x); int(y)
ans =
3/2 * exp(2 * x)
```

3.

$$\int_1^2 (2x^5 - 6x^3 + 3x) dx = \left(\frac{2}{6}x^6 - \frac{6}{4}x^4 + \frac{3}{2}x^2 \right) \Big|_1^2 = \left(\frac{2}{6} \cdot 2^6 - \frac{6}{4}2^4 + \frac{3}{2}2^2 \right) - \left(\frac{2}{6}1^6 - \frac{6}{4}1^4 + \frac{3}{2}1^2 \right)$$

a.

$$= \frac{1 \times 64}{3} - \frac{3 \times 16}{2} + \frac{3 \times 4}{2} - \frac{1}{3} + \frac{3}{2} - \frac{3}{2} = 21 \cdot \frac{1}{3} - 24 + 6 - 1/3$$
$$= 3 \text{ square units}$$

Check with the MATLAB **int(f,lower_limit,upper_limit)** function:

$$y=2*x^5-6*x^3+3*x; \text{int}(y,1,2)$$

ans =

3

b.

$$\int_3^5 (3x^2 + 6x - 8) dx = \left(\frac{3}{3}x^3 + \frac{6}{2}x^2 - 8x \right) \Big|_3^5 = (5^3 + 3 \times 5^2 - 8 \times 5) - (3^3 + 3 \times 3^2 - 8 \times 3)$$
$$= 125 + 75 - 40 - 27 - 27 + 24 = 130 \text{ square units}$$

Check with the MATLAB **int(f,lower_limit,upper_limit)** function:

$$y=3*x^2+6*x-8; \text{int}(y,3,5)$$

ans =

130

4.

Let $p_1 = 8000$, $p_2 = 16000$, $r = 0.06$, and we wish to find t . Then,

$$\int_{8000}^{16000} \frac{1}{p} dp = \int_0^t 0.06 dt$$

Integrating both sides and rearranging, we obtain

$$0.06t = \ln(16000 - 8000) = \ln\left(\frac{16000}{8000}\right) = \ln 2$$

Solving for t , we obtain

$$t = \frac{\ln 2}{0.06} = 11.55 \text{ years}$$

Chapter 6

Mathematics of Finance and Economics

This chapter is an introduction to the mathematics used in the financial community. The material presented in this appendix is indispensable to all business and technology students as well as those enrolled in continuing and adult education. Also, an executive lacking a background in finance is seriously considered handicapped in this vital respect.

6.1 Common Terms

6.1.1 Bond

A *bond* is a debt investment. That is, we loan money to an entity (company or government) that needs funds for a defined period of time at a specified interest rate. In exchange for our money, the entity will issue us a certificate, or bond, that states the interest rate that we are to be paid and when our loaned funds are to be returned (maturity date). Interest on bonds is usually paid every six months, i.e., semiannually.

6.1.2 Corporate Bond

A *corporate bond* is a bond issued by a corporation.

6.1.3 Municipal Bond

A *municipal bond* is a bond issued by a municipality and that generally is tax-free. That is, we pay no taxes on the interest that we earn. Because it is tax-free, the interest rate is usually lower than for a taxable bond.

6.1.4 Treasury Bond

A treasury bond is a bond issued by the US Government. These are considered safe investments because they are backed by the taxing authority of the US government. The interest on Treasury bonds is not subject to state income tax. Treasury bonds, or T-bonds for short, have maturities greater than 10 years, while notes and bills have lower maturities.

6.1.5 Perpetuity

A perpetuity is a constant stream of identical cash flows with no maturity date.

6.1.6 Perpetual Bond

A *perpetual bond* is a bond with no maturity date. Perpetual bonds are not redeemable and pay a steady stream of interest forever. Such bonds have been issued by the British government.

6.1.7 Convertible Bond

A *convertible bond* is a bond that can be converted into a predetermined amount of a company's equity at certain times during its life. Usually, convertible bonds offer a lower rate of return in exchange for the option to trade the bond into stock.

The conversion ratio can vary from bond to bond. We can find the terms of the convertible, such as the exact number of shares or the method of determining how many shares the bond is converted into, in the *indenture**. For example a conversion ratio of 45:1 means that every bond we hold (with a \$1000 par value) can be exchanged for 45 shares of stock. Occasionally, the indenture might have a provision that states the conversion ratio will change over the years, but this is uncommon.

6.1.8 Treasury Note

A *treasury note* is essentially a treasury bond. The only difference is that a treasury note is issued for a shorter time (e.g., two to five years) than a treasury bond.

6.1.9 Treasury Bill

A *treasury bill* is a treasury note that is held for a shorter time (e.g., three, six, or nine months to two years) than either a treasury bond or a treasury note. Interest on T-bills are paid at the time the bill matures, and the bills are priced accordingly.

6.1.10 Face Value

Face value is the dollar amount assigned to a bond when first issued.

6.1.11 Par Value

The *par value* of a bond is the face value of a bond, generally \$1,000 for corporate issues, and as much as \$10,000 for some government issues. The price at which a bond is purchased is generally not the same as the face value of the bond. When the purchase price of a bond is its par value, it is said to be purchased at par; when the purchase price exceeds the par value, it is said to be purchased above par, or at a premium. When the purchase price is below the par value, it is said to be purchased below par, or at a discount. Thus, the difference between the par value of a bond and its purchase price is termed the premium or discount whichever applies.

For example, if the face value of a bond is \$5,000 and it sells for \$4,600, it sells at a discount of \$400; if the same bond sells for \$5,300, then it sells at a premium of \$300.

* *Indenture* is a contract between an issuer of bonds and the bondholder stating the time period for repayment, amount of interest paid, if the bond is convertible and if so at what price or what ratio, and the amount of money that is to be repaid.

6.1.12 Book Value

Book value is the value of a bond at any date intermediate between its issue and its redemption.

6.1.13 Coupon Bond

A *coupon bond* is a debt obligation with coupons representing semiannual interest payments attached. No record of the purchaser is kept by the issuer, and the purchaser's name is not printed on the certificate.

6.1.14 Zero Coupon Bond

A *zero coupon bond* is a bond that generates no periodic interest payments and is issued at a discount from face value. The Return is realized at maturity. Because it pays no interest, it is traded at a large discount.

6.1.15 Junk Bond

A *junk bond* is a bond purchased for speculative purposes. They have a low rating and a higher default risk. Typically, junk bonds offer interest rates three to four percentage points higher than safer government bonds. Generally, a junk bond is issued by a corporation or municipality with a bad credit rating. In exchange for the risk of lending money to a bond issuer with bad credit, the issuer pays the investor a higher interest rate. "High-yield bond" is a name used by the for junk bond issuer.

6.1.16 Bond Rating Systems

Standard and Poor's and Moody's are the bond-rating systems used most often by investors. Their ratings are as shown in Table 6.1.

TABLE 6.1 Bond-Ratings

Moody's	Standard & Poor's	Grade	Risk
Aaa	AAA	Investment	Lowest risk/highest quality
Aa	AA	Investment	Low risk/high quality
A	A	Investment	Low risk/upper medium quality
Baa	BBB	Investment	Medium risk/medium quality
Ba, B	BB, B	Junk	High risk/low quality
Caa, Ca, C	CCC, CC, C	Junk	Highest risk/highly speculative
C	D	Junk	In default

6.1.17 Promissory Note

A *promissory note* is an unconditional written promise made by one party to another, to pay a stipulated sum of money either on demand or at a definite future date. The sum of money stated in the note is its face value and if the note is payable at a definite date, this date is termed its “due date” or “date of maturity”. If no mention of interest appears, the maturity value of the note is its face value. But if payment of interest is specified, the maturity value is the face value of the note plus the accrued interest.

For example, let us assume that a promissory note dated January 10, 2007, reads as follows: “Six months after date I promise to pay to the order of John Doe the sum of \$10,000 with interest at 8% per annum. Here, the face value of the note is \$10,000, its maturity date is July 10, 2007, and the maturity value is computed as

$$\text{Interest} = \$10,000 \times 0.08 \times \frac{1}{2} = \$400$$

Therefore, the maturity value is

$$\text{Maturity value} = \$10,000 + \$400 = \$10,400$$

6.1.18 Discount Rate

Discount rate is the interest rate charged by Reserve Banks when they extend credit to depository institutions either through advances or through the discount of certain types of paper, including ninety-day commercial paper. Discount rate also refers to the fee a merchant pays its merchant bank for the privilege to deposit the value of each day's credit purchases. This fee is usually a small percentage of the purchase value.

6.1.19 Prime Rate

Prime rate is the interest rate charged by banks to their most creditworthy customers (usually the most prominent and stable business customers). The rate is almost always the same amongst major banks. Adjustments to the prime rate are made by banks at the same time; although, the prime rate does not adjust on any regular basis.

6.1.20 Mortgage Loan

In the case of many types of loans, regardless of the duration, the creditor will require some form of security for the repayment of the loan, such as property owned by the debtor. The legal instrument under which the property is pledged is termed a “mortgage,” or “trust deed.” In general, title and possession of the property remain with the mortgagor, or borrower. However, since the mortgagee, or lender, has a vested interest in maintaining the value of this property unimpaired, at least to the extent of the unpaid balance of the loan, the terms of the mortgage generally require

that the mortgagor keep the property adequately insured, make necessary repairs, and pay all taxes as they become due.

Should it occur that there is more than one lien applying to a particular piece of property, then the mortgages are referred to as the first mortgage, second mortgage, home-equity, etc., corresponding to the priority of the claim. In the event of default in repaying the loan, or in the event of a violation of the terms of the mortgage on the part of the owner, the mortgagee can request a court of law to undertake the sale of the property in order to recover the balance of his loan.

Property is often purchased with borrowed funds, and the property thus acquired constitutes the security behind the loan. This is frequently true, for example, in the purchase of homes. Mortgage loans of this type generally possess an amortization feature. Although the term “amortization” means simply the liquidation of a debt, it is generally used to denote the process of gradually extinguishing a debt through a series of equal periodic payments extending over a stipulated period of time. Each payment can be regarded as consisting of a payment of the interest accrued for that particular period on the unpaid balance of the loan and a partial payment of principal. With each successive payment, the interest charge diminishes, thereby accelerating the reduction of the debt.

6.1.21 Predatory Lending Practices

Several types of predatory lending have been identified by federal investigators. The most common are:

Equity stripping: Occurs when a loan is based on the equity of a home rather than the borrower’s ability to repay. These loans often have high fees, prepayment penalties, and more strict terms than a regular loan.

Packing. The practice of adding credit insurance and other extras to the loan. The supplements to the loan are very profitable to lenders and are generally financed in a single up-front of balloon payment.

Flipping: A form of equity stripping, this occurs when a lender persuades a borrower to repeatedly refinance a loan within a short period of time. The lender typically charges high fees each time.

6.1.22 Annuity

An *annuity* is series of equal payments made at equal intervals or periods of time. When paid into a fund which is invested at compound interest for a specified number of years, the annuity is referred to as a sinking fund.

6.1.23 Ordinary Annuity

An *ordinary annuity* is a series of equal payments or receipts occurring over a specified number of periods with the payments or receipts occurring at the end of each period.

6.1.24 Sinking Fund

A *sinking fund* is an interest-earning fund in which equal deposits are made at equal intervals of time. Thus, if the sum of \$1,000 is placed in a fund every three months, a sinking fund has been established. See also annuity.

As an example, let us assume that a company plans to purchase a new asset five years from now, and the purchase price at that time will be \$10,000. To accumulate this sum, it will make annual deposits in a sinking fund for the next five years and the fund will be earning interest at 4% per annum. If the fund did not earn interest, the required deposit would be \$2,000 per annum. However, since the principle in the fund will be continuously augmented through the accrual of interest, the actual deposit required will be less than \$2,000.

6.2 Interest

A *paid interest* is charge for a loan, usually a percentage of the amount loaned. If we have borrowed money, from a bank or credit union, we have to pay them interest. An *earned interest* is a percentage of an amount that we receive from a bank or credit union for the amount of money that we have deposited. In other words, if we put money into a bank or credit union they will pay us interest on this money. In this section we will discuss simple interest, simple interest over multiple years, and simple interest over a fraction of a year.

6.2.1 Simple Interest

Simple interest is calculated on a yearly basis (annually) and depends on the interest rate. The rate is often given per annum (p.a.) which means per year.

Let

P_0 = principal of the loan, i.e., present value

i = interest rate per period (simple interest) expressed as a decimal

n = number of periods constituting the life of the loan

P_n = maturity value of the loan, i.e., future value

Then, the maturity (future) value of a loan is computed as

$$P_n = P_0 + P_0 ni = P_0(1 + ni) \quad (6.1)$$

If we know the future value P_n , we can compute the present value P_0 using the formula

$$P_0 = \frac{P_n}{1 + ni} \quad (6.2)$$

Example 6.1

We deposit \$500.00 into a bank account with an interest rate of 2% per annum. We want to find how much money we have in the account after one year.

Solution:

For this example, $P_0 = 500$, $n = 1$, and $i = 0.02$. Therefore, after one year the maturity value P_n will be

$$P_n = P_0(1 + ni) = 500(1 + 1 \times 0.02) = \$510$$

That is, we have earned \$10.00 interest.

Example 6.2

We deposit \$350.00 in a simple interest account for 3 years. The account pays interest at a rate of 3% per annum. How much do we have in this account after 3 years?

Solution:

$$P_n = P_0(1 + ni) = 350(1 + 3 \times 0.03) = \$381.50$$

That is, after 3 years we will have \$381.50 in this account.

If money is not left in a bank account for a whole year then only a fraction of the interest is paid.

Example 6.3

We deposit \$50,000 in a simple interest account for 6 months. This account pays interest at a rate of 8.5% per annum. How much do we have in this account after 6 months?

Solution:

Since the annual rate is 8.5%, the semiannual rate is

$$\frac{1}{2} \times 0.085 = 0.0425$$

Then,

$$A = P(1 + ni) = 50000(1 + 0.0425) = \$52125$$

That is, after 6 months we will have \$52125 in this account.

6.2.2 Compound Interest

Compound interest includes interest earned on interest. In other words, with compound interest, the interest rate is applied to the original principle and any accumulated interest.

Let

P_0 = principal of the loan, i.e., present value

i = interest rate per period (simple interest) expressed as a decimal

n = number of periods constituting the life of the loan

P_n = maturity value of the loan, i.e., future value

Then, for:

Period 1

$$\begin{aligned}\text{Principal at start of period} &= P_0 \\ \text{Interest earned} &= P_0 i \\ \text{Principal at end of Period 1} &= P_0 + P_0 i \\ P_1 &= P_0 (1 + i)\end{aligned}$$

Period 2

$$\begin{aligned}\text{Principal at start of period} &= P_0 (1 + i) \\ \text{Interest earned} &= P_0 (1 + i) i \\ \text{Principal at end of Period 2} &= P_0 (1 + i) + P_0 (1 + i) i = P_0 (1 + i)(1 + i) \\ P_2 &= P_0 (1 + i)^2\end{aligned}$$

Period 3

$$\begin{aligned}\text{Principal at start of period} &= P_0 (1 + i)^2 \\ \text{Interest earned} &= P_0 (1 + i)^2 i \\ \text{Principal at end of Period 3} &= P_0 (1 + i)^2 + P_0 (1 + i)^2 i = P_0 (1 + i)^2 (1 + i) \\ P_3 &= P_0 (1 + i)^3\end{aligned}$$

and so on. Thus, for **Period n**

$$P_n = P_0 (1 + i)^n \tag{6.3}$$

The formula of (6.3) assumes that the interest is compounded annually. But if the interest is compounded q times per year, the maturity value is computed using the formula

$$P_n = P_0 \left(1 + \frac{i}{q}\right)^{nq} \quad (6.4)$$

Example 6.4

We deposit \$350.00 in a compound interest account for 3 years now. The account pays interest at a rate of 3% per annum compounded annually. How much will we have in this account after 3 years?

Solution:

Since the interest rate is compounded annually, the formula of (6.3), Page 6–8, applies for this example. Thus,

$$P_n = P_0 (1 + i)^n = 350(1 + 0.03)^3 = \$382.45$$

That is, after 3 years we will have \$382.45 in this account.

The Microsoft Excel financial functions are described in Section 6.8. For our present discussion, we will use the appropriate function to verify our answers.

The function

`=FV(0.03,3,0,-350,0)`

returns \$382.45. In the Excel formula above 0.03 represents the annual interest rate expressed in decimal form, 3 represents the number of periods (3 years for our example), 0 represents the payments made in each period. For this example it is zero since there are no periodic payments made. The present value of \$350.00 is shown as a negative number since it is an outgoing cash flow. Cash we receive, such as dividends, is represented by positive numbers. The last 0 in the formula indicates that maturity of this investment occurs at the end of the 3-year period. Cash we receive, such as dividends, is represented by positive numbers.

Example 6.5

We deposit \$350.00 in a compound interest account for 3 years, now. The account pays interest at a rate of 3% per annum compounded monthly. How much do we have in this account after three years?

Solution:

Since the interest rate is compounded monthly, the formula of (6.4), Page 6–9, applies for this example. Thus,

$$P_n = P_0 \left(1 + \frac{i}{q}\right)^{nq} = 350 \left(1 + \frac{0.03}{12}\right)^{3 \times 12} = \$382.92$$

Check with Microsoft Excel:

The function

=FV(0.03/12,3*12,0,-350,0)

returns \$382.92. In the above formula we divided the interest rate by 12 to represent the compounded interest on a monthly basis, and we multiplied the period (3 years) by 12 to convert to 36 monthly periods.

Example 6.6

This example computes interest on a principal sum and illustrates the differences between simple interest and compound interest. We have used a \$100.00 principal and 7% interest.

Simple Interest

The interest rate is applied only to the original principal amount in computing the amount of interest as shown on Table 6.2.

TABLE 6.2 Simple Interest

Year	Principal(\$)	Interest(\$)	Ending Balance(\$)
1	100.00	6.00	106.00
2	100.00	6.00	114.00
3	100.00	6.00	121.00
4	100.00	6.00	128.00
5	100.00	6.00	135.00
	Total interest	35.00	

Compound Interest

The interest rate is applied to the original principle and any accumulated interest as shown on Table 6.3.

TABLE 6.3 Compound Interest

Year	Principal(\$)	Interest(\$)	Ending Balance(\$)
1	100.00	6.00	106.00
2	106.00	6.49	114.49
3	114.49	8.01	122.50
4	122.50	8.58	131.08
5	131.08	9.18	140.26
	Total interest	40.26	

Quite often, both the present value P_0 and maturity (future) value P_n are known and we are interested in finding the interest rate i or the period n . Since equation (6.3), Page 6–8, contains the interest rate and the period, we can express it in logarithmic form by taking the common log of both sides of that equation. Thus,

$$\log P_n = \log[P_0(1+i)^n] = \log P_0 + \log[(1+i)^n] = \log P_0 + n \log(1+i) \quad (6.5)$$

or

$$\log \frac{P_n}{P_0} = n \log(1+i) \quad (6.6)$$

Example 6.7

On January 1, 2002, we borrowed \$5000.00 from a bank and the debt was discharged on December 31, 2006, with a payment of \$6100.00. Assuming that the interest was compounded annually, what was the interest rate that we paid?

Solution:

For the example, $P_0 = 5000$, $P_5 = 6100$, and $n = 5$. By substitution in (6.6) above, we obtain

$$\log \frac{6100}{5000} = 5 \log(1+i)$$

or

$$\log \frac{6.1}{5} = 5 \log(1+i)$$

Using Microsoft Excel, we find that =LOG10(6.1/5) returns 0.0864

Then,

$$\log(1 + i) = \frac{0.0864}{5} = 0.0173$$

We recall* that $\log_{10} M = x$ implies that $M = 10^x$. Using Excel we find that $=10^{0.0173}$ returns 1.041. Thus,

$$1 + i = 10^{0.0173} = 1.041$$

and solving for i we obtain

$$i = 1.041 - 1 = 0.041$$

Therefore, the interest rate that we paid is 4.1% to the nearest tenth of one per cent.

Check with Microsoft Excel:

The function $=\text{RATE}(5,0,5000,-6100,0.05)$ returns 4.06%. In this Excel formula 5 represents the number of the periods (5 years), 0 represents the number of periodic payments made during the 5-year period which is zero for this example, 5000 is the amount of the loan we received, -6100 is the amount we paid, and 0.05 is our guess of what the interest rate might be.

Example 6.8

On January 1, 2007, we deposited \$5000.00 to an account paying 4.5% annual interest. How long will it take for this amount to grow to \$7500.00?

Solution:

For this example, $P_0 = 5000$, $P_n = 7500$, and $i = 0.045$. By substitution in (6.6), Page 6–11, we obtain

$$\log \frac{7500}{5000} = n \log(1 + 0.045)$$

or

$$\log 1.5 = n \log(1.045)$$

or

$$n = \frac{\log 1.5}{\log(1.045)}$$

Using Excel we find that $=\text{LOG10}(1.5)$ returns 0.1761 and $=\text{LOG10}(1.045)$ returns 0.0191. Then,

$$n = \frac{0.1761}{0.0191} = 9.22$$

* For a review of logarithms and powers, please refer to Appendix A.

That is, it will take 9.22 years for our \$5000.00 investment to grow to \$7500.00.

Check with Microsoft Excel:

The function =NPER(0.045,0,-5000,7500,0) returns 9.21 years.

Often, we want to translate a given sum of money to its equivalent value at a prior valuation date. To adhere to our previous notation, we shall let P_0 denote the value of a given sum of money at a specified valuation or zero date, and let P_{-n} denote the value of P_0 at a valuation date n periods prior to that of P_0 . We will assume that the sum of money P_{-n} is deposited in a fund at interest rate i per period. Thus, its value at the end of n periods will be P_0 . Applying equation (6.3), Page 6–8, but substituting P_{-n} for P_0 , and P_0 for P_{-n} , we obtain

$$P_0 = P_{-n}(1 + i)^n$$

or

$$P_{-n} = \frac{P_0}{(1 + i)^n}$$

or

$$P_{-n} = P_0(1 + i)^{-n} \quad (6.7)$$

Example 6.9

We possess two promissory notes each having a maturity value of \$1000.00. The first note is due 2 years hence, and the second 3 years hence. At an annual interest rate of 6% what proceeds will we obtain by discounting the notes at the present date?

Solution:

Equation (6.7) applies to this example. Thus, the value of the first note is

$$P_{-2} = P_0(1 + i)^{-2} = 1000(1 + 0.06)^{-2} = 1000 \times 0.89000 = 890.00$$

and the value second note is

$$P_{-3} = P_0(1 + i)^{-3} = 1000(1 + 0.06)^{-3} = 1000 \times 0.83962 = 839.62$$

Thus, the total discount value is

$$890.00 + 839.62 = 1729.62$$

Chapter 6 Mathematics of Finance and Economics

Check with Microsoft Excel: The function =PV(0.06,2,0,-1000) returns \$890.00, and the function =PV(0.06,3,0,-1000) returns \$839.62.

The discount applicable to the sum P_0 for n periods at interest rate i is

$$P_0 - P_{-n} = P_0 - P_0(1+i)^{-n} = P_0[1 - (1+i)^{-n}] \quad (6.8)$$

Example 6.10

A note whose maturity value of \$5000.00 is to be discounted 1 year before maturity at an interest rate of 6%. Compute the discount value using:

- Equation (6.7), Page 6–13.
- Equation (6.8), this page

Solution:

a.

$$P_{-1} = P_0(1+i)^{-1} = 5000(1+0.06)^{-1} = 5000(0.94340) = 4716.98$$
$$\text{Discount value} = 5000.00 - 4716.98 = 283.02$$

b.

$$P_0 - P_{-1} = P_0[1 - (1+i)^{-1}] = 5000[1 - (1+0.06)^{-1}] = 5000(1 - 0.94340) = 283.02$$

Check with Microsoft Excel:

We will subtract the function =PV(0.06,1,1,-5000) from \$5000.00. Thus, =5000-PV(0.06,1,1,-5000) returns \$283.96.

Example 6.11

Let us assume that the following transactions occurred in a fund whose annual interest rate is 8%.

- An initial deposit of \$1000.00 was made on January 1, 1996
- A withdrawal of \$600.00 was made on January 1, 2000
- A deposit of \$800.00 was made on January 1, 2005

If the account was closed on December 31, 2006, what was the final principal?

Solution:

For convenience, the history of the principal in the fund is shown in Figure 6.1.

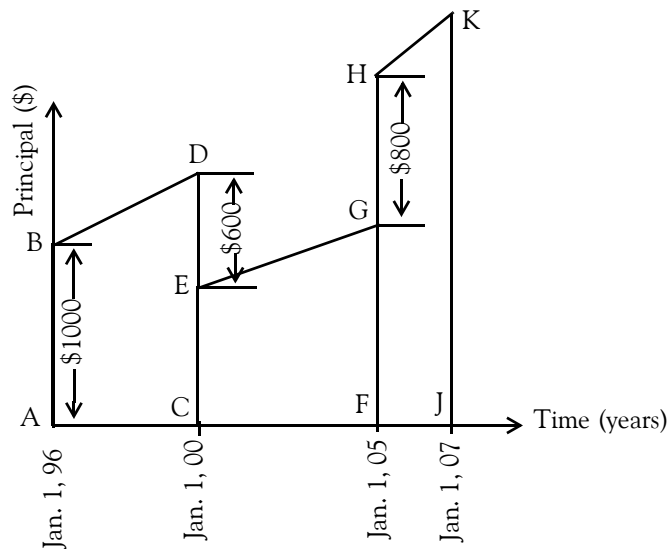


Figure 6.1. Plot for Example 6.11.

The four significant dates and the principal at these dates are computed as follows:

$$AB = 1000$$

$$CD = 1000(1.08)^4 = 1000(1.36049) = 1360.49$$

$$CE = CD - ED = 1360.49 - 600 = 760.49$$

$$FG = 760.49(1.08)^5 = 760.49(1.46933) = 1117.41$$

$$FH = FG + GH = 1117.41 + 800 = 1917.41$$

$$JK = 1917.41(1.08)^2 = 1917.41(1.1664) = 2236.47$$

Therefore, the principal on December 31, 2006 was \$2236.47.

The principal at any nonsignificant date can be found by equation (6.3), Page 6–8.

Quite often, it is desirable to translate a given group of money values to a second group equivalent to the first at another date. The procedure is illustrated with the following example.

Example 6.12

Johnson owes Taylor \$3000 due December 31, 2007, and \$2000 due on December 31, 2009. By mutual consent, the terms of payment are altered to allow Johnson to discharge the debt by mak-

Chapter 6 Mathematics of Finance and Economics

ing a payment of \$4000 on December 31, 2010, and a payment for the balance on December 31, 2011. If the annual interest rate is 5%, what will be the amount of Johnson's final payment?

Solution:

The given data are tabulated in Table 6.4, where the final payment whose value is to be computed, is denoted as x .

TABLE 6.4 Table for Example 6.12

Payments—Original Plan	Payments—Revised Plan
\$3,000 on 12/31/07	\$4,000 on 12/31/10
\$2,000 on 12/31/09	x on 12/31/11

Since the two groups of payments are equivalent to one another, based on an interest rate of 5%, the total value of one group must equal the total value of the other group at every instant of time. To determine x , we select some standard date for evaluating all sums of money involved. For convenience, we shall choose Dec. 31, 2011, as the valuation date. Then,

$$\begin{aligned}4000(1.05) + x &= 3000(1.05)^4 + 2000(1.05)^2 \\4200 + x &= 3000(1.21551) + 2000(1.10250) \\4200 + x &= 3646.53 + 2205\end{aligned}$$

or

$$x = 3646.53 + 2205 - 4200 = 1651.53$$

That is, with the revised payment plan Johnson must pay Taylor \$4000 on December 31, 2010, and \$1651.53 on December 31, 2011.

Had we selected December 31, 2007 as our valuation date, our solution would be

$$\begin{aligned}4000(1.05)^{-3} + x(1.05)^{-4} &= 3000 + 2000(1.05)^{-2} \\3455.36 + x(1.05)^{-4} &= 3000 + 1814.06 \\x(1.05)^{-4} &= 4814.06 - 3455.36 = 1358.70 \\x &= 1358.70(1.05)^4 = 1358.70(1.21551) = 1651.51\end{aligned}$$

We can also verify the result by application of a chronological sequence as shown in Table 6.5.

TABLE 6.5 Chronological sequence in table form for Example 6.12

Value of first debt on 12/31/05 = $3000(1.05)^2$	=	\$3,307.50
Value of second debt on 12/31/05	=	2,000.00
Total principal on loan on 12/31/05	=	5,307.50
Interest earned in 2006 = $5307.50(0.05)$	=	265.38
Principal on loan on 12/31/06	=	5,572.88
Payment on 12/31/06	=	4,000.00
Principal on loan on 01/01/07	=	1,572.88
Interest earned in 2007 = $1572.88(0.05)$	=	78.64
Principal on loan on 12/31/07	=	1,651.52
Payment required on 12/31/07	=	1,651.52

We should emphasize that for Example 6.12, these two groups of payments are equivalent to one another only for an interest rate of 5%. Implicit in our reasoning is the assumption that the creditor can reinvest each sum of money he receives in a manner that continues to yield 5%. Should there occur any variation of the interest rate, then the equivalence of the two groups of payments becomes invalid. For example, assume that the creditor is able to reinvest his capital at a rate of only 4% per cent. For this situation, the revised terms of payment are more advantageous to him since, by deferring the collection of the money due him, he enables his capital to earn the higher rate of interest for a longer period of time.

In the preceding example, the replacement of one group of money values with an equivalent one was accomplished with a known interest rate. Many problems arise in practice, however, in which the equivalent groups of money values are known, and it is necessary to determine the interest rate on which their equivalence is predicated. Problems of this nature can only be solved by a trial-and-error method, and if the exact interest rate is not one of those listed in the interest tables, it can be approximated by means of straight-line interpolation.

Example 6.13

Smith owed Jones the sum of \$1000 due Dec. 31, 2002, and \$4,000, due Dec. 31, 2004. Because Smith was unable to meet these obligations as they became due, the debt was discharged by means of a payment of \$2000 on Dec. 31, 2005, and a second payment of \$3850 on Dec. 31, 2006. What annual interest rate was intrinsic in these payments?

Solution:

Let i denote the interest rate. If the expression $(1 + i)^n$ is expanded by the binomial theorem* and all terms beyond the second are discarded, we obtain

$$(1 + i)^n \approx 1 + ni \quad (6.9)$$

We will apply this approximate value to obtain a first approximation of the value of i . This procedure is tantamount to basing our first calculation on the use of simple rather than compound interest. This approximation understates the value of $(1 + i)^n$, and the degree of error varies in proportion to n .

Since the two groups of payments in this problem are equivalent, the value of one group equals the value of the other at any valuation date that we select. Equating the two groups at the end of 2004 on the basis of simple interest, we obtain

$$\begin{aligned} [1,000 + 4i(1000)] + [4,000 + 2i(4000)] &= [2000 + i(2000)] + 3850 \\ 4000i + 8000i - 2000i &= 2000 + 3850 - 1000 - 4000 \\ 10000i &= 850 \\ i &= 0.085 \end{aligned}$$

or $i = 8.5\%$ as a first approximation.

But because application of equation (6.9) yields an approximation, our computations have produced an understatement of the value of each sum of money except the last. Moreover, since the amount of error increases as n increases, it is evident that we have reduced the importance of the money values having early valuation dates and inflated the importance of those having late valuation dates. The true rate, therefore, will be less than 8.5%. Let us assume an 8% rate. If this were the actual rate, then the payment x required at the end of 2004 is determined as follows:

$$\begin{aligned} 1000(1.08)^4 + 4000(1.08)^2 &= 2000(1.08) + x \\ x &= 1000(1.08)^4 + 4000(1.08)^2 - 2000(1.08) = 3866.09 \end{aligned}$$

Since the actual payment was \$3850.00, the true interest rate is less than 8%. Let us try a 7.5% rate.

$$\begin{aligned} 1000(1.075)^4 + 4000(1.075)^2 &= 2000(1.075) + x \\ x &= 1000(1.075)^4 + 4000(1.075)^2 - 2000(1.075) = 3807.97 \end{aligned}$$

* The binomial theorem is discussed in Chapter 10.

Since this value is less than \$3850.00, we expect the interest rate to lie between 7.5% and 8% as shown on the graph of Figure 6.2.

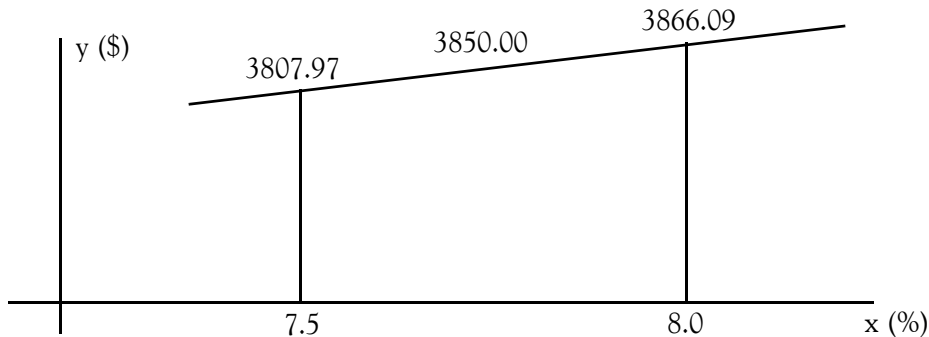


Figure 6.2. Graph for Example 6.13

The interest rate between 7.5% and 8% can be found by applying straight-line interpolation as we discussed on Chapter 1, Relation (1.48), that is,

$$x_i = \frac{1}{\text{slope}} \times (y_i - y_1) + x_1 = \frac{1}{m} \times (y_i - y_1) + x_1$$

For our example,

$$\begin{aligned} x|_{y=3850} &= \frac{1}{\text{slope}} \times (3850.00 - 3807.97) + 0.075 = \frac{1}{\frac{3866.09 - 3807.97}{0.080 - 0.075}} \times 42.03 + 0.075 \\ &= \frac{0.005}{58.12} \times 42.03 + 0.075 = 0.0786 \end{aligned}$$

Therefore, the actual interest rate was 7.86%.

The problem of calculating an interest rate is often encountered where an investment is made that produces certain known returns at future dates and we wish to determine the rate of return earned by the investment. The following example illustrates the procedure.

Example 6.14

An undeveloped lot was purchased in January, 1999, for \$2000. Taxes and assessments were charged at the end of each year are as shown in Table 6.6.

The owner paid the charges for 2001 and 2002, but not for subsequent years. At the end of 2006, the lot was sold for \$3000, the seller paying the back charges at 7% interest compounded annu-

Chapter 6 Mathematics of Finance and Economics

ally and also paying a commission of 5% to a real estate agent. What rate of return was realized on the investment?

TABLE 6.6 Data for Example 6.14

Year	2001	2002	2003	2004	2005	2006
Taxes Paid	30	30	200	30	30	30

The payment made by the owner at the end of 2006 for the back charges consisted of the following:

TABLE 6.7 Computations for Example 6.14

Year	Computations	\$ Amount
2003	$200 \times (1.07)^3$	\$245.01
2004	$30 \times (1.07)^2$	34.35
2005	$30 \times (1.07)$	32.10
	Total back charges	\$311.46
2006	Taxes	\$30.00
2006	Real Estate Commission $\$3,000 \times 5\%$	\$150.00
	Total payment at end of 2002	\$491.46
2006	Selling Price	\$3,000.00
2006	Net profit $(\$3,000.00 - \$491.46)$	\$2,508.54

Next, we form the following two equivalent groups of money shown in Table 6.8.

TABLE 6.8 Disbursements and Receipts equivalence for Example 6.14

Group 1 – Disbursements		Group 2 – Receipts	
Date	Amount	Date	Amount
1/1/2001	\$2,000.00	12/31/2006	\$2508,54
12/31/2001	30.00		
12/31/2002	30.00		

Let i denote the investment rate. Using simple interest to obtain a first approximation and selecting Dec. 31, 2006, as our valuation date, we obtain

$$[2000 + 2000 \times 6i] + [30 + 30 \times 5i] + [30 + 30 \times 4i] = 2508.54$$

or

$$12270i = 2508.54 - 2060 = 448.54$$

or

$$i = \frac{448.54}{12270} = 0.03655583 \approx 3.66\%$$

That is, as a first approximation, the interest rate is 3.66%. But our approximation has given insufficient weight to the disbursements and therefore, the true investment rate is less than this. To understand this, let us consider the factor $(1 + i)^n$. Using the binomial theorem and discarding all terms beyond the second, we obtain

$$(1 + i)^n \approx 1 + ni \quad (6.10)$$

It is obvious that this approximation understates the value of $(1 + i)^n$ and the error varies in proportion to n .

Let us try a 3.5% approximation. The income y required to yield this rate is found as follows:

$$2000 \times (1.035)^6 + 30 \times (1.035)^5 + 30 \times (1.035)^4 = y$$

or

$$y|_{x=0.035} = 2458.11 + 35.63 + 34.43 = 2528.57$$

Since the actual income was \$2508.54, the investment rate is less than 3.5%. Let us try a 3.0% rate. Then,

$$2000 \times (1.03)^6 + 30 \times (1.03)^5 + 30 \times (1.03)^4 = y$$

or

$$y|_{x=0.030} = 2388.10 + 34.78 + 33.77 = 2456.65$$

Since this value is less than \$2508.54, we expect the interest rate to lie between 3.0% and 3.5% as shown on the graph of Figure 6.3.

The interest rate between 3.0% and 3.5% can be found by applying straight-line interpolation as we discussed on Chapter 1, Relation (1.47), Page 1–21, that is,

$$x_i = \frac{1}{\text{slope}} \times (y_i - y_1) + x_1 = \frac{1}{m} \times (y_i - y_1) + x_1$$

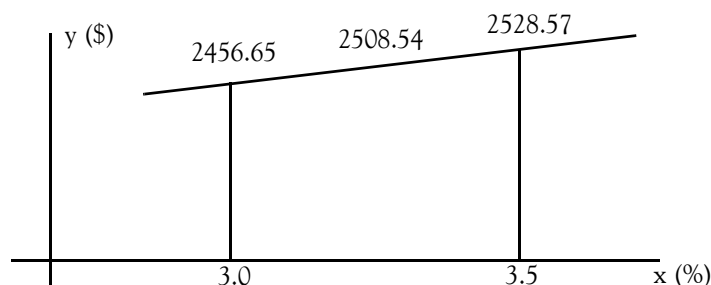


Figure 6.3. Graph for Example 6.14

For our example,

$$\begin{aligned}x|_{y=2508.54} &= \frac{1}{\text{slope}} \times (2508.54 - 2456.65) + 0.03 = \frac{1}{\frac{2528.57-2456.65}{0.035-0.030}} \times 51.89 + 0.03 \\ &= \frac{0.005}{71.92} \times 51.89 + 0.03 = 0.0336\end{aligned}$$

Therefore, the actual interest rate is 3.36%.

6.2.3 Effective Interest Rate

Consider the sum of \$1000 to be deposited in a fund earning interest at the rate of 8% per annum compounded quarterly. Since the interest is compounded quarterly, 8% is merely a nominal interest rate; the true interest rate is 2% per quarterly period. At the end of 6 months, or two interest periods, this sum has increased to

$$P_2 = 1000(1 + 0.02)^2 = 1040.40$$

If this fund, instead of earning 2% interest per quarterly period, were earning interest at the rate of 4.04% per cent per semiannual period, then the principal at the end of any 6-month period would be the same. Hence, assuming that the principal is not withdrawn from the fund except at the end of a 6-month period, we may regard the two interest rates as being equivalent to one another.

At the expiration of 1 year, or four interest periods, the original sum of \$1000 has increased to

$$P_4 = 1000(1 + 0.02)^4 = 1082.43$$

Similarly, if this fund earned interest at the rate of 8.243% per cent compounded annually, the principal at the end of each year would be the same. This interest rate, then, is also equivalent to the given rate.

We have thus illustrated the equivalence of these three interest rates:

2% per quarterly period

4.04% per semiannual period

8.243% per annual period

If a given interest rate applies to a period less than 1 year, then its equivalent rate for an annual period is referred to as its *effective rate*. Thus, the effective rate corresponding to a rate of 2% per

quarterly period is 8.243% per cent. The effective rate is numerically equal to the interest earned by a principal of \$1.00 for 1 year.

In general, let

j = nominal interest rate

r = effective interest rate

m = number of compoundings per year

Then,

$$\text{Effective rate} = r = \left(1 + \frac{j}{m}\right)^m - 1 \quad (6.11)$$

Example 6.15

We wish to make a ratio comparison of the following two interest rates:

First rate, 6% per annum compounded monthly = 1/2% per month

Second rate, 4% per annum compounded semiannually = 2% per semiannual period

Since the two rates apply to unequal periods of time, a direct comparison is not possible. However, we can establish a basis of comparison by determining the effective rate corresponding to each:

Effective rate for 6% compounded monthly:

$$r_1 = \left(1 + \frac{0.06}{12}\right)^{12} - 1 = 6.168\%$$

Effective rate for 4% compounded semiannually

$$r_2 = \left(1 + \frac{0.04}{2}\right)^2 - 1 = 4.04\%$$

Therefore,

$$\frac{r_1}{r_2} = \frac{6.168\%}{4.04\%} = 1.527$$

The calculation of effective interest rates is therefore highly useful for comparative purposes.

6.3 Sinking Funds

We defined a sinking fund in Section 6.1. As a review, a sinking fund is an interest-earning fund in which equal deposits are made at equal intervals of time. Thus, if the sum of \$100 is placed in a fund every 3 months, a sinking fund has been established. A sinking fund is generally created for the purpose of gradually accumulating a specific sum of money required at some future date. For example, when a corporation has floated an issue of bonds it will often set aside a portion of its annual earnings to assure itself sufficient funds to retire the bonds at their maturity. The money reserve need not remain idle but can be invested in an interest-earning fund.

Assume that a business firm plans to purchase a new machine 5 years hence, the purchase price being \$10000. To accumulate this sum, it will make annual deposits in a sinking fund for the next 5 years, the fund earning interest at the rate of 4% per annum. How much should each deposit be? (In reality, it would be pointless to deposit the fifth sum of money, since the fund will be closed at the same date. However, it will simplify our discussion if we consider that the deposit is actually made.)

Now, if the fund did not earn interest, the required annual deposit would be simply $\$10000/5$ or \$2000. However, since the principal in the fund will be continuously augmented through the accrual of interest, the actual deposit required will be somewhat less than \$2000. We shall soon derive a formula for calculating the required deposit.

In the discussion that follows, we shall assume, if nothing is stated to the contrary, that the following conditions prevail:

1. The sinking fund is established at the beginning of a specific interest period. We shall identify the date on which the fund is established as the *origin date* of the fund.
2. The interval between successive deposits, known as the *deposit period*, is equal in length to an interest period.
3. Each deposit is made at the end of an interest period. Consequently, there is a time interval of one interest period between the date the fund is established and the date the first deposit is made.
4. The date at which a sinking fund is to terminate is always the last day of a specific interest period. We shall call this date the *terminal date* of the fund. The principal in the fund at its termination will, of course, include the interest earned during the last period and the final deposit in the fund, made on the terminal date.

A sinking fund satisfying the above requirements is referred to as an *ordinary sinking fund*. The duration of the fund is referred to as the *term of the fund*.

To illustrate the above definitions, assume that a sinking fund is created on Jan. 1, 2007, to consist of 10 deposits of \$1000 each. The interest (and deposit) period of the fund is 1 year. Then

Origin date = Jan. 1, 2007

Date of first deposit = Dec. 31, 2007

Terminal date (and date of last deposit) = Dec. 31, 2016

Term of fund = 10 years

At the end of each interest period, the principal in the fund increases abruptly as a result of the compounding of the interest earned during that period and the receipt of the periodic deposit. Where the principal in a sinking fund is to be calculated at a date intermediate between its origin and termination, we shall in all cases determine the principal on the last day of a particular interest period, immediately after these two events have transpired.

Example 6.16

The sum of \$500 is deposited in a sinking fund at the end of each year for 4 years. If the interest rate is 6% compounded annually, what is the principal in the fund at the end of the fourth year?

Solution:

The development of the principal in the fund is recorded graphically in Figure 6.4 by the method of chronological sequence. This graph is intended solely as an aid in visualizing the growth of the principal, not as a means of achieving a graphical solution of the problem.

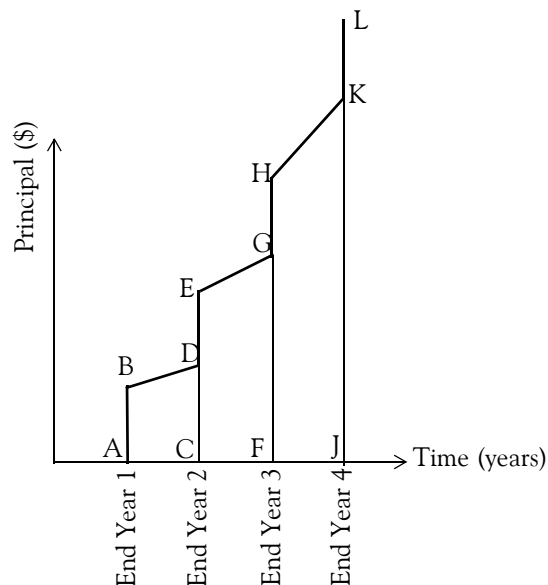


Figure 6.4. Graph for Example 6.16

At the end of the first year, the sum of \$500, represented by line segment AB, is placed in the fund. During the second year, the principal earns interest at the rate of 6% and amounts to \$530

Chapter 6 Mathematics of Finance and Economics

at the end of the second year represented by line segment CD. A deposit of \$500 at this time (represented by line segment DE) increases the principal to the amount of \$1030 (represented by line segment CE). This two-cycle process is repeated each year until the final deposit (line segment KL) is made. The growth of the principal is recorded in Table 6.9.

TABLE 6.9 Computations for Example 6.16

Year	Principal in fund at beginning of period	Interest earned	Deposit at end of period	Principal in fund at end of period
1	0.00	0.00	500.00	500.00
2	500.00	30.00	500.00	1030.00
3	1030.00	61.80	500.00	1591.80
4	1591.80	95.51	500.00	2187.31

Thus, the principal in the sinking fund at the end of the fourth year is \$2187.31.

Of course, we could have computed the principal at the end of the fourth year as shown in Table 6.10.

TABLE 6.10 Alternate Computations for Example 6.16

Deposit Number	Amount	Number of years in fund	Value at the end of the fourth year		
1	\$500.00	3	$500(1.06)^3$	=	\$595.51
2	500.00	2	$500(1.06)^2$	=	561.80
3	500.00	1	$500(1.06)$	=	530.00
4	500.00	0	500	=	500.00
			Total	=	2187.31

We can derive an equation that computes the principal at the end of a period in a sinking fund as follows:

Let

R = periodic deposit

i = interest rate

n = number of deposits made (number of interest periods contained in the term of the fund)

S_n = principal in the sinking fund at the end of the n th period

To apply the method of chronological sequence, we must observe that the principal at the end of any period equals the principal at the end of the preceding period multiplied by the factor $(1 + i)$ and augmented by the periodic deposit R . That is,

$$S_r = S_{r-1}(1 + i) + R \quad (6.12)$$

where r has any integral value less than n . Hence, we obtain the following relations:

$$\begin{aligned} S_1 &= R \\ S_2 &= R(1 + i) + R \\ S_3 &= R(1 + i)^2 + R(1 + i) + R \\ &\dots \\ S_n &= R(1 + i)^{n-1} + R(1 + i)^{n-2} + \dots + R(1 + i) + R \end{aligned}$$

For convenience, we shall reverse the sequence of the terms, thereby obtaining

$$S_n = R + R(1 + i) + R(1 + i)^2 + \dots + R(1 + i)^{n-2} + R(1 + i)^{n-1} \quad (6.13)$$

We recognize the right side of this equation as a geometric series where R is the first term and the ratio of each term to the preceding term is $(1 + i)$. In Chapter 1, we derived Equation (1.61) which is repeated here for convenience.

$$S_n = a_1 \frac{1 - r^n}{1 - r} = a_1 \frac{r^n - 1}{r - 1} \quad (6.14)$$

Applying this equation to the sinking-fund principal, but substituting R for a_1 and $(1 + i)$ for r , we obtain

$$S_n = R \frac{(1 + i)^n - 1}{(1 + i) - 1}$$

or

$$S_n = R \frac{(1 + i)^n - 1}{i} \quad (6.15)$$

For the special case where the periodic deposit R is \$1, we shall use the notation $s_{\bar{n}}$ to denote the principal in the fund at the end of the n th period. Hence,

$$s_{\bar{n}} = \frac{(1 + i)^n - 1}{i} \quad (6.16)$$

Chapter 6 Mathematics of Finance and Economics

For the general case where the periodic deposit R has any value other than \$1, the principal S_n can be expressed in terms of $s_{\bar{n}}$ as

$$S_n = R s_{\bar{n}} \quad (6.17)$$

Equation (6.17) is useful when a math tables book such as *CRC Standard Mathematical Tables* that contains financial tables is available.

Example 6.17

A sinking fund consists of 15 annual deposits of \$1000 each, with interest earned at the rate of 4% compounded annually. What is the principal in the fund at its terminal date?

Solution:

For this example, $R = 1000$, $i = 0.04$, and $n = 15$. Using Equation (6.15), Page 6–27, we obtain

$$S_n = R \frac{(1+i)^n - 1}{i} = 1000 \times \frac{(1+0.04)^{15} - 1}{0.04} = 1000 \times 20.05359 = 20,023.59$$

Check with Microsoft Excel:

The function =FV(0.04,15,-1000) returns \$20,023.59.

In many problems, the principal in the sinking fund at its terminal date is the known quantity, and we must determine the periodic deposit required to accumulate this principal. Solving Equation (6.15) for R we obtain

$$R = S_n \frac{i}{(1+i)^n - 1} \quad (6.18)$$

Example 6.18

A corporation is establishing a sinking fund for the purpose of accumulating a sufficient capital to retire its outstanding bonds at maturity. The bonds are redeemable in 10 years, and their maturity value is \$150,000. How much should be deposited each year if the fund pays interest at the rate of 3%?

Solution:

For this example, $S_n = 150000$, $i = 0.03$, and $n = 10$. Using Equation (6.18), this page, we obtain

$$R = S_n \frac{i}{(1+i)^n - 1} = 150000 \times \frac{0.03}{(1+0.03)^{10} - 1} = 150000 \times 0.08723 = \$13,084.58$$

Check with Microsoft Excel:

The function =PMT(0.03,10,0,-150000) returns \$13,084.58.

6.4 Annuities

Annuity: A series of equal payments or receipts occurring over a specified number of periods.

Ordinary annuity: A series of equal payments or receipts occurring over a specified number of periods with the payments or receipts occurring at the end of each period.

Annuity due: A series of equal payments or receipts occurring over a specified number of periods with the payments or receipts occurring at the beginning of each period.

NOTE: While technically correct, the last two definitions shown above can be a somewhat confusing. Whether a cash flow appears to occur at the end or the beginning of a period often depends on our perspective. For example, the end of year 2 is also the beginning of year 3.

Therefore, the real key to distinguishing between an ordinary annuity and an annuity due is the point at which either a future or present value is to be calculated. Remembering the following characteristics should help us identify the type of annuity that we are dealing with:

1. For an ordinary annuity, future value is calculated as of the last cash flow, while present value is calculated as of one period before the first cash flow.
2. For an annuity due, future value is calculated as of one period after the last cash flow, while present value is calculated as of the first cash flow.

The equations used in the computations of sinking funds apply also to annuities.

Example 6.19

Suppose that on January 1, 2007 our bank account was \$6000, and withdrawals of \$500 each were made at the end of each year for 4 years. If the account earned 4% annual interest, what would be the balance in the account at the end of 2010 immediately after the last withdrawal is made?

Solution:

The computations are shown in Table 6.11.

TABLE 6.11 Computations for Example 6.19

Year	Principal at beginning	Interest Earned	End-of-year withdrawal	Principal at end of period
2007	\$6000.00	$6000.00 \times 0.04 = \$240.00$	\$500.00	$6000.00 + 240.00 - 500.00 = \5740.00
2008	5740.00	$5740.00 \times 0.04 = \$229.60$	\$500.00	$5740.00 + 229.60 - 500.00 = \5469.60
2009	5469.60	$5469.60 \times 0.04 = \$218.78$	\$500.00	$5469.60 + 218.78 - 500.00 = \5188.38
2010	5188.38	$5188.38 \times 0.04 = \$207.54$	\$500.00	$5188.38 + 207.54 - 500.00 = \4895.92

As shown in the table above, the principal at the end of 2010 would be \$4895.92.

In many instances, it is desirable to find the value of these n sums of money at the origin date of the annuity. The equation that will compute the amount at the origin date is derived as follows:

We recall from (6.7), Page 6–13, that

$$P_{-n} = P_0(1 + i)^{-n} \tag{6.19}$$

and from (6.15), Page 6–27, that

$$S_n = R \frac{(1 + i)^n - 1}{i} \tag{6.20}$$

As stated earlier, these equations apply to both sinking funds and annuities. For convenience, we will denote (6.19) as

$$PV_{\text{annuity}} = FV_{\text{annuity}}(1 + i)^{-n} \tag{6.21}$$

and (6.20) as

$$FV_{\text{annuity}} = \text{Payment} \frac{(1 + i)^n - 1}{i} \tag{6.22}$$

By substitution of (6.22) into (6.21) we obtain

$$PV_{\text{annuity}} = \text{Payment} \frac{(1 + i)^n - 1}{i} (1 + i)^{-n} = \text{Payment} \frac{1 - (1 + i)^{-n}}{i} \tag{6.23}$$

Example 6.20

Suppose that an annuity has been established that will enable us to receive \$300 at the end of each year for 5 years starting with December 31, 2006. If the interest rate is 6%, what amount can we expect to receive on January 1, 2012 should we decide to close this annuity?

Solution:

Equation (6.23), Page 6–30, is applicable for this example. Thus,

$$PV_{\text{annuity}} = 300 \frac{1 - (1 + 0.06)^{-5}}{0.06} = 300 \times 4.21236 = 1263.71$$

Example 6.21

A business firm must make a disbursement of \$800 at the end of each year from 2008 to 2011 inclusive. In order to meet the obligations as they fall due, it must deposit in a fund at the beginning of 2008 an amount of money that will be just sufficient to provide the necessary payments. If the interest rate is 3% compounded annually, what sum should be deposited?

Solution:

Let x denote the sum to be deposited in the fund on Jan. 1, 2008. Then,

$$x = \text{Value of deposit} = \text{Value of annuity}$$

This relationship holds for any valuation date we select; the most convenient date to use is obviously the beginning of 2008, which is both the date of the deposit and the origin date of the annuity. Then,

$$PV_{\text{annuity}} = 800 \frac{1 - (1 + 0.03)^{-3}}{0.03} = 800 \times 3.7171 = 2973.68$$

Example 6.22

A father wishes to establish a fund for his newborn son's college education. The fund is to pay \$2,000 on the 18th, 19th, 20th, and 21st birthdays of the son. The fund will be built up by the deposit of a fixed sum on the son's 1st to 17th birthdays, inclusive. If the fund earns 2%, what should the yearly deposit into the fund be?

Solution:

Let R denote the periodic deposit in the fund. Selecting the end of the 17th year as our valuation date, we have

$$PV_{\text{annuity}} = \text{Payment} \frac{1 - (1 + i)^{-n}}{i} = 2000 \frac{1 - (1 + 0.02)^{-4}}{0.02} = 2000 \times 3.80773 = \$7615.46$$

That is, at the end of the 17th year, the fund must have accumulated \$7615.46.

Chapter 6 Mathematics of Finance and Economics

Now, we must find the annual payments that must be deposited in order to accumulate this amount in 17 years. Solving (6.22) for Payment, we obtain

$$\text{Payment} = FV_{\text{annuity}} \frac{i}{(1+i)^n - 1} = 7615.46 \frac{0.02}{(1+0.02)^{17} - 1} = 7615.46 \times 0.04997 = \$380.54$$

Determining the Interest Rate of an Annuity

In many problems pertaining to annuities that arise in practice, the known quantities are the value of the annuity at either the origin or terminal date and the amount of the periodic payment, while the interest rate by which the given quantities are related is unknown. This is often true, for example, where an asset is purchased on the installment plan. The purchaser knows the purchase price of the asset and the periodic payment he is obligated to make, but he is not directly aware of the interest rate implicit in the loan. Only an approximate solution of this type of problem is possible, and a solution by straight-line interpolation yields results of a sufficient degree of accuracy.

Example 6.23

An asset costing \$5,000 was purchased on the installment plan, the terms of sale requiring that the buyer make a down payment of \$2,000 and five annual payments of \$720 each. The first of these periodic payments is to be made 1 year subsequent to the date of purchase. What is the interest rate pertaining to this loan?

Solution:

From (6.23), Page 6–30,

$$PV_{\text{annuity}} = \text{Payment} \frac{1 - (1+i)^{-n}}{i}$$

Assuming that PV_{annuity} is the value at the origin date, this value is $5000 - 2000 = 3000$ and the Payment is \$720. Substitution of these values into the above relation and rearranging yields

$$\frac{1 - (1+i)^{-n}}{i} = \frac{3000}{720}$$

Obviously, it is a formidable task to solve this equation for the interest rate i . Instead, we construct the following spreadsheet where we observe that the interest rate corresponding to the value at the origin is between 6% and 7%.

i	R	n	Value at origin date
1%	720	5	3494.47
2%	720	5	3393.69
3%	720	5	3297.39
4%	720	5	3205.31
5%	720	5	3117.22
6%	720	5	3032.90
7%	720	5	2952.14
8%	720	5	2874.75
9%	720	5	2800.55
10%	720	5	2729.37

We will use the graph of Figure 6.5 below to apply straight-line interpolation.

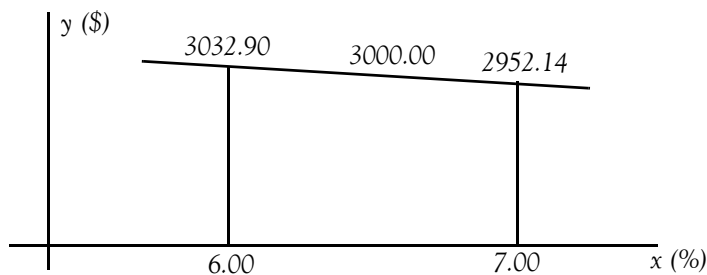


Figure 6.5. Graph for Example 6.23

$$x_i = \frac{1}{\text{slope}} \times (y_i - y_1) + x_1 = \frac{1}{m} \times (y_i - y_1) + x_1$$

For our example,

$$\begin{aligned} x_{y=3000.00} &= \frac{1}{\text{slope}} \times (3000.00 - 3032.90) + 0.06 = \frac{1}{\frac{2952.14 - 3032.90}{0.07 - 0.06}} \times (-32.90) + 0.06 \\ &= \frac{0.01}{-80.76} \times (-32.90) + 0.06 = 0.06407 \end{aligned}$$

Therefore, the interest rate is 6.4%.

Check with the Microsoft Excel IRR function:

In a blank spreadsheet we enter the values 3000 (5000–2000), –720, –720, –720, –720, and –720 in A1 through A6 respectively, and we use the formula =IRR(A1:A6) and Excel returns 6.402%.

6.5 Amortization

Amortization is the liquidation (a debt, such as a mortgage) by installment payments or payment into a sinking fund. In the course of doing business, we will likely acquire what are known as intangible assets. These assets can contribute to the revenue growth of your business and, as such, they can be expensed against these future revenues. An example of an intangible asset is when you buy a patent for an invention.

Calculating amortization

The formula for calculating the amortization on an intangible asset is similar to the one used for calculating straight-line depreciation*. We divide the initial cost of the intangible asset by the estimated useful life of the intangible asset. The amortization per year is computed with the formula

$$\text{Amortization per year} = \frac{\text{Initial cost}}{\text{Useful life}} \quad (6.24)$$

Example 6.24

Suppose that it costs \$10,000 to acquire a patent and it has an estimated useful life of ten years. The amortization per year is

$$\text{Amortization per year} = \frac{\text{Initial cost}}{\text{Useful life}} = \frac{\$10000}{10} = \$1000$$

The amount of amortization accumulated since the asset was acquired appears on the balance sheet as a deduction under the amortized asset.

The formula of (6.24) above assumes that the interest rate is zero. However, when a debt such as a mortgage is amortized over a number of years, there is always an interest rate associated with that debt. In this case, the following formula is used.

$$\text{Payment} = \text{Principal} \times \frac{\text{Interest}}{1 - (1 + \text{Interest})^{-n}} \quad (6.25)$$

Amortization tables can be easily constructed with spreadsheets such as that shown on Figure 6.6. For this spreadsheet we have assumed an annual interest rate of 7.5% and thus the monthly interest rate is computed and shown in cell C8 as =0.075/12.

* Depreciation is discussed in Chapter 7.

30 YEAR MORTGAGE LOAN					
Inputs					
Terms of Loan					
	Principal	\$500,000			
	Term	360			
	Interest (Month)	0.625%			
	Payment	\$3,496.07			
Loan Payment Schedule					
Payment	Regular Payment	Interest	Principle	Extra	Principal
Period	Year 2007	Expense	Payment	Principal Paid	Balance
1	\$3,496.07	\$3,125.00	\$371.07	\$0.00	\$499,628.93
2	\$3,496.07	\$3,122.68	\$373.39	\$0.00	\$499,255.54

Figure 6.6. 30-Year Amortization Table

With Microsoft Excel, in all arguments, cash we pay out such as deposits to savings, is represented by negative numbers; cash we receive, such as dividend checks, is represented by positive numbers.

Other formulas in other cells are as follows:

C9: =PMT(C8,C7,-C6)

B16: =C9

C16: =C6*C8

D16: =B16-C16

E16: =IF(AND(\$F\$7<=A16,\$F\$8>=A16),MIN(\$F\$6,C6-D16),0)

F16: =C6-D16-E16

B17: =MIN(\$C\$9,F16+C17)

C17: =F16*\$C\$8

D17: =B17-C17

E17: =IF(AND(\$F\$7<=A17,\$F\$8>=A17),MIN(\$F\$6,F16-D17),
IF(AND(\$F\$10<=A17,\$F\$11>=A17),MIN(\$F\$9,F16-D17),0))

F17: =F16-D17-E17

The remaining entries in columns B though F are obtained by copying B17:F17 to B18:F549.

Example 6.25

A \$10,000 second mortgage is being amortized by means of 20 equal yearly installments at an interest rate of 6%. The agreement provides for paying off the mortgage in a lump sum at any time with an amount equal to the unpaid balance plus a charge of 1% of the unpaid balance. What would have to be paid to discharge the mortgage after 10 payments have been made?

Solution:

Each yearly installment is found by

$$\text{Payment} = \text{Principal} \times \frac{\text{Interest}}{1 - (1 + \text{Interest})^{-n}}$$

For our example,

$$\text{Payment} = 10000 \times \frac{0.06}{1 - (1 + 0.06)^{-20}} = 10000 \times \frac{0.06}{1 - 0.3118} = 871.85$$

To find the unpaid balance at the end of 10 years, we construct the spreadsheet of Figure 6.7.

We can also use the equation

$$\text{Unpaid balance} = \text{Payment} \times \frac{1 - (1 + i)^{-(n-r)}}{i}$$

where

n = specified period of the loan

r = actual period when the loan is paid in full

For this example,

$$\text{Unpaid balance} = 871.85 \times \frac{1 - (1 + 0.06)^{-(20-10)}}{0.06} = 6486.89$$

6.6 Perpetuities

As stated earlier, a perpetuity is a constant stream of identical cash flows with no maturity date. In other words, a perpetuity is an annuity that continues indefinitely. To provide for a perpetuity, it is necessary to deposit in a fund a sum of money of such magnitude that the interest earned between successive payments exactly equals the amount of the periodic payment, thereby maintaining the original principal permanently intact. Thus,

$$\text{Amount deposited} = \frac{\text{Earned interest}}{\text{Interest rate}} \quad (6.26)$$

20 YEAR MORTGAGE LOAN PAID IN 10 YEARS

Inputs

Terms of Loan		Extra Payments	
Principal	\$10,000	Extra Amount	\$0.00
Term	20	Start Period	\$0.00
Interest (Month)	6.00%	End Period	\$0.00
Payment	\$871.85		

Loan Payment Schedule

Payment Period	Regular Payment	Interest Expense	Principle Payment	Extra Principal Paid	Principal Balance
1	\$871.85	\$600.00	\$271.85	\$0.00	\$9,728.15
2	\$871.85	\$583.69	\$288.16	\$0.00	\$9,440.00
3	\$871.85	\$566.40	\$305.45	\$0.00	\$9,134.55
4	\$871.85	\$548.07	\$323.77	\$0.00	\$8,810.78
5	\$871.85	\$528.65	\$343.20	\$0.00	\$8,467.58
6	\$871.85	\$508.05	\$363.79	\$0.00	\$8,103.79
7	\$871.85	\$486.23	\$385.62	\$0.00	\$7,718.17
8	\$871.85	\$463.09	\$408.76	\$0.00	\$7,309.42
9	\$871.85	\$438.57	\$433.28	\$0.00	\$6,876.14
10	\$871.85	\$412.57	\$459.28	\$0.00	\$6,416.86

Figure 6.7. Spreadsheet for Example 6.25

Example 6.26

An individual wishes to establish a scholarship for a university of \$5,000 annually and that the fund in which the money is to be held earns 4% interest per annum. To earn \$5,000 interest annually, the sum to be deposited is

$$\text{Amount deposited} = \frac{5000}{0.04} = \$125000$$

Of course, this sum must be deposited 1 year prior to the date of the first payment.

In the above example, the payment period coincided with the interest period, but this may not be always the case. In general, let

R = periodic payment of a perpetuity

Chapter 6 Mathematics of Finance and Economics

i = interest rate

m = number of interest periods contained in one payment period

M = principal to be deposited in the fund one payment period prior to the date of the first payment

Equating the periodic payment R to the interest earned by M in one payment period ($= m$ interest periods), we obtain

$$M(1+i)^m - M = R$$

$$M[(1+i)^m - 1] = R$$

or

$$M = \frac{R}{(1+i)^m - 1} \quad (6.27)$$

Example 6.27

What amount must be donated to build an institution having an initial cost of \$500,000, provide an annual upkeep of \$50,000, and have \$500,000 at the end of each 50-year period to rebuild the institution. Assume that invested funds return 4% earned interest.

Solution:

The annual upkeep of \$50,000 can be regarded as a single disbursement made at the end of each year. The required payments can be grouped in the following manner:

1. An initial payment of \$500,000 to construct the building.
2. A perpetuity consisting of payments of \$50,000 each made at 50-year intervals.
3. A perpetuity consisting of annual payments of \$50,000 each for maintenance.

Then,

Payments = Initial payment + Perpetuity at 50-year intervals + Perpetuity for maintenance

or

$$\begin{aligned} M &= 500000 + \frac{500000}{(1+0.04)^{50} - 1} + \frac{50000}{0.04} = 500000 + \frac{500000}{6.107} + \frac{50000}{0.04} \\ &= 500000 + 81873.26 + 1250000 = 1,831,873.26 \end{aligned}$$

6.7 Valuation of Bonds

Generally, the price at which a bond is purchased is not identical with the face value of the bond. When the purchase price of a bond does coincide with its par value, it is said to be purchased at par; when the purchase price exceeds the par value it is said to be purchased above par, or at a premium; finally, when the purchase price is below the par value, it is said to be purchased below par, or at a discount.

The difference between the par value of a bond and its purchase price is termed the premium or discount, whichever applies. For example, if the face value of a bond is \$1000 and it sells for \$940, it has been sold at a discount of \$60.00; if the same bond sells for \$1020, then it has been sold at a premium of \$20. We shall hereafter refer to the periodic interest payments received by the bondholder as the *dividends* and to the interval between successive dividends as the *dividend period*.

We will assume that the economic conditions remain unchanged. Accordingly, we will assume that both the price and the prevailing interest rate remain unaltered during the entire period. Therefore, if an investor has purchased a bond at its date of issue at a price to yield a return of 6% on his investment, it follows that at any date prior to the maturity of the bond he can transfer the bond to a new purchaser at such price as to maintain a 6% return on his original investment and to yield the new investor the same rate of return. The value of a bond at any date intermediate between its issue and its redemption is termed its *book value*.

Example 6.28

Assume that a bond of \$1000 face value, redeemable at par, and earning interest at the rate of 6% annually is purchased by an investor for \$980 at a date 1 year prior to its maturity. What is the rate of return at the end of the year?

Solution:

At the end of the year the investor receives the following sums of money:

Dividend 6% of \$1000	=	\$60.00
Redemption price (face value)	=	1000.00
Total income at end of year	=	1060.00
Investment at beginning of year	=	980.00
Earning for year	=	80.00
Rate of return on investment	=	$80/980 = 8.16\%$

6.7.1 Calculating the Purchase Price of a Bond

We will now present an example to find out the price we should pay for a bond to yield a stipulated investment rate. We will assume that the date of purchase coincides with the beginning of a dividend period.

Example 6.29

A \$1000 bond, redeemable at par in 5 years, pays dividends at the rate of 5% per year. Compute the purchase price of the bond and the book value at the beginning of each dividend period if the bond is purchased to yield a rate of return of

- a. 5%
- b. 8%
- c. 3%

Solution:

- a. Since the investment rate equals the dividend rate, the purchase price of the bond at its date of issue equals its par value of \$1000. At date of issue the bond is worth \$1000 and during the first year, the bond earns interest at the rate of 5% and therefore attains the value of \$1050 at the end of that date. At this date, however, a dividend of \$50 is received by the bondholder restoring the bond to its original value of \$1000. This process is repeated every period. The result is that at the beginning of each dividend period the value of the bond has reverted to its par value.
- b. The sums of money which the bondholder is to receive are the following:
 - 1. The series of five annual dividends, each of which is 5% of the face value, or \$50. These dividends constitute an annuity whose value at date of issue of the bond, based on a rate of 8%, is given by Equation (6.23), Page 6–30, as

$$PV_{\text{annuity}} = \text{Payment} \frac{1 - (1 + i)^{-n}}{i} = 50 \times \frac{1 - (1 + 0.08)^{-5}}{0.08} = 50 \times 3.99271 = 199.64$$

- 2. The redemption value of the bond, \$1000, which will be received 5 years after the date of purchase. The value of this sum at date of purchase is by Equation (6.19)

$$P_{-n} = P_0(1 + i)^{-n} = 1000(1 + 0.08)^{-5} = 1000 \times 0.68058 = 680.58$$

Then, the purchase price of the bond is

$$199.64 + 680.58 = 880.22$$

Hence, the bond is purchased at a discount of

$$1000.00 - 880.22 = 119.78$$

The book value of the bond at the commencement of each subsequent year is computed by repeating the above steps. The results are shown in Table 6.12.

TABLE 6.12 Computations for Part (b), Example 6.29

Beginning of year	Value of future dividends	Redemption price	Book value of bond
1	$50 \times \frac{1 - (1.08)^{-5}}{0.08} = \199.64	$1000(1.08)^{-5} = \$680.58$	$199.64 + 680.58 = \$880.22$
2	$50 \times \frac{1 - (1.08)^{-4}}{0.08} = 165.61$	$1000(1.08)^{-4} = 735.03$	$165.61 + 735.03 = 900.64$
3	$50 \times \frac{1 - (1.08)^{-3}}{0.08} = 128.85$	$1000(1.08)^{-3} = 793.83$	$128.85 + 793.83 = 922.68$
4	$50 \times \frac{1 - (1.08)^{-2}}{0.08} = 89.16$	$1000(1.08)^{-2} = 857.34$	$89.16 + 857.34 = 946.50$
5	$50 \times \frac{1 - (1.08)^{-1}}{0.08} = 46.30$	$1000(1.08)^{-1} = 925.93$	$46.30 + 925.93 = 972.23$
End of 5th yr	0.00	1000.00	1000.00

- c. We follow the same method of solution as in Part (b) with all computations based on an interest rate of 3%. The computations are shown in Table 6.13.

The first line in Table 6.13 indicates that the purchase price is

$$\text{Purchase price} = 50 \times \frac{1 - (1 + 0.03)^{-5}}{0.08} + 1000(1 + 0.03)^{-5} = 228.99 + 862.61 = \$1091.60$$

Therefore, the bond is purchased at a premium of $1091.60 - 1000.00 = \$91.60$.

TABLE 6.13 Computations for Part (c), Example 6.29

Beginning of year	Value of future dividends	Redemption price	Book value of bond
1	$50 \times \frac{1 - (1.03)^{-5}}{0.03} = \228.99	$1000(1.03)^{-5} = \$862.61$	$228.99 + 862.61 = \$1091.60$
2	$50 \times \frac{1 - (1.03)^{-4}}{0.03} = 185.85$	$1000(1.03)^{-4} = 888.49$	$185.85 + 888.49 = 1074.34$
3	$50 \times \frac{1 - (1.03)^{-3}}{0.03} = 141.43$	$1000(1.03)^{-3} = 915.14$	$141.43 + 915.14 = 1056.57$
4	$50 \times \frac{1 - (1.03)^{-2}}{0.03} = 95.67$	$1000(1.03)^{-2} = 942.60$	$95.67 + 942.60 = 1038.27$
5	$50 \times \frac{1 - (1.03)^{-1}}{0.03} = 48.54$	$1000(1.03)^{-1} = 970.87$	$48.54 + 970.87 = 1019.41$
End of 5th yr	0.00	1000.00	1000.00

6.7.2 Total Periodic Bond Disbursement

To accumulate the funds necessary to retire an issue of bonds at its maturity, the seller must make periodic deposits in a sinking fund. Let us assume that the deposit period of the fund coinciding with the dividend period of the bonds. Also let

F = total face value of the bonds

M = total redemption value of the bonds

d = dividend rate of the bonds

i = interest rate of sinking fund

The total disbursement pertaining to the bonds which the seller must make at the end of each period consists of the following:

1. The periodic dividend D payable to the bondholders. This amount is computed as

$$D = Fd \tag{6.28}$$

2. The periodic contribution R to the sinking fund

$$R = M \frac{i}{(1+i)^n - 1} \tag{6.29}$$

Therefore, the total periodic disbursements (annual costs) are

$$\text{Periodic disbursements} = D + R = Fd + M \frac{i}{(1+i)^n - 1} \quad (6.30)$$

Example 6.30

A community wishes to purchase an existing utility valued at \$5,000,000 by selling 5% bonds that will mature in 30 years. The money to retire the bonds will be raised by paying equal annual amounts into a sinking fund that will earn 4% per cent. The bonds will be sold at par. What will be the total annual cost of the bonds until they mature?

Solution:

For this example $F = M = 5,000,000$, $d = 0.05$, $i = 0.04$, and $n = 30$ years. By substitution into (6.30) above we obtain

$$\begin{aligned} \text{Annual disbursements} &= 5,000,000 \times 0.05 + 5,000,000 \times \frac{0.04}{(1+0.04)^{30} - 1} \\ &= 250,000 + 89,150.50 = 339,150.50 \end{aligned}$$

Example 6.31

A municipality wishes to raise funds for improvements by issuing 5.5% bonds. There is \$200,000 available per year for interest payments and retirement of the bonds at 10% above face value. The interest rate of the sinking fund is 3%. What should be the amount of the bond issue if all the bonds are to be retired in 20 years?

Solution:

For this example $M = 1.10F$, $d = 0.055$, $i = 0.03$, and $n = 20$ years. By substitution into (6.30) above we obtain

$$\begin{aligned} \text{Annual disbursements} = 200,000 &= 0.055F + 1.10F \frac{0.03}{(1+0.03)^{20} - 1} \\ &= 0.055F + 1.10F(0.03722) \end{aligned}$$

or

$$\begin{aligned} 0.055F + 0.041F &= 200000 \\ 0.096F &= 200000 \end{aligned}$$

and solving for F we obtain $F = \$2,083,333.33$.

6.7.3 Calculation of Interest Rate of Bond

In the previous examples we calculated the purchase price of a bond corresponding to a particular interest rate. In many cases, however, the purchase price of a bond is the known quantity, and it is necessary to determine the investment rate secured by the purchaser. A calculation of this nature is also required to enable the issuing corporation to determine the true rate of interest it is paying for borrowed money. The basis for calculating this interest rate is not the actual selling price of the bonds but rather the sum of money remaining after the payment of administrative and legal expenses.

Assuming there are no bond tables available for obtaining the interest rate corresponding to a given purchase price, we shall calculate the approximate rate using straight-line interpolation.

Example 6.32

Williams paid \$11,000 for a \$10,000 bond that pays \$400 per year. In 20 years the bond will be redeemed for \$10,500. What net rate of interest will Williams receive on his investment?

Solution:

Let i denote the interest rate. Evaluating all sums at the maturity date, we obtain

$$11000(1+i)^{20} = 400 \frac{(1+i)^{20} - 1}{i} + 10500$$

To obtain a first approximation, we substitute $1 + ni$ for the expression $(1+i)^n$. Then,

$$11000(1+20i) = 400 + 10500 = 400 \times 20 + 10500$$

Solving for i we obtain as a first approximation

$$\begin{aligned} 220000i &= 18500 - 11000 \\ i &= \frac{7500}{220000} = 0.0341 = 3.41\% \end{aligned}$$

Since we have understated the value both of the \$11,000 investment and of the periodic dividends, it is reasonable to presume that this value does not deviate markedly from the true value. Therefore, let us compute at 3.0% and at 3.5% and perform straight line approximation as before.

At 3.0% we have

$$11000(1.03)^{20} = 11000 \times 1.80611 = 19,867.20$$

$$400 \frac{(1 + 0.03)^{20} - 1}{0.03} = 400 \times 26.87037 = 10,748.20$$

Redemption price for 3.0% rate = 19,867.20 – 10,748.20 = \$9,119.00

At 3.5% we have

$$11000(1.035)^{20} = 11000 \times 1.98979 = 21,887.70$$

$$400 \frac{(1 + 0.035)^{20} - 1}{0.035} = 400 \times 28.27968 = 11,319.90$$

Redemption price for 3.5% rate = 21,887.70 – 11,319.90 = \$10,567.80

Applying straight-line interpolation, we obtain

$$x_i = \frac{1}{\text{slope}} \times (y_i - y_1) + x_1 = \frac{1}{m} \times (y_i - y_1) + x_1$$

For our example,

$$\begin{aligned} x|_{y=10,500} &= \frac{1}{\text{slope}} \times (10,500.00 - 9,119.00) + 0.03 = \frac{1}{\frac{10,578.80 - 9,119.00}{0.035 - 0.030}} \times 1,381.00 + 0.03 \\ &= \frac{0.005}{1,456.80} \times 1,381.00 + 0.03 = 0.03474 \end{aligned}$$

or 3.47% approximately.

6.8 Spreadsheet Financial Functions

Spreadsheets such as Microsoft Excel, Lotus 1–2–3, and Quattro Pro have built-in financial functions that we can use to obtain quick answers. To appreciate the usefulness of these functions, we present the following example which we will solve by classical methods.

Microsoft Excel solves for one financial argument in terms of the others. If the argument **rate** is not zero, then,

$$pv \times (1 + \text{rate})^{\text{nper}} + \text{pmt}(1 + \text{rate} \times \text{type}) \times \left(\frac{(1 + \text{rate})^{\text{nper}} - 1}{\text{rate}} \right) + \text{fv} = 0 \quad (6.31)$$

provided that rate ≠ 0

If rate = 0, the following formula is used.

$$(\text{pmt} \times \text{nper}) + \text{pv} + \text{fv} = 0 \quad (6.32)$$

In equations (6.31) and (6.32) the arguments are defined as follows:

pv = present value or the lump sum amount that a series of future payments is worth at the present.

rate = the interest rate per period. For example, if we obtain an automobile loan at a 10% annual interest rate and make monthly payments, our interest rate per month is $10\%/12$, or 0.83%. Accordingly, we must enter $10\%/12$, or 0.83%, or 0.0083, into the formula as the rate.

nper = the total number of payment periods in an annuity. For example, if we obtain a four-year car loan and make monthly payments, our loan has $4 \times 12 = 48$ periods. We must enter 48 into the formula for **nper**.

pmt = the payment made each period and cannot change over the life of the annuity. Typically, **pmt** includes principal and interest but no other fees or taxes. For example, the monthly payments on a \$10,000, 4-year car loan at 12% are \$263.33. We must enter -263.33 into the formula as the **pmt**.

type = the number 0 or 1 which indicates when payments are due. If 0, the payments are due at the end of the period. If 1, the payments are due at the beginning of the period.

fv = the future value, or a cash balance we want to attain after the last payment is made. If **fv** is omitted, it is assumed to be 0 (the future value of a loan, for example, is 0). For example, if we want to save \$50,000 to pay for a special project in 18 years, then \$50,000 is the future value. We could then make a conservative guess at an interest rate and determine how much we must save each month.

The Microsoft Excel financial functions are presented in Subsections 6.8.1 through 6.8.11 below.

6.8.1 PV Function

The PV function returns the present value of an investment. The present value is the total amount that a series of future payments is worth now. For example, when we borrow money, the loan amount is the present value to the lender.

This function also computes the present value of an ordinary annuity. An *ordinary annuity* is a series of payments made at equally spaced intervals.

The present value **pv** of an annuity allows us to compare different investment alternatives or potential obligations.

The formula for computing the present value of an ordinary annuity is

$$pv = \text{Payment} \times \frac{1 - (1 + \text{interest})^{-n}}{\text{interest}} \quad (6.33)$$

The syntax for Excel is

=PV(rate,nper,pmt,fv,type) where *

rate = the interest rate per period. For example, if we obtain an automobile loan at a 10% annual interest rate and make monthly payments, our interest rate per month is $10\%/12$, or 0.83%. Accordingly, we must enter $10\%/12$, or 0.83%, or 0.0083, into the formula as the rate.

nper = the total number of payment periods in an annuity. For example, if we obtain a four-year car loan and make monthly payments, our loan has $4 \times 12 = 48$ periods. We must enter 48 into the formula for **nper**.

pmt = the payment made each period and cannot change over the life of the annuity. Typically, **pmt** includes principal and interest but no other fees or taxes. For example, the monthly payments on a \$10,000, 4-year car loan at 12% are \$263.33. We must enter -263.33 into the formula as the **pmt**. If **pmt** is omitted, we must include the **fv** argument.

fv = the future value, or a cash balance we want to attain after the last payment is made. If **fv** is omitted, it is assumed to be 0 (the future value of a loan, for example, is 0). For example, if we want to save \$50,000 to pay for a special project in 18 years, then \$50,000 is the future value. We could then make a conservative guess at an interest rate and determine how much we must save each month. If **fv** is omitted, we must include the **pmt** argument.

type = the number 0 or 1 which indicates when payments are due. If 0, the payments are due at the end of the period. If 1, the payments are due at the beginning of the period. If omitted, it is assumed to be 0.

Remarks:

1. We must make sure that we are consistent with the units that we use for specifying **rate** and **nper**. For example, if we make monthly payments on a four-year loan at 11% annual interest rate, we must use $11\%/12$ for **rate** and 4×12 for **nper**. If we make annual payments on the same loan, we must use 11% for **rate** and 4 for **nper**.
2. As stated earlier, for all the arguments, cash we pay out, such as deposits to savings, is represented by negative numbers; cash we receive, such as dividend checks, is represented by positive numbers.

* These terms were defined on the previous page but are repeated here for convenience.

Example 6.33

We wish to receive \$12,000 at the end of each year for the next 20 years. If the interest rate is 7%, what amount must we deposit now to achieve this goal?

Solution:

We use the Excel PV function =PV(0.07,20,12000,0,0) and this returns -\$127,128.17. This result is negative because it represents the amount that we must pay. In other words, this is an outgoing cash flow.

6.8.2 FV Function

The FV function computes the future value of an ordinary annuity. An ordinary annuity is a series of payments made at equally spaced intervals.

The future value of an annuity allows us to compare different investment alternatives or potential obligations.

The formula for computing the future value of an ordinary annuity is

$$FV = \text{Payment} \times \frac{(1 + \text{interest})^n - 1}{\text{interest}} \quad (6.34)$$

The syntax for Excel is

=FV(rate,nper,pmt,pv,type) where

rate = interest rate per period

nper = the total number of payment periods in an annuity

pmt = payment made each period; it cannot change over the life of the annuity. Typically, **pmt** contains principle and interest but no other fees or taxes. If **pmt** is omitted, we must include the **pv** argument.

pv = present value or the lump sum amount that a series of future payments is worth at the present. If **pv** is omitted, it is assumed to be 0 (zero) and we must include the **pmt** argument.

type = the number 0 or 1 which indicates when payments are due. If 0, the payments are due at the end of the period. If 1, the payments are due at the beginning of the period. If omitted, it is assumed to be 0.

Remarks:

1. The payment returned by PMT includes principal and interest but no taxes, reserve payments, or fees sometimes associated with loans.

2. We must make sure that we are consistent with the units that we use for specifying **rate** and **nper**. For example, if we make monthly payments on a four-year loan at 11% annual interest rate, we must use 11%/12 for **rate** and 4 × 12 for **nper**. If we make annual payments on the same loan, we must use 11% for **rate** and 4 for **nper**.

Example 6.34

Suppose we deposit \$2,000 each year for the next 20 years at 7.5% interest rate and we make each year's contribution on the last day of the year. We want to find out the amount accumulated at the end of the 20-year period.

Solution:

We use the Excel FV function =FV(0.075,20,-2000,0,0) and this returns \$86,609.36. This is the amount that we have accumulated over the 20-year period.

6.8.3 PMT Function

The PMT function calculates the payment for a loan based on constant payments and a constant interest rate.

The formula for computing the payment is

$$PMT = \text{Principal} \times \frac{\text{interest}}{1 - (1 + \text{interest})^{-n}} \tag{6.35}$$

The syntax for Excel is

=PMT(rate,nper,pv,fv,type) where

rate = interest rate for the loan per period

nper = the total number of payments for the loan

pv = present value or the total amount that a series of future payments is worth at the present. It is also referred to as the principal.

fv = future value or the cash balance that we want to attain after the last payment is made. If **fv** is omitted, it is assumed to be 0 (zero), that is, the future value of a loan is zero.

type = the number 0 or 1 which indicates when payments are due. If 0, the payments are due at the end of the period. If 1, the payments are due at the beginning of the period. If omitted, it is assumed to be 0.

Remarks:

1. The payment returned by PMT includes principal and interest but no taxes, reserve payments, or fees sometimes associated with loans.
2. We must make sure that we are consistent with the units that we use for specifying **rate** and **nper**. For example, if we make monthly payments on a four-year loan at 11% annual interest rate, we must use 11%/12 for **rate** and 4×12 for **nper**. If we make annual payments on the same loan, we must use 11% for **rate** and 4 for **nper**.

Example 6.35

Suppose we buy an automobile and we finance a loan of \$15,000 for the next 5 years at 8% interest rate. We want to find out the monthly payment for this loan.

Solution:

We use the Excel PMT function `=PMT(0.08/12,5*12,15000,0,0)` and this returns -304.15 . This represents our monthly payment. This amount is shown as a negative number since it represents an outgoing cash flow. The interest rate is divided by 12 to obtain the monthly payment, and the years that the money is paid out is multiplied by 12 to obtain the number of payments.

6.8.4 RATE Function

The **RATE** function returns the interest rate per period of an annuity. We can use **RATE** to determine the yield of a zero-coupon bond that is sold at a discount of its face value. This function is also useful in forecasting applications to calculate the compound growth rate between current and projected revenues and earnings.

RATE is calculated by iteration and can have zero or more solutions. If the successive results of **RATE** do not converge to within 0.0000001 after 20 iterations, **RATE** returns the **#NUM!** error value.

The syntax for Excel is

`=RATE(nper,pmt,pv,fv,type,guess)` where

nper = the total number of payment periods in an annuity.

pmt = the payment made each period and cannot change over the life of the annuity. Typically, **pmt** includes principal and interest but no other fees or taxes. If **pmt** is omitted, we must include the **fv** argument.

pv = present value or the total amount that a series of future payments is worth at the present.

fv = future value or the cash balance that we want to attain after the last payment is made. If **fv** is omitted, it is assumed to be 0 (zero), that is, the future value of a loan is zero.

type = the number 0 or 1 which indicates when payments are due. If 0, the payments are due at the end of the period. If 1, the payments are due at the beginning of the period. If omitted, it is assumed to be 0.

guess = our guess for what the rate will be. If we omit **guess**, it is assumed to be 10%. If **RATE** does not converge, we can try different values for **guess**. **RATE** usually converges if **guess** is between 0 and 1.

Remark:

We must make sure that we are consistent with the units that we use for specifying **rate** and **nper**. For example, if we make monthly payments on a four-year loan at 11% annual interest rate, we must use 11%/12 for **rate** and 4×12 for **nper**. If we make annual payments on the same loan, we must use 11% for **rate** and 4 for **nper**.

Example 6.36

Suppose that for \$3,500 we can purchase a zero-coupon bond with \$10,000 face value maturing in 10 years. We want to determine the annual interest rate that this bond will yield.

Solution:

We use the Excel **RATE** function `=RATE(10,0,-3500,10000,0,0.1)` and this returns 11.07%. This represents the annual interest rate that this bond will yield.

6.8.5 NPER* Function

The **NPER** function returns the number of periods for an investment based on periodic, constant payments, and a constant interest rate.

The syntax for Excel is

`=NPER(rate,pmt,pv,fv,type)` where

rate = the interest rate per period.

pmt = the payment made each period and cannot change over the life of the annuity. Typically, **pmt** includes principal and interest but no other fees or taxes.

* This function is also discussed in Chapter 1.

Chapter 6 Mathematics of Finance and Economics

pv = present value or the total amount that a series of future payments is worth at the present.

fv = future value or the cash balance that we want to attain after the last payment is made. If **fv** is omitted, it is assumed to be 0 (zero), that is, the future value of a loan is zero.

type = the number 0 or 1 which indicates when payments are due. If 0, the payments are due at the end of the period. If 1, the payments are due at the beginning of the period. If omitted, it is assumed to be 0.

Example 6.37

Suppose that we deposit \$2,000 at the end of each year into a savings account. This account earns 7.5% per year compounded annually. How long will it take to accumulate \$100,000?

Solution:

We use the Excel NPER function =NPER(0.075,-2000,0,100000) and this returns 21.54 years. This represents the time it will take to accumulate \$100,000.

6.8.6 NPV Function

The NPV function calculates the net present value of an investment by using a discount rate and a series of future payments (negative values) and income (positive values).

The syntax for Excel is

=NPV(rate,value1,value2,...) where

rate = the rate of discount over the length of one period.

value1, value2,... = 1 to 29 arguments representing the payments and income and must be equally spaced in time and occur at the end of each period.

Remarks:

1. NPV uses the order of **value1, value2,...** to interpret the order of cash flows. We must enter our payment and income values in the correct sequence.
2. Arguments that are numbers, empty cells, logical values, or text representations of numbers are counted; arguments that are error values or text that cannot be translated into numbers are ignored.
3. If an argument is an array or reference, only numbers in that array or reference are counted. Empty cells, logical values, text, or error values in the array or reference are ignored.

4. The NPV investment begins one period before the date of the valued cash flow and ends with the last cash flow in the list. The NPV calculation is based on future cash flows. If our first cash flow occurs at the beginning of the first period, the first value must be added to the NPV result, not included in the values arguments. For more information, see the examples below.

5. The formula for NPV is

$$NPV = \sum_{i=1}^n \frac{\text{value}_i}{(1 + \text{rate})^i} \quad (6.36)$$

6. NPV is similar to the PV function (present value). The primary difference between PV and NPV is that PV allows cash flows to begin either at the end or at the beginning of the period whereas in NPV the arguments value1, value2,... must occur at the end of each period as stated above. Also, unlike the variable NPV cash flow values, PV cash flows must be constant throughout the investment. For information about annuities and financial functions, see PV.

7. NPV is also related to the IRR function (Internal Rate of Return) which is discussed below. IRR is the rate for which NPV equals zero, that is,

$$NPV(\text{IRR}(\dots), \dots) = 0 \quad (6.37)$$

Example 6.38

Suppose that in a year we make 12 irregular distributions shown as value1, value2,...,value12 on the spreadsheet of Figure 6.8 at 11.5% annual interest rate, and we want to compute the net present value of those distributions.

	A	B
1	Distributions	Amount
2	value1	0
3	value2	0
4	value3	2500
5	value4	2500
6	value5	3000
7	value6	5000
8	value7	6000
9	value8	9000
10	value9	3000
11	value10	2500
12	value11	0
13	value12	7500
14	Annual Interest rate	11.5%
15	=NPV(B15/12,B2:B13)	\$38,084.13

Figure 6.8. Spreadsheet for Example 6.38

Solution:

We use the formula `=NPV(B15/12,B2:B13)` which returns \$38,084.13. This value represents the net present value of these distributions. As expected, this amount is less than the sum of \$41,000 of the 12 irregular monthly distributions.

6.8.7 IIR Function

The IIR function returns the internal rate of return for a series of cash flows represented by the numbers in values. These cash flows need not be even as they would be for an annuity. However, the cash flows must occur at regular intervals, such as monthly or annually. The internal rate of return is the interest rate received for an investment consisting of payments (negative values) and income (positive values) that occur at regular periods.

The syntax for Excel is

`=IRR(values,guess)` where

values = an array or a reference to cells that contain numbers for which we want to calculate the internal rate of return. This argument must contain at least one positive value and one negative value to calculate the internal rate of return. IRR uses the order of values to interpret the order of cash flows. We must enter our payment and income values in the sequence we want. Normally, the first cash flow value in the range is a negative number that represents an outgoing cash flow. If an array or reference argument contains text, logical values, or empty cells, those values are ignored.

guess = a number that we guess is close to the result of IRR.

Excel uses an iterative method for calculating IRR. Starting with **guess**, IRR cycles through the calculation until the result is accurate within 0.00001%. If IRR cannot find a result that works after 20 tries, the **#NUM!** error value is returned.

In most cases we need not provide **guess** for the IRR calculation. If **guess** is omitted, it is assumed to be 0.1 (10%).

If IRR gives the **#NUM!** error value, or if the result is not close to what we expected, we should try again with a different value for **guess**.

Example 6.39

Suppose that we make an initial investment of \$10,000 and we receive 12 unequal payments as shown in Column A of the spreadsheet of Figure 6.9 below. What is the internal rate of return?

Solution:

In Cell C2 of the spreadsheet of Figure 6.9 below we enter =IRR(A2:A14) and this returns 10.10%.

	A	B	C
1	Cash Flows		IRR
2	-10000		10.10%
3	1800		
4	1500		
5	1250		
6	1400		
7	1375		
8	1280		
9	1625		
10	1560		
11	1230		
12	1425		
13	1580		
14	1620		

Figure 6.9. Spreadsheet for Example 6.39

Remark:

IRR is closely related to NPV, the net present value function. The rate of return calculated by IRR is the interest rate corresponding to a 0 (zero) net present value. The following formula illustrates how NPV and IRR are related:

$\text{NPV}(\text{IRR}(\text{A2:A14}), \text{A2:A14})$ returns 9.29×10^{-13} .

Example 6.40

Suppose we started our business with an initial investment of \$60,000 and our net income for the first five years is as shown on the spreadsheet of Figure 6.10. We want to compute:

- a. the internal rate of return after 4 years
- b. the internal rate of return after 5 years
- c. the internal rate of return after 2 years

Solution:

The formulas that we used for this example are shown in Cells A10:A13. The internal rate of return for the first 4, first 5, and first 2 years are shown in Cells B10, B11, and B13. We observe that for the last computation it was necessary to enter a guess of -10% .

	A	B
1	Description	Cash Flows
2	Initial Cost of Business	-60000
3	Net Income for 1st year	10000
4	Net Income for 2nd year	13000
5	Net Income for 3rd year	15000
6	Net Income for 4th year	18000
7	Net Income for 5th year	23000
8		
9	Formula	Result
10	=IRR(B2:B6)	-2%
11	=IRR(B2:B7)	9%
12	=IRR(B2:B4)	#NUM!
13	=IRR(B2:B4, -0.10)	-44%

Figure 6.10. Spreadsheet for Example 6.40

6.8.8 MIRR Function

The MIRR function returns the modified internal rate of return for a series of periodic cash flows. MIRR considers both the cost of the investment and the interest received on reinvestment of cash.

The syntax for Excel is

=MIRR(values,finance_rate,reinvest_rate) where

values = an array or a reference to cells that contain numbers for which we want to calculate the internal rate of return. This argument must contain at least one positive value and one negative value to calculate the internal rate of return; otherwise, MIRR returns the #DIV/0! error value. MIRR uses the order of values to interpret the order of cash flows. We must enter our payment and income values in the sequence we want. Normally, the first cash flow value in the range is a negative number that represents an outgoing cash flow. If an array or reference argument contains text, logical values, or empty cells, those values are ignored; however, cells with zero value are included.

finance_rate = the interest rate that we pay on the money used in the cash flows.

reinvest_rate = the interest rate that we receive on the cash flows as we reinvest them.

MIRR uses the following formula:

$$\text{MIRR} = \left(\frac{-\text{NPV}(\text{reinvest_rate}, \text{values}[\text{positive}]) \times (1 + \text{reinvest_rate})^n}{\text{NPV}(\text{finance_rate}, \text{values}[\text{negative}]) \times (1 + \text{finance_rate})} \right)^{\frac{1}{n-1}} - 1 \quad (6.38)$$

where n is the number of cash flows in values.

Example 6.41

Suppose we started our business with a loan of \$120,000 and our net income for the first five years including interest, is as shown on the spreadsheet of Figure 6.11.

	A	B
1	Description	Cash Flows & Interest
2	Initial Cost of Business (Loan)	-120000
3	Return for 1st year	39000
4	Return for 2nd year	30000
5	Return for 3rd year	21000
6	Return for 4th year	37000
7	Return for 5th year	46000
8	Annual interest rate for the \$120,000 loan	10%
9	Annual interest rate for the reinvested profits	12%
10		
11	Formula	Result
12	=MIRR(B2:B7,B8,B9)	13%
13	=MIRR(B2:B5,B8,B9)	-5%
14	=MIRR(B2:B7,B8,0.14)	13%

Figure 6.11. Spreadsheet for Example 6.41

We want to compute:

- a. the modified rate of return after 5 years based on a reinvest rate of 12%
- b. the modified rate of return after 3 years based on a reinvest rate of 12%
- c. the 5-year modified rate of return based on a reinvest rate of 14% instead of 12%

Solution:

The formulas that we used for this example are shown in Cells A12:A14. The modified rate of return after 5 years and 3 years based on a reinvest rate of 12% is shown on B12 and B13 respectively. The modified rate of return after 5 years based on a reinvest rate of 14% is shown on B14.

6.8.9 IPMT Function

The IPMT function returns the interest payment for a given period for an investment based on periodic, constant payments, and a constant interest rate.

The syntax for Excel is

=IPMT(rate,per,nper,pv,fv,type) where

rate = interest rate per period

per = the period for which we want to find the interest. It must be in the range 1 to **nper**.

nper = the total number of payment periods

pv = present value or the total amount that a series of future payments is worth at the present.

fv = future value or the cash balance that we want to attain after the last payment is made. If **fv** is omitted, it is assumed to be 0 (zero), that is, the future value of a loan is zero.

type = the number 0 or 1 which indicates when payments are due. If 0, the payments are due at the end of the period. If 1, the payments are due at the beginning of the period. If omitted, it is assumed to be 0.

Remarks:

1. For all arguments, cash that we pay out, such as deposits to savings, is represented by negative numbers. Cash we receive, such as dividends, is represented by positive numbers.
2. We must make sure that we are consistent with the units that we use for specifying **rate** and **nper**. For example, if we make monthly payments on a four-year loan at 11% annual interest rate, we must use 11%/12 for **rate** and 4×12 for **nper**. If we make annual payments on the same loan, we must use 11% for **rate** and 4 for **nper**.

Example 6.42

We took out an \$8,000 loan for 3 years at an annual interest rate of 10%. We want to compute:

- a. the interest due in the first month of the first year
- b. the interest due for the last year of the loan assuming that the payments are made annually.

Solution:

- a. We use the formula =IPMT(0.1/12,1,3,8000) where the interest rate is divided by 12 to obtain the monthly rate. This formula returns -\$66.67 and this amount represents the interest paid in the first month of the first period.

b. We use the formula `=IPMT(0.1,3,3,8000)` where the interest rate is not divided by 12 since we are interested in the annual interest payment for the last (third) period. This formula returns \$292.45 and this amount represents the interest paid for the entire third year.

6.8.10 PPMT Function

The PPMT function returns the payment on the principal for a given period for an investment based on periodic, constant payments, and a constant interest rate.

The syntax for Excel is

`=PPMT(rate,per,nper,pv,fv,type)` where

rate = interest rate per period

per = specifies the period. It must be in the range 1 to **nper**.

nper = the total number of payment periods

pv = present value or the total amount that a series of future payments is worth at the present.

fv = future value or the cash balance that we want to attain after the last payment is made. If **fv** is omitted, it is assumed to be 0 (zero), that is, the future value of a loan is zero.

type = the number 0 or 1 which indicates when payments are due. If 0, the payments are due at the end of the period. If 1, the payments are due at the beginning of the period. If omitted, it is assumed to be 0.

Remark:

We must make sure that we are consistent with the units that we use for specifying **rate** and **nper**. For example, if we make monthly payments on a four-year loan at 11% annual interest rate, we must use 11%/12 for **rate** and 4×12 for **nper**. If we make annual payments on the same loan, we must use 11% for **rate** and 4 for **nper**.

Example 6.43

We took out an \$6,000 loan for 5 years at an annual interest rate of 11%. We want to compute the principal payment for the first month of the first year.

Solution:

We use the formula `=PPMT(0.11/12,1,5*12,6000)` where the interest rate is divided by 12 to obtain the monthly interest rate, and the number of years is multiplied by 12 to obtain the total

number of monthly payments. The formula above returns $-\$75.45$ and this amount represents the interest paid for the first month of the first year.

Example 6.44

We took out an \$300,000 loan for 10 years at an annual interest rate of 9%. We want to compute the principal payment for the last year of the 10-year period.

Solution:

We use the formula `=PPMT(0.09,10,10,300000)`. This formula above returns $-\$42,886$ and this amount represents the principal payment for the last year of the 10-year period.

6.8.11 ISPMT Function

The ISPMT function calculates the interest paid during a specific period of an investment.

The syntax for Excel is

`=ISPMT(rate,per,nper,pv)` where

rate = interest rate for the investment

per = specifies the period for which we want to find the interest and must be in the range 1 to **nper**.

nper = the total number of payment periods

pv = present value of the investment. For a loan, **pv** is the loan amount.

Remarks:

1. For all arguments, cash that we pay out, such as deposits to savings, is represented by negative numbers. Cash we receive, such as dividends, is represented by positive numbers.
 2. We must make sure that we are consistent with the units that we use for specifying **rate** and **nper**. For example, if we make monthly payments on a four-year loan at 11% annual interest rate, we must use 11%/12 for **rate** and 4×12 for **nper**. If we make annual payments on the same loan, we must use 11% for **rate** and 4 for **nper**.
-

Example 6.45

We have decided to take out a \$6,000,000 loan for 8 years at an annual interest rate of 8.5%. We want to compute

- a. the interest that we must pay the first month of the first year
- b. the interest that we must pay for first year of the 8 –year period

Solution:

- a. We use the formula `=ISPMT(0.085/12,1,8*12,6000000)` where the interest rate is divided by 12 to obtain the monthly interest rate, and the number of years is multiplied by 12 to obtain the total number of monthly payments. The formula above returns $-\$42,057.29$ and this amount represents the interest that we must pay the first month of the first year.
- b. We use the formula `=ISPMT(0.085,1,8,6000000)`. The formula above returns $-\$446,250.00$ and this amount represents the interest that we must pay the entire first year.

6.9 The MATLAB Financial Toolbox

The MATLAB Financial Toolbox that contains several cash–flow functions to compute interest rates, rates of return, present or future values, depreciation, and annuities. These are very similar to the financial functions provided by Microsoft Excel. It also includes the Financial Derivatives Toolbox and the Financial Time Series Toolbox. However, in this section, we will restrict our discussion on the most common functions of the Financial Toolbox.

Like the Microsoft Excel functions, with all of the MATLAB financial functions investments are considered negative cash flows, and return payments are considered positive cash flows.

6.9.1 irr MATLAB Function

The **irr** function computes the internal rate of return of periodic cash flows. Its syntax is **r=irr(cf)** where **cf** is a cash flow vector.

The first entry in **cf** is the initial investment and it is entered as a negative number. If **cf** is entered as a matrix, each column is treated as a separate cash flow.

Example 6.46

Suppose that an initial investment of \$150,000 is made and the annual income for this investment is as shown in Table 6.14.

TABLE 6.14 Income for Example 6.46

Year	1	2	3	4	5
Income	15,000	25,000	35,000	50,000	60,000

We can compute the internal rate of return with the MATLAB statement

`internal_rate=irr([-150000 15000 25000 35000 45000 60000])` and this returns

```
internal_rate =  
    0.05
```

6.9.2 **effrr** MATLAB Function

The **effrr** function computes the effective rate of return given an annual interest rate, also known as Annual Percentage Rate (APR), and number of compounding periods per year. Its syntax is

r=effrr(apr,per) where

apr is the annual interest rate and **per** is the number of compounding periods per year.

Example 6.47

Suppose that the annual percentage interest rate is 7% and it is compounded monthly. We can compute the effective annual rate with the MATLAB statement

`effective_rate=effrr(0.07,12)` and this returns

```
effective_rate =  
    0.0723
```

That is, the effective rate is 7.23%.

6.9.3 **pvfix** MATLAB Function

The **pvfix** function computes the present value of cash flows at regular time intervals with equal or unequal payments. Its syntax is

p=pvfix(rate,nper,p,fv,due) where

rate is the periodic interest rate, **nper** is the number of periods, **p** is the periodic payment, **fv** is a payment received other than **p** in the last period, and **due** specifies whether the payments are made at the beginning (**due=1**) or end (**due=0**) of the period. The default arguments are **fv=0** and **due=0**.

Example 6.48

A \$400 payment is made monthly into a savings account earning 4%. The payments are made at the end of the month for 6 years. To compute the present value of these payments in a fixed format with two decimal places, we use the statements

```
format bank
```

`present_value_equal_fixed_payments=pvfix(0.04/12,6*12,400,0,0)` and these return

```
present_value_equal_fixed_payments =
    25566.97
```

Example 6.49

Payments of \$300, \$325, \$350, \$375, \$400, \$425, \$450, and \$475 are made semiannually into a savings account earning 5% per year. The payments are made at the end of June and at the end of December for 4 years. To compute the present value of these payments in a fixed format with two decimal places, we use the statements

`format bank`

```
payments=[300 325 350 375 400 425 450 475];
```

`present_value_unequal_fixed_payments=pvfix(0.05/2,4*2,payments,0,0)` and these return

```
present_value_unequal_fixed_pay =
    2151.04    2330.29    2509.55    2688.80    2868.05    3047.31    3226.56
    3405.82
```

These values represent the present value at the end of each semiannual period.

6.9.4 pvvar MATLAB Function

The **pvvar** function computes the present value of cash flows at irregular time intervals with equal or unequal payments. Its syntax is

```
pv=pvvar(cf,rate,df)
```

where **cf** is a vector representing the cash flows, **rate** is the periodic interest rate, and **df** represents the dates that the cash flows occur. If **df** omitted, it is understood that the cash flows occur at the end of the period.

Example 6.50

Suppose we make an initial investment of \$15,000 at an annual interest rate of 7% and the yearly income realized by this investment is as shown in Table 6.15.

TABLE 6.15 Cash flows for Example 6.50

Year	1	2	3	4	5
Return	\$3,500	\$4,000	\$3,750	\$4,200	\$4,100

Chapter 6 Mathematics of Finance and Economics

To compute the net present value of the periodic cash flows, we use the following statements:

format bank

```
CF=[-15000 3500 4000 3750 4200 4100]; present_value=pvvar(CF,0.07) and these return
```

```
present_value =
```

```
953.30
```

Example 6.51

An investment of \$15,000 returns the cash flows shown in Table 6.16 where the first cash flow value is shown as the first cash flow value. The annual interest rate is 8%.

TABLE 6.16 Cash flows for Example 6.51

Cash Flow	Date
\$-15,000	January 23, 2001
3,500	February 27, 2002
4,000	March 18, 2002
3,750	July 15, 2002
4,200	September 9, 2002
4,100	November 20, 2002

To compute the present value of these cash flows we use the following statements:

format bank

```
CF=[-15000 3500 4000 3750 4200 4100];
```

```
DF=['01/23/2001'; '02/27/2002'; '03/18/2002'; '07/15/2002'; '09/09/2002'; '11/20/2002'];
```

```
present_value=pvvar(CF,0.08, DF) and these return
```

```
present_value =
```

```
2494.95
```

6.9.5 fvmf MATLAB Function

The **fvmf** function computes the future value of cash flows at regular time intervals with equal or unequal payments. Its syntax is

```
f=fvmf(rate,nper,p,pv,due)
```

where **rate** is the periodic interest rate, **nper** is the number of periods, **p** is the periodic payment, **pv** is the initial value, and **due** specifies whether the payments are made at the beginning (**due=1**) or end (**due=0**) of the period. The default arguments are **pv=0** and **due=0**.

Example 6.52

Suppose that we already have \$2,000 in a savings account and we add \$250 at the end of each month for 10 years. This account pays 6% interest compounded monthly. To compute the future value of our investment at the end of 10 years, we use the following statements:

```
format bank
future_value=fvfix(0.06/12,10*12,250,2000,0) and these return
future_value =
    44608.63
```

6.9.6 fvvar MATLAB Function

The **fvvar** function computes the future value of cash flows at irregular time intervals with equal or unequal payments. Its syntax is

```
fv=fvvar(cf,rate,df)
```

where **cf** is a vector representing the cash flows, **rate** is the periodic interest rate, and **df** represents the dates that the cash flows occur. If **df** omitted, it is understood that the cash flows occur at the end of the period. The initial investment is included as the initial cash flow value.

Example 6.53

Suppose we make an initial investment of \$15,000 at an annual interest rate of 7% and the yearly income realized by this investment is as shown in Table 6.17.

TABLE 6.17 Cash flows for Example 6.53

Year	1	2	3	4	5
Return	\$3,500	\$4,000	\$3,750	\$4,200	\$4,100

To compute the future value of the periodic cash flows, we use the following statements:

```
format bank
CF=[-15000 3500 4000 3750 4200 4100];
future_value=fvvar(CF,0.07) and these return
```

```
future_value =  
    1337.06
```

Example 6.54

An investment of \$15,000 returns the cash flows shown in Table 6.18 where the first cash flow value is shown as the first cash flow value. The annual interest rate is 8%.

TABLE 6.18 Cash flows for Example 6.54

Cash Flow	Date
\$-15,000	January 23, 2001
3,500	February 27, 2002
4,000	March 18, 2002
3,750	July 15, 2002
4,200	September 9, 2002
4,100	November 20, 2002

To compute the present value of these cash flows we use the following statements:

```
format bank
```

```
CF=[-15000 3500 4000 3750 4200 4100];
```

```
DF=['01/23/2001'; '02/27/2002'; '03/18/2002'; '07/15/2002'; '09/09/2002'; '11/20/2002'];
```

```
future_value=fvvar(CF,0.08, DF) and these return
```

```
future_value =  
    2871.10
```

6.9.7 annurate MATLAB Function

The **annurate** function computes the periodic interest rate paid on a loan or annuity. Its syntax is

```
r=annurate(nper,pv,fv,due)
```

where **rate** is the periodic interest rate, **nper** is the number of periods, **p** is the periodic payment, **pv** is the present value of the loan or annuity, **fv** is the future value of the loan or annuity, and **due** specifies whether the payments are made at the beginning (**due=1**) or end (**due=0**) of the period. The default arguments are **fv=0** and **due=0**.

Example 6.55

Suppose that we have taken a loan of \$8,000 for 5 years and we make monthly payments of \$225 at the end of each month. To compute the interest rate that we pay for this loan, we use the following statements:

format bank

monthly_rate=annurate(5*12,225,8000,0,0) and these return

```
monthly_rate =
    0.02
```

Therefore, the annual interest rate is $0.02 \times 12 = 0.24$ or 24%.

6.9.8 amortize MATLAB Function

The **amortize** function computes the principal and interest portions of a loan paid by a periodic payment and returns the remaining balance of the original loan amount and the periodic payment. Its syntax is

```
[principal, interest, balance, periodic_payment]=amortize(rate,nper,pv,fv,due)
```

where **rate** is the periodic interest rate that we pay, **nper** is the number of periods, **pv** is the present value of the loan or annuity, **fv** is the future value of the loan or annuity, and **due** specifies whether the payments are made at the beginning (**due=1**) or end (**due=0**) of the period. The default arguments are **fv=0** and **due=0**. On the left side of above statement **principal** is a vector of the principal paid in each period, **interest** is a vector of the interest paid in each period, **balance** is the remaining balance of the loan in each payment period, and **periodic_payment** is the periodic payment.

Example 6.56

Suppose we took a loan of \$2,000 to be paid in 6 bimonthly installments at an annual interest rate of 7.5%. The following statements will compute the principal paid in each period, the interest paid in each period, the remaining balance in each period, and the periodic payment.

format bank

```
[principal, interest, balance, periodic_payment]=amortize(0.075/6, 6, 2000) and these return
```

```
principal =
    323.07    327.11    331.19    335.33    339.53    343.77
interest =
```

```
25.00 20.96 16.87 12.73 8.54 4.30
balance =
1676.93 1349.83 1018.63 683.30 343.77 -0.00
periodic_payment =
348.07
```

The MATLAB Financial Toolbox includes the Securities Industry Association (SIA) * compliant functions to compute accrued interest, determine prices and yields, and calculate duration of fixed-income securities. It also includes a set of functions to generate and analyze term structure of interest rates.

6.10 Comparison of Alternate Proposals

Quite often, one must perform economic analysis to select equipment or property from two or more proposals. In this section we will discuss the annual cost comparison method by comparing the total annual costs for each proposal.

Let

PP = Purchase Price (initial cost) of asset

SV = Salvage Value of asset

n = life of asset in years

i = Annual interest rate

AE = Annual Expenditures

The annual expenditures include all costs except depreciation and interest. These expenditures include taxes, insurance, materials, labor, repairs and maintenance, supplies, and utilities.

The effective purchase price at the retirement date of an asset is computed as

$$PP(1 + i)^n - SV$$

Converting this to an equivalent annuity, we obtain

* SIA-compliant functions are normally used with U.S. Treasury bills, bonds (corporate and municipal), and notes.

$$\begin{aligned} \text{Total Annual Cost} &= [PP(1+i)^n - SV] \frac{i}{(1+i)^n - 1} + AE \\ &= \frac{PPi(1+i)^n - SVi + PPi - PPi}{(1+i)^n - 1} + AE = \frac{(PP - SV)i + PPi[(1+i)^n - 1]}{(1+i)^n - 1} + AE \end{aligned}$$

or

$$\text{Total Annual Cost} = (PP - SV) \frac{i}{(1+i)^n - 1} + PPi + AE \quad (6.39)$$

The total annual cost of an asset is a highly useful tool in arriving at a business decision where an economic choice is to be made among several alternative courses of action. The following examples will illustrate the applicability of annual cost for comparison purposes.

Example 6.57

A manufacturer has bought a piece of property on which he will build a storage warehouse. Two types of construction are considered, brick and galvanized iron. The data available are shown in Table 6.19.

TABLE 6.19 Data for Example 6.57

	Brick	Galvanized Iron
Initial cost (purchase price)	\$24,000	\$10,000
Salvage Value	5,000	2,000
Estimated life	50 years	15 years
Annual maintenance	200	300
Annual insurance premium	48	50
Annual taxes	300	125
Interest	4%	4%

Which type of construction is the most economical?

Solution:

The equation of (6.40) applies to this example. For convenience, we construct Table 6.20 below to consolidate the data shown in Table 6.19 above.

TABLE 6.20 Consolidation of the data of Table 6.19

	Brick	Galvanized Iron
PP (Purchase Price)	\$24,000	\$10,000
SV (Salvage Value)	5,000	2,000
n (Estimated life)	50 years	15 years
AE (Maintenance, Insurance, Taxes)	200+48+300=548	300+50+125=475
Interest	4%	4%

Then,

$$\begin{aligned} \text{Total Annual Cost}_{\text{brick}} &= (24000 - 5000) \frac{0.04}{(1 + 0.04)^{50} - 1} + 24000 \times 0.04 + 548 \\ &= 19000 \times 0.00655 + 960 + 548 = \$1, 632.45 \end{aligned}$$

Likewise,

$$\begin{aligned} \text{Total Annual Cost}_{\text{galvanized iron}} &= (10000 - 2000) \frac{0.04}{(1 + 0.04)^{15} - 1} + 10000 \times 0.04 + 475 \\ &= 8000 \times 0.04994 + 400 + 475 = 1, 274.53 \end{aligned}$$

Therefore, the galvanized construction is the most economical choice.

Example 6.58

Two possible routes for a power line are under consideration. Route A is around a lake, 15 miles in length. The first cost will be \$6,000 per mile, the yearly maintenance will be \$2,000 per mile, and there will be a salvage value at the end of 15 years of \$3,000 per mile. Route B is a submarine cable across the lake, 5 miles long. The first cost will be \$31,000 per mile, the annual maintenance \$400 per mile, and the salvage value at the end of 15 years will be \$6,000 per mile. The yearly power loss will be \$550 per mile for both routes. Interest rate is 4.5%; taxes are 3% of the first cost. Compare the two routes on the basis of annual costs.

Solution:

For convenience, we tabulate the given data on Table 6.21 below.

Then, for Route A

$$\begin{aligned} \text{Total Annual Cost}_{\text{Route A}} &= (90000 - 45000) \frac{0.045}{(1 + 0.045)^{15} - 1} + 90000 \times 0.045 + 40950 \\ &= 45000 \times 0.04811 + 4050 + 40950 = \$47, 165.12 \end{aligned}$$

TABLE 6.21 Given data for Example 6.58

	Route A	Route B
PP (Purchase Price, i.e., Initial Cost)	$\$6,000 \times 15 = \$90,000$	$\$31,000 \times 5 = \$155,000$
SV (Salvage Value)	$3,000 \times 15 = 45,000$	$6,000 \times 5 = 30,000$
Interest	4.5%	4.5%
Annual Expenditures		
Maintenance	$2,000 \times 15 = 30,000$	$400 \times 5 = 2,000$
Power loss (\$550/mile)	$550 \times 15 = 8,250$	$550 \times 5 = 2,750$
Taxes at 3% of initial cost	$90,000 \times 0.03 = 2,700$	$155,000 \times 0.03 = 4,650$
Total Annual Expenditures	$30,000 + 8,250 + 2,700 = \$40,950$	$2,000 + 2,750 + 4,650 = \$9,400$

Likewise, for Route B

$$\begin{aligned} \text{Total Annual Cost}_{\text{Route B}} &= (155000 - 30000) \frac{0.045}{(1 + 0.045)^{15} - 1} + 155000 \times 0.045 + 9400 \\ &= 125000 \times 0.04811 + 4050 + 9400 = \$22,389.23 \end{aligned}$$

Therefore, Route B is the most economical choice.

6.11 Kelvin's Law

On the previous section, we discussed how we can make an economic choice by computing the annual costs of two or more alternatives. In this section, we will discuss a class of problems in which the total cost can be found by the following formula:

$$y = ax + \frac{b}{x} + c \quad (6.40)$$

where

y = total cost

a , b , and c are constants

x = a variable

The total cost y becomes a minimum when

$$\frac{dy}{dx} = 0 = a - \frac{b}{x^2}$$

or when

$$x = \sqrt{\frac{b}{a}} \quad (6.41)$$

The pair of the equations of (6.40) and (6.41) are known as *Kelvin's law*. This law can be applied to special types of problems such as the most economical span length of a bridge, most economical lot size, and most economical diameter of an oil pipeline.

Example 6.59

A railroad company proposes to build a double-track half-through plate-girder bridge 980 feet long, with spans of equal length. The cost of steel in place is estimated at \$0.10 cents per pound. Disregarding the cost of the track as constant, find the economical length of span. For superstructure use the equation

$$\text{Weight}_{\text{lb/linear foot}} = 25s + 2050$$

where s is the span length in feet.

For the center piers use the equation

$$\text{Cost}_{\text{center piers/pier}} = 50s + 30000$$

and for the end piers use the equation

$$\text{Cost}_{\text{end piers/pier}} = 50s + 20000$$

Solution:

The total weight of steel in the superstructure is

$$\text{Steel}_{\text{total weight}} = 980 \text{ feet} \times \text{Weight/linear foot} = 980(25s + 2050) = 24500s + 2009000$$

Cost of steel at \$0.10/lb is

$$\text{Cost}_{\text{steel}} = 0.10(24500s + 2009000) = 2450s + 200900$$

As stated, s is the span length in feet. Then, the number of spans is

$$\text{Number of spans} = \frac{\text{Bridge length}}{\text{Span length}} = \frac{980}{s}$$

There are 2 end piers and the number of center piers is one less than the number of spans. Thus,

$$\text{Cost}_{\text{end piers}} = 2(50s + 20000) = 100s + 40000$$

and

$$\begin{aligned}\text{Cost}_{\text{center piers}} &= \left(\frac{980}{s} - 1\right)(50s + 30000) = 49000 - 50s + \frac{29400000}{s} - 30000 \\ &= \frac{29400000}{s} - 50s + 19000\end{aligned}$$

Therefore, denoting the total cost of bridge as y , we obtain

$$\begin{aligned}y &= \text{Cost}_{\text{steel}} + \text{Cost}_{\text{end piers}} + \text{Cost}_{\text{center piers}} \\ &= 2450s + 200900 + 100s + 40000 + \frac{29400000}{s} - 50s + 19000 \\ &= 2500s + \frac{29400000}{s} + 259900\end{aligned}$$

By comparison of the above expression with the equation of (6.40), Page 6–71, we observe that

$$a = 2500 \quad b = 29400000 \quad x = s$$

Therefore, the most economical length of each bridge span is when y is minimum, and by equation (6.41), Page 6–72, it occurs when

$$x = s = \sqrt{\frac{b}{a}} = \sqrt{\frac{29400000}{2500}} = 108.44353 \approx 108.5 \text{ feet}$$

and the number of spans is

$$\text{Number of spans} = \frac{\text{Bridge length}}{\text{Span length}} = \frac{980}{108.5} = 9.03 \approx 9 \text{ spans}$$

Example 6.60

A manufacturer must to produce 50,000 parts a year and wishes to select the most economical lot size to be equally spaced during the year. The annual cost of storage and interest on investment varies directly with the lot size and is \$1,000 when the entire annual requirement is manufactured in one lot. The cost of setting up and dismantling the machine for each run is \$30. What is the most economical lot size?

Solution:

Let x be the most economical lot size. Then, the number of machine setups will be $50000/x$.

Chapter 6 Mathematics of Finance and Economics

Annual cost of setups will be

$$30 \times \frac{50000}{x} = \frac{1500000}{x}$$

and since the entire annual cost of storage and interest varies directly with the lot size and is \$1,000 when manufactured in one lot, by proportion we have

$$\frac{\$1000}{50000 \text{ parts}} \times x = 0.02x$$

Now, denoting the total cost to produce 50,000 units as y , we have

$$y = 0.02x + \frac{1500000}{x}$$

Therefore, $a = 0.02$, $b = 1500000$, and per equation (6.41), Page 6–72, y will be a minimum cost when

$$x = \sqrt{\frac{b}{a}} = \sqrt{\frac{1500000}{0.02}} = 8660 \text{ parts per lot}$$

and the number of lots required will be

$$\frac{50000}{8660} = 5.8$$

If we round this number to 6, we find that the most economical lot size is

$$\frac{50000}{6} = 8333 \text{ parts}$$

6.12 Summary

- A bond is a debt investment. That is, we loan money to an entity (company or government) that needs funds for a defined period of time at a specified interest rate.
- A corporate bond is a bond issued by a corporation.
- A municipal bond is a bond issued by a municipality and that generally is tax-free.
- A treasury bond is a bond issued by the US Government.
- A perpetuity is a constant stream of identical cash flows with no maturity date.
- A perpetual bond is a bond with no maturity date. Perpetual bonds are not redeemable and pay a steady stream of interest forever. Such bonds have been issued by the British government.
- A convertible bond is a bond that can be converted into a predetermined amount of a company's equity at certain times during its life.
- A treasury note is a treasury bond that is issued for a shorter time.
- A treasury bill is a treasury bond that is held for a shorter time than either a treasury bond or a treasury note.
- Face value is the dollar amount assigned to a bond when first issued.
- Par value is the face value of a bond.
- Book value is the value of a bond at any date intermediate between its issue and its redemption.
- A coupon bond is a debt obligation with coupons representing semiannual interest payments attached.
- A zero coupon bond is a bond that generates no periodic interest payments and is issued at a discount from face value. The return is realized at maturity.
- A junk bond is a bond purchased for speculative purposes. It has a low rating and a high default risk.
- Bonds are assigned a rating by Standard and Poor's and Moody's.
- A promissory note is an unconditional written promise made by one party to another, to pay a stipulated sum of money either on demand or at a definite future date.
- The discount rate is the rate charged by Reserve Banks when they extend credit to depository institutions either through advances or through the discount of certain types of paper, including ninety-day commercial paper.
- The prime rate is the interest rate charged by banks to their most creditworthy customers.

Chapter 6 Mathematics of Finance and Economics

- An annuity is a series of equal payments made at equal intervals or periods of time. When paid into a fund which is invested at compound interest for a specified number of years, the annuity is referred to as a sinking fund.
- An ordinary annuity is a series of equal payments or receipts occurring over a specified number of periods with the payments or receipts occurring at the end of each period.
- A sinking fund is an interest-earning fund in which equal deposits are made at equal intervals of time. It is also referred to as an annuity.
- Interest is charge for a loan.
- Simple interest is calculated on a yearly basis (annually) and depends on the interest rate.
- Compound interest includes interest earned on interest.
- If a given interest rate applies to a period less than 1 year, then its equivalent rate for an annual period is referred to as its effective rate. Thus, the effective rate corresponding to a rate of 2% per quarterly period is 8.243% per cent. The effective rate is numerically equal to the interest earned by a principal of \$1.00 for 1 year.
- The Excel PV function returns the present value of an investment. The present value is the total amount that a series of future payments is worth now. This function also computes the present value of an ordinary annuity.
- The Excel FV function computes the future value of an ordinary annuity.
- The Excel PMT function calculates the payment for a loan based on constant payments and a constant interest rate.
- The Excel RATE function returns the interest rate per period of an annuity. We can use RATE to determine the yield of a zero-coupon bond that is sold at a discount of its face value.
- The Excel NPER function returns the number of periods for an investment based on periodic, constant payments, and a constant interest rate.
- The Excel NPV function calculates the net present value of an investment by using a discount rate and a series of future payments (negative values) and income (positive values).
- The Excel IIR function returns the internal rate of return for a series of cash flows represented by the numbers in values. These cash flows need not be even as they would be for an annuity.
- The Excel MIRR function returns the modified internal rate of return for a series of periodic cash flows. MIRR considers both the cost of the investment and the interest received on reinvestment of cash.
- The Excel IPMT function returns the interest payment for a given period for an investment based on periodic, constant payments, and a constant interest rate.

- The Excel **PPMT** function returns the payment on the principal for a given period for an investment based on periodic, constant payments, and a constant interest rate.
- The Excel **ISPMT** function calculates the interest paid during a specific period of an investment.
- The MATLAB **irr** function computes the internal rate of return of periodic cash flows.
- The MATLAB **effrr** function computes the effective rate of return given an annual interest rate, also known as Annual Percentage Rate (APR), and number of compounding periods per year.
- The MATLAB **pvfix** function computes the present value of cash flows at regular time intervals with equal or unequal payments.
- The MATLAB **pvvar** function computes the present value of cash flows at irregular time intervals with equal or unequal payments.
- The MATLAB **fvfix** function computes the future value of cash flows at regular time intervals with equal or unequal payments.
- The MATLAB **fvvar** function computes the future value of cash flows at irregular time intervals with equal or unequal payments.
- The MATLAB **annurate** function computes the periodic interest rate paid on a loan or annuity.
- The MATLAB **amortize** function computes the principal and interest portions of a loan paid by a periodic payment and returns the remaining balance of the original loan amount and the periodic payment.
- The material presented in this chapter allows us to perform economic analysis to select equipment or property from two or more proposals.
- Kelvin's law can be applied to special types of problems such as the most economical span length of a bridge, most economical lot size, and most economical diameter of an oil pipeline.

6.13 Exercises

1. Jack's bank account pays interest at a rate of 4.3% per year. If he puts \$850 into his account, how much will Jack have after a year?
2. Mary earns 4.5% interest per year on the money she has saved in her bank account. Her initial investment was \$200. How much will she have on her account after 5 years?
3. Lisa has \$560. She deposits it into a bank account that pays 3.7% interest per annum. She leaves it in the bank for 2 months. How much money will Lisa have now?
4. If we deposit \$1,500 into an account earning interest rate of 5% compounded quarterly, what will our principal be at the end of 7 years?
5. An individual possesses a promissory note, due 2 years hence, whose maturity value is \$3,200. What is the discount value of this note based on an interest rate of 7%?
6. On July 1, 2001, a bank account had a balance of \$8000. A deposit of \$1000 was made on Jan. 1, 2002, and a deposit of \$750 on Jan. 1, 2003. On Apr. 1, 2004, the sum of \$1,200 was withdrawn. What was the balance in the account on Jan. 1, 2007, if interest is earned at the rate of 4% per cent compounded quarterly?
7. A manufacturing firm contemplates retiring an existing machine at the end of 2016. The new machine to replace the existing one will have an estimated cost of \$10,000. This expense will be partially defrayed by sale of the old machine as scrap for \$750. To accumulate the balance of the required capital, the firm will deposit the following sums in an account earning interest at 5% compounded quarterly:
 $\$1,500$ at the end of 2013 $\$1,500$ at the end of 2014 $\$2,000$ at the end of 2015
What cash disbursement will be necessary at the end of 2016 to purchase the new machine?
8. A father wishes to have \$50,000 available at his son's eighteenth birthday. What sum should be set aside at the son's fifth birthday if it will earn interest at the rate of 3% per cent compounded semiannually?
9. In Exercise 8, if the father will set aside equal sums of money at the son's fifth, sixth, and seventh birthdays, what should each sum be?
10. Brown owes Smith the following sums:
 $\$1,000$ due 2 years hence $\$1,500$ due 3 years hence $\$1,800$ due 4 years hence
Having received an inheritance, he has decided to liquidate the debts at the present date. If the two parties agree on a 5% interest rate, what sum must Brown pay? In order to verify the

solution obtained, evaluate all sums of money, including the required payment, at any date other than the present.

11. How much interest will \$1,000 earn if it is invested at 6% for 5 years?
12. A business firm contemplating the installation of labor-saving machinery has a choice between two different models. Machine A will cost \$36,500, while machine B will cost \$36,300. The repairs required for each machine are as follows:

Machine A: \$1,500 at end of 5th year and \$2,000 at end of 10th year

Machine B: \$3,800 at end of 9th year

The machines are alike in all other respects. If this firm is earning a 7% per cent return on its capital, which machine should be purchased?

13. A fund was established in 2001, whose history is recorded below:

Deposit of \$1,000 on 1/1/2001

Deposit of \$2,000 on 1/1/2003

Withdrawal of \$300 on 7/1/2003

Deposit of \$1,600 on 7/1/2004

Withdrawal of \$1,200 on 1/1/2005

The fund earned interest at the rate of 3.5% compounded semiannually until the end of 2003. At that date, the interest rate was augmented to 4% compounded semiannually. What was the principal in the fund at the end of 2007?

14. If the sum of \$2,000 is invested in a fund earning 7.5% interest compounded semiannually, what will be its final value at the end of 5 years?
15. In Exercise 14, how long will it take the original sum of \$2,000 to expand to \$3,000?
16. On Apr. 1, 2005, an investor purchased stock of the XYZ Corp. at a total cost of \$3,600. He then received the following semiannual dividends:
- \$105 on Oct. 1, 2005 \$110 on Apr. 1, 2006
\$110 on Oct. 1, 2006 \$100 on Apr. 1, 2007

After receipt of the last dividend, the investor sold his stock, receiving \$3,800 after deduction of brokerage fees. What semiannual rate did this individual realize on his investment?

17. A group of investors is considering alternate plans for the investment of \$200,000,000.

Plan A involves the purchase of non-appreciating tax-free securities paying 1.5% per annum. Dividends will be invested at a net return of 2%, and the securities will be sold at the end of 20 years.

Plan B calls for the purchase of income property consisting of land worth \$40,000,000 and apartments worth \$160,000,000. It is estimated that the units will have an average occupancy of 90%, a useful life of 20 years and a salvage value at the end of that time of 10%. The land is expected to triple in value during this period. Property taxes are \$60 per thousand on a 50% valuation. Insurance is \$600,000 per year, repairs will be \$1,000,000 annually, and income at full occupancy will be \$1,200,000 per month. A real estate agency will be paid 3% of gross income. Income taxes are estimated at 20% of net income. The annual net profit is invested at a net return of 2%. Depreciation as a credit against income taxes is assumed to be straight line.

Calculate the value of each plan at the end of 20 years, state which plan would be more profitable.

18. Fourteen years ago a 12000-Kw steam electric plant was constructed at a cost of \$2,200 per Kw. Annual operating expenses have been \$310,000 to produce the annual demand of 54,000,000 Kw-hours. It is estimated that the annual operating expenses and demand for current will continue at the current level. The original estimate of a 20-year life with a 5% salvage value at that time is still correct. The present salvage value is estimated as \$7,500,000.

The company is contemplating the replacement of the old steam plant with a new 12000-Kw diesel plant. The old plant can be sold now for \$750,000, while the new diesel plant will cost \$2,450 per Kw to construct. The diesel plant will have a life of 25 years with a salvage value of 10% at the end of that time and will cost \$230,000 annually to operate. Annual taxes and insurance will be 2.3% of the first cost of either plant. Using an interest rate of 5%, determine whether the company is financially justified in replacing the old steam plant now.

19. The yearly requirement of a manufacturer is 1000 units of a part that is used at a uniform rate throughout the year. The machine set-up cost per lot is \$400. Production cost is \$52.00 per unit. Interest, insurance and taxes are estimated at 12% on average inventory. The cost of storing the parts is estimated at \$8.00 per unit per year. Calculate the economic lot size.

6.14 Solutions to End-of-Chapter Exercises

1.

$$850 \times (1 + 0.043) = \$886.55$$

or with Excel =FV(0.043,1,0,-850,0)= 886.55

2.

$$200 \times (1 + 0.045)^5 = \$249.24$$

or with Excel =FV(0.045,5,0,-200,0) = 249.24

3.

$$560 \times (1 + (0.037/12) \times 2) = \$563.45$$

or with Excel =FV(0.037/12,1*2,0,-560,0)=563.46

4.

$$1500 \times \left(1 + \frac{0.05}{4}\right)^{7 \times 4} = 1500 \times (1.0125)^{28} = \$2,123.99$$

or with Excel =FV(0.05/4,7*4,0,-1500)=2123.99

5.

$$3200 \times (1 + 0.07)^{-2} = \$2,795$$

or with Excel =PV(0.07,2,0,3200) = -2,795

6.

$$8000(1.01)^{22} + 1000(1.01)^{20} + 750(1.01)^{16} - 1200(1.01)^{11} = x$$

$$8000(1.24472) + 1000(1.22019) + 750(1.17258) - 1200(1.11567) = x$$

$$9957.76 + 1220.19 + 879.43 - 1338.80 = x$$

$$x = 10718.58$$

and with Excel

$$\begin{aligned} &= \text{FV}(0.04/4, (5*4+2), 0, -8000, 0) + \text{FV}(0.04/4, 5*4, 0, -1000, 0) + \text{FV}(0.04/4, 4*4, 0, -750, 0) \\ &+ \text{FV}(0.04/4, (2*4+3), 0, 1200, 0) = 10,718.55 \end{aligned}$$

7.

$$1500(1.0125)^{12} + 1500(1.0125)^8 + 2000(1.0125)^4$$

$$1500(1.16075) + 1500(1.10449) + 2000(1.05095)$$

$$1741.125 + 1656.735 + 2101.90 = 5499.76$$

Cost of new machine	10,000
Trade-in	750
Net cost	9250

$$9250 - 5499.76 = 3750.24$$

8.

$$P_n = P_0(1 + i)^n$$

$$P_0 = \frac{P_n}{(1 + i)^n} = P_n(1 + i)^{-n}$$

$$P_0 = 50000(1 + 0.15)^{-26} = 50000(0.67902) = \$33,951$$

9.

$$x(1.015)^{26} + x(1.015)^{24} + x(1.015)^{20} = 50000$$

$$x(1.47271) + x(1.42950) + x(1.34686) = 50000$$

$$4.25x = 50000$$

$$x = 11,765$$

10.

$$P_{-n} = P_0(1 + i)^{-n}$$

$$P_{-n1} = 1000(1 + 1.05)^{-2} = 1000(0.90703) = 907.03$$

$$P_{-n2} = 1500(1 + 1.05)^{-3} = 1500(0.86384) = 1295.22$$

$$P_{-n3} = 1800(1 + 1.05)^{-4} = 1800(0.82270) = 1480.86$$

$$907.03 + 1295.22 + 1480.86 = 3683.11$$

11.

$$P_n = P_0(1 + i)^n = 1000(1.06)^5 = 1000(1.41852) = 1418.52$$

Therefore, interest earned is

$$1418.52 - 1000 = 418.52$$

12.

Assumptions:

1. Purchase date is selected as our valuation date
2. Interest of 7% is compounded annually.

Machine A:

$$\begin{aligned} 36500 + 1500(1.07)^{-5} + 2000(1.07)^{-10} &= 36500 + 1500(0.71299) + 2000(0.50835) \\ &= 36500 + 1069.50 + 1016.70 = 38586.20 \end{aligned}$$

Machine B:

$$36300 + 3800(1.07)^{-9} = 36300 + 3800 \times 0.54393 = 38367$$

Savings by choosing Machine B,

$$38586 - 38367 = 219$$

13.

$$\begin{aligned} &1000(1.0175)^6 + 2000(1.0175)^4 - 300(1.0175) \\ &1000(1.10970) + 2000(1.07186) - 300(1.0175) = 2948.17 \\ &2948.17(1.02)^8 + 1600(1.02)^7 - 1200(1.02)^6 \\ &2948.17(1.17166) + 1600(1.14869) - 1200(1.12616) \\ &3454.25 + 1837.90 - 1351.39 = 3940.76 \end{aligned}$$

14.

$$P_n = P_0(1+i)^n = 2000(1.0375)^{10} = 2890.08$$

15.

$$\begin{aligned} P_n &= P_0(1+i)^n \\ \log P_n &= \log P_0 + n \log(1+i) \end{aligned}$$

Solving for n we obtain:

$$n = \frac{\log P_n - \log P_0}{\log(1+i)} = \frac{\log 3000 - \log 2000}{\log(1.0375)} = \frac{3.477 - 3.301}{0.016} = \frac{0.176}{0.016} = 11$$

Therefore, it will take 11 semiannual periods or 5.5 years.

16.

$$3600 + 4i(3600) = [105 + 3i(105)] + [110 + 2i(110)] + [110 + i(110)] + 100 + 3800$$

This yields $i = 4.15\%$. Since our approximation has given insufficient weight to receipts, the true investment rate is higher than this. Let us try an interest rate of 4.3% .

The amount y required to yield this rate is

$$3600(1.043)^4 - 105(1.043)^3 + 110(1.043)^2 + 110(1.043) = y$$

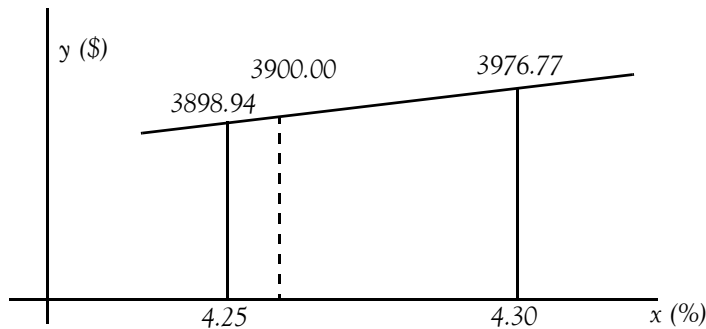
$$y = 3906.77$$

But this value is higher than $100 + 3800 = 3900$ which is the amount the investor received at the end of the period. Therefore, the interest rate must be slightly lower. Let us try an interest rate of 4.25% .

$$3600(1.0425)^4 - 105(1.0425)^3 + 110(1.0425)^2 + 110(1.0425) = y$$

$$y = 3898.94$$

We will use the graph below to apply straight-line interpolation



$$x_i = \frac{1}{\text{slope}} \times (y_i - y_1) + x_1 = \frac{1}{m} \times (y_i - y_1) + x_1$$

$$x|_{y=3900.00} = \frac{1}{\text{slope}} \times (3900.00 - 3898.94) + 0.0425 = \frac{1}{\frac{3906.77 - 3898.94}{0.0430 - 0.0425}} \times 1.06 + 0.0425$$

$$= \frac{0.0005}{7.83} \times 1.06 + 0.0425 = 0.0426$$

Therefore, the actual interest rate is 4.26% . This is the semiannual rate. The annual rate would be 8.52% .

17.

Plan A:

$$FV = \text{Payment} \times \frac{(1 + \text{interest})^n - 1}{\text{interest}}$$

The lump sum at the end of 20 years will be

$$\begin{aligned} \text{Lump sum} &= 200,000,000 + (200,000,000 \times 0.015) \times \frac{(1 + 0.02)^{20} - 1}{0.02} \\ &= 200,000,000 + 3,000,000 \times 24.2974 = \$272,892,100 \end{aligned}$$

Plan B:

	PP (Purchase Price, i.e., Initial Investment)	SV (Salvage Value)	PP-SV
Apartment Units	\$160,000,000	10%×160,000,000 = \$16,000,000	\$160,000,000–16,000,000 =\$144,000,000
Land	\$40,000,000	\$120,000,000	-80,000,000
Total	\$200,000,000	\$136,000,000	\$64,000,000

$$\text{Taxable Income} = \text{Gross Revenue} - (\text{Operating Expenses} + \text{Depreciation})$$

where

$$\text{Gross Revenue} = 12 \text{ months} \times 1,200,000/\text{month} \times 0.9 \text{ occupancy} = \$12,960,000$$

Operating Expenses

$$= \text{Real Estate Fee at } 3\% \text{ of } \$12,960,000 + \text{Property Taxes} + \text{Insurance} + \text{Repairs}$$

$$= 0.03 \times 12,960,000 + \frac{60}{1000} \times (200,000,000 \times 0.50) + 600,000 + 1,000,000 = \$7,988,800$$

Depreciation is allowable on the apartment units only and it will subject the investor to a capital gains tax at the end of 20 years. Thus,

$$\text{Straight Line Depreciation} = \frac{160,000,000 - 16,000,000}{20} = \$7,200,000$$

and

$$\begin{aligned} \text{Total credits against income taxes} &= \text{Total Operating Expenses} + \text{Depreciation} \\ &= \$7,988,800 + \$7,200,000 = \$15,188,800 \end{aligned}$$

Since total credits of \$15,188,800 exceed the gross revenue of \$12,960,000, there would be no annual income tax and the amount available for investment would be

Chapter 6 Mathematics of Finance and Economics

$$\begin{aligned}\text{Amount Available for Investment} &= \text{Gross Revenue} - \text{Operating Expenses} \\ &= \$12,960,000 - \$7,988,800 = \$4,971,200\end{aligned}$$

If this amount is invested for a return of 2%, the lump sum available at the end of 20 years would be

$$\begin{aligned}\text{Lump Sum} &= \text{Amount Available for Investment} + \text{Salvage Value} \\ &= \$4,971,200 \times \frac{(1 + 0.02)^{20}}{0.02} + 136,000,000 \\ &= \$4,971,200 \times 24.2974 + 136,000,000 = \$256,787,100\end{aligned}$$

The lump sum in Plan A was found to be \$272,892.10. Therefore, before capital gains tax, Plan A is more profitable by $\$272,892,100 - \$256,787,100 = \$16,105,000$.

18.

For convenience, we tabulate the given data as shown below.

	Electric Plant	Diesel Plant
Plant Capacity	12,000 Kw	12,000 Kw
Investment per Kw	\$2,200	\$2,450
Original Investment (PP)	26,400,000	29,400,000
Estimated Life	20 years	25 years
End of Life Salvage Value (ESV)	5%PP=1,320,000	10%PP=2,940,000
Present Age of Investment	14 years	N/A
Remaining Life (RL)	6 years	N/A
Present Salvage Value (PSV)	7,500,000	N/A
Annual Expenditures (AE)		
Operating Expenses	3,100,000	2,300,000
Taxes and Insurance	2.3%PP=607,200	2.3%PP=676,200
Total Annual Expenditures	3,707,200	2,976,200

We will use (6.39), Page 6–69, to compute the annual costs for each plant, i.e.,

$$\text{Total Annual Cost} = (PP - SV) \frac{i}{(1 + i)^n - 1} + PPi + AE$$

where for the electric plant $n = 20 - 14 = 6$, and the diesel plant $n = 25$

For the existing (electric) plant, let $SA =$ Annual Savings if existing plant is replaced

Then,

$$\begin{aligned}
 SA &= (PSV - ESV) \frac{i}{(1+i)^n - 1} + PSVi + AE \\
 &= (7,500,000 - 1,320,000) \frac{0.05}{(1+0.05)^6} + 7,500,000 \times 0.05 + 3,707,200 \\
 &= 6,180,000 \times 0.14702 + 375,000 + 3,707,200 = \$4,990,768
 \end{aligned}$$

For the diesel plant

$$\begin{aligned}
 AC &= (PP - SV) \frac{i}{(1+i)^n - 1} + PPi + AE \\
 &= (29,400,000 - 2,940,000) \frac{0.05}{(1+0.05)^{25}} + 29,400,000 \times 0.05 + 2,976,200 \\
 &= 26,460,000 \times 0.020952 + 1,470,000 + 2,976,200 = \$5,000,602
 \end{aligned}$$

From a practical point of view, the economics indicate a standoff due to the small differential between values of SA and AC.

19.

Let us denote the economic lot size as x where lot size is the number of units manufactured with one machine set-up. Then

$$\text{Number of lots per year} = \text{Number of machine setups} = \frac{1000}{x}$$

and

$$\text{Annual cost for all setups} = \$400 \times \frac{1000}{x} = \frac{400,000}{x}$$

Next,

$$\text{Average number of units in storage} = \text{Average inventory} = \frac{x}{2}$$

and

$$\begin{aligned}
 \text{Inventory cost per year} &= \text{Production cost per unit} \times \text{Average inventory} \\
 &= \$52.00 \times \frac{x}{2} = \$26.00x
 \end{aligned}$$

Also,

$$\text{Interest, Insurance, and Taxes on Average Inventory} = 12\% \times \$26.00x = \$3.12x$$

and

Chapter 6 Mathematics of Finance and Economics

$$\text{Storage costs} = \$8.00 \times \frac{x}{2} = \$4.00x$$

The annual production costs are

$$AC = \$52.00 \times 1000 = \$52,000$$

Therefore, total annual costs

$$y = (3.12 + 4.00)x + \frac{400,000}{x} + 52,000 = 7.12x + \frac{400,000}{x} + 52,000$$

From (6.40), Page 6–71,

$$y = ax + \frac{b}{x} + c$$

Therefore,

$$a = 7.12 \quad b = 400,000$$

and from (6.41), Page 6–72, the economic lot size is

$$x = \sqrt{\frac{b}{a}} = \sqrt{\frac{400,000}{7.12}} = 237 \text{ units}$$

This lot size would require

$$\frac{1000}{237} = 4.22 \text{ lots}$$

and rounding this number to 4 lots, we find that the 1000 units should be manufactured in 4 lots of 250 units each lot.

Chapter 7

Depreciation, Impairment, and Depletion

This chapter introduces the concept of depreciation as applied to business and financial applications. The different methods of depreciation are discussed and are illustrated with several examples.

7.1 Depreciation Defined

A *capital asset* is a piece of equipment, a building, or a vehicle for use in our business. When we buy a capital asset, we expect to use it for several years. The cost of these capital assets should be written off over the same period of time we expect them to earn income for us. Therefore, we must spread the cost over several tax years and deduct part of it each year as a business expense. As these assets wear out, lose value, or become obsolete, we recover our cost as a business expense. This method of deducting the cost of business property is called *depreciation*.

The concept of depreciation is really pretty simple. For example, let's say we purchase a truck for our business. The truck loses value the minute we drive it out of the dealership. The truck is considered an operational asset in running our business. Each year that we own the truck, it loses some value, until the truck finally stops running and has no value to the business. Measuring the loss in value of an asset is known as *depreciation*.

Depreciation is considered an expense and is listed in an income statement under expenses. In addition to vehicles that may be used in our business, we can depreciate office furniture, office equipment, any buildings we own, and machinery we use to manufacture products.

There are two types of property that can be depreciated:

- I Real Property – Buildings (real estate) or anything else built on or attached to land. However, land can never be depreciated.
- II Personal Property – Cars, trucks, equipment, furniture, or almost anything that is not real property.

As stated above, land is not considered an expense; therefore, it cannot be depreciated. This is because land does not wear out like vehicles or equipment.

To find the annual depreciation cost for our assets, we need to know the initial cost of the assets. We also need to determine how many years we think the assets will retain some value for our business. In the case of the truck, it may only have a useful life of ten years before it wears out and loses all value.

Chapter 7 Depreciation, Impairment, and Depletion

7.1.1 Items that Can Be Depreciated

We can depreciate many different kinds of property, for example, machinery, buildings, vehicles, patents, copyrights, furniture and equipment. We can depreciate property if it meets all the following requirements:

1. It must be used in business or held to produce income
2. It must be expected to last more than one year
3. It must be something that wears out, decays, gets used up, becomes obsolete, or loses its value from natural causes.

7.1.2 Items that Cannot Be Depreciated

We cannot depreciate the following items

1. Property placed in service and disposed of in the same year
2. Inventory
3. Land
4. Repairs and replacements that do not increase the value of our property, make it more useful, or lengthen its useful life.
5. Items which do not decrease in value over time are not depreciated. These include antiques over 100 years old, expensive solid wood furniture such as that made of oak or walnut, and fine china.

7.1.3 Depreciation Rules

Under the claims statute, an individual is paid the actual value of an item at the time of its loss. Certainly, one should not expect to be reimbursed more than an item was worth when it was lost or destroyed beyond repair. That would put us in a better position than we were in before the incident. For example, if we owned a five year old computer, we would not expect someone to pay us for a brand new computer. Although our computer may have been working, it was still a used computer. Accordingly, we should be paid for the actual value of our used item. We can then use the money to buy a similar used item, or, we can apply the money toward the cost of a newer item if we choose.

The *actual value* of an item is the current replacement cost minus depreciation, if any. Current replacement cost takes inflation and local unavailability into account. If the item costs more now than when we bought it, or is not available in the local area, we provide the current replacement price of the item where it can be found. Therefore, the actual value is a fair measure of what a claimant should be paid. The actual value rule in effect does pay us the replacement cost. One can obtain insurance if he/she wants full replacement cost coverage.

7.1.4 When Depreciation Begins and Ends

We begin to depreciate our property:

1. When we place it in service for use in our trade or business
2. For the production of income.

We stop depreciating property either:

1. When we have fully recovered our cost or other basis, or
2. When we retire it from service, whichever happens first

For depreciation purposes, we place property in service when it is ready and available for a specific use, whether in a trade or business, the production of income, a tax-exempt activity, or a personal activity.

Example 7.1

Suppose we bought a home and used it as our personal home several years before we converted it to rental property. Although its specific use was personal and no depreciation was allowable, we placed the home in service when we began using it as our home. We can claim a depreciation deduction in the year that we converted it to rental property because its use changed to an income-producing use at that time.

Example 7.2

We bought a computer for our tax preparation business after April 15. We take a depreciation deduction for the computer for that year because it was ready and available for its specific use.

7.1.5 Methods of Depreciation

Our annual depreciation expense depends on the depreciation method that we choose. These are:

1. **Straight-line depreciation**
2. **Sum of the Year's Digits**
3. **Fixed Declining Balance**
4. **General Declining Balance (125%, 150%, 200%)**
5. **Variable-Declining Balance**

Chapter 7 Depreciation, Impairment, and Depletion

6. Units of Production

We can choose a different method of depreciation for each asset that we own. The method that we choose depends on the type of property and the needs of our business. For certain types of assets, an accelerated method of depreciation will be more appropriate. For others, we may decide to keep it simple by depreciating the same amount each year.

We will discuss each in the following subsections.

7.1.5.1 Straight-Line (SL) Depreciation Method

Straight-line depreciation is considered to be the most common method of depreciating assets. To compute the amount of annual depreciation expense using the straight-line method requires two numbers:

1. The initial cost of the asset less its salvage value
2. Its estimated useful life.

In other words, straight line depreciation is computed with the use of the following formula:

$$\text{Annual depreciation expense} = \frac{\text{Cost} - \text{Salvage value}}{\text{Estimated useful life}} \quad (7.1)$$

Example 7.3

Suppose that we purchase a truck for \$30,000 and we expect to use it in our business for 10 years. Using the straight-line method for determining depreciation, and assuming that the salvage value is \$1,000 we would divide the initial cost of the truck by its useful life.

The \$30,000 becomes a depreciation expense that is reported on our income statement under operation expenses at the end of each year. The useful life is 10 years; therefore, the depreciation is

$$\text{Annual depreciation expense} = \frac{30000 - 1000}{10} = 2900 \quad (7.2)$$

Thus, the depreciation is \$2,900 per year for the next ten years.

We can use Excel's **SLN** function to compute the straight line depreciation of an asset for one period. It's syntax is

`=SLN(Cost,Salvage_Value,Life)`

For this example, `=SLN(30,000,1000,10)` returns \$2,900

7.1.5.2 Sum of the Years Digits (SYD) Method

Sum of the Years Digits (SYD) is a method of accelerated depreciation. Accelerated methods assume that a fixed asset loses a greater proportion of its value in the early years of use. Consequently, the amount of depreciation expense is higher at the beginning of the useful life, and declines over time.

Some kinds of property – cars, for instance – are more efficiently productive in the initial years of use. Over time, they become more costly to maintain. Rising maintenance and repair costs tend to offset the lower depreciation expense in later years. Property with this characteristic makes a good candidate for an accelerated method.

SYD is computed with the use of the following formula:

$$\text{SYD} = \frac{(\text{Cost} - \text{Salvage value}) \times (\text{Useful life} - \text{Period} + 1)}{(\text{Useful life}) \times (\text{Useful life} + 1) / 2} \quad (7.3)$$

where period is the year in which depreciation is computed, and we must use the same units as useful life, that is, years, months, and so on.

The expression $\text{Useful life} - \text{Period} + 1$ in the numerator of (7.3) indicates the life of the depreciation in the first period decreased by 1 in each subsequent period. This reflects the declining pattern of depreciation over time. The expression $(\text{Useful life}) \times (\text{Useful life} + 1) / 2$ in the denominator is equal to the sum of the digits, that is, if we denote $\sum x$ the sum of the years digits and the useful life as n , then

$$\sum x = 1 + 2 + 3 + \dots + (n - 2) + (n - 1) + n = \frac{n(n + 1)}{2}$$

We must remember that (7.3) yields the depreciation for some particular year, not the total depreciation up to that year. To compute the total depreciation to the end of the n th year, we use the equation

$$\text{SYD}_{\text{Total to } n\text{th year}} = \frac{\text{Period} \times (\text{Cost} - \text{Salvage value}) \times (2 \times \text{Useful life} - \text{Period} + 1)}{(\text{Useful life}) \times (\text{Useful life} + 1)} \quad (7.4)$$

Example 7.4

Suppose that we purchase a truck for \$30,000 and we expect to use it in our business for 10 years. Using the SYD method for determining depreciation, and assuming that the salvage value is \$1,000, for the first period (first year) the depreciation is:

Chapter 7 Depreciation, Impairment, and Depletion

$$\text{SYD} = \frac{(30000 - 1000) \times (10 - 1 + 1) \times 2}{(10) \times (10 + 1)} = \frac{580000}{110} = 5272.73 \quad (7.5)$$

Thus, the depreciation for the first year is \$5,272.73

We can use Excel's SYD function to compute the SYD depreciation of an asset for any period. It's syntax is

=SYD(Cost,Salvage_Value,Life,Period)

For this example,

=SYD(30,000,1000,10,1) returns \$5,272.73

For the seventh period (seventh year) the depreciation is found by using the Excel formula

=SYD(30,000,1000,10,7) and this returns \$2,109.09

7.1.5.3 Fixed-Declining Balance (FDB) Method

The *Fixed-Declining Balance* (FDB) method computes depreciation at a fixed rate. For any period other than the first and the last, this method uses the following formula:

$$\text{FDB} = (\text{Cost} - \text{Total depreciation from prior periods}) \times \left(1 - \left(\frac{\text{Salvage value}}{\text{Cost}} \right)^{\frac{1}{\text{Life}}} \right) \quad (7.6)$$

For the first period, FDB uses the following formula:

$$\text{FDB}_{\text{1st year}} = \text{Cost} \times \left(1 - \left(\frac{\text{Salvage value}}{\text{Cost}} \right)^{\frac{1}{\text{Life}}} \right) \times \frac{\text{Month}}{12} \quad (7.7)$$

where Month is the number of months in the first year.

For the last period, FDB uses the following formula:

$$\begin{aligned} \text{FDB}_{\text{Last year}} &= (\text{Cost} - \text{Total depreciation from prior periods}) \\ &\times \left(1 - \left(\frac{\text{Salvage value}}{\text{Cost}} \right)^{\frac{1}{\text{Life}}} \right) \times \frac{12 - \text{Month}}{12} \end{aligned} \quad (7.8)$$

Example 7.5

Suppose that we purchase a truck for \$30,000 in June and we expect to use it in our business for 10 years. Using the FDB method for determining depreciation, and assuming that the salvage value is \$1,000, for the first period (first year) the depreciation is computed using the formula of (7.7) as follows:

$$\text{FDB}_{1\text{st year}} = 30000 \times \left(1 - \left(\frac{1000}{30000} \right)^{\frac{1}{10}} \right) \times \frac{7}{12} = \frac{210000}{12} \times (1 - 0.712) = 5040 \quad (7.9)$$

We can use Excel's DB function to compute the fixed-declining balance method of depreciation. It's syntax is

`=DB(Cost,Salvage_Value,Life,Period,Month)`

If the Month is omitted, it is assumed to be 12

For this example,

`=DB(30000,1000,10,1,7)` returns \$5,040

The depreciation for the second through the tenth year is computed with the formula of (7.8). For instance, the depreciation for the second year is found from

$$\text{FDB}_{2\text{nd year}} = (30000 - 5040) \times \left(1 - \left(\frac{1000}{30000} \right)^{\frac{1}{10}} \right) = 24960 \times (1 - 0.712) = 7188.48 \quad (7.10)$$

We can verify this result with Excel's DB function as follows:

`=DB(30000,1000,10,2,7)` and this returns \$7,188.48 .

The depreciation for the last period is computed with the formula of (7.8). But this formula requires that we compute the depreciation for all previous years. For convenience, we use Excel's DB function and we construct the following table:

Period	Depreciation
1	\$5,040.00
2	\$7,188.48
3	\$5,118.20
4	\$3,644.16
5	\$2,594.64
6	\$1,847.38
7	\$1,315.34
8	\$936.52
9	\$666.80
10	\$474.76
11	\$140.85

Chapter 7 Depreciation, Impairment, and Depletion

The sum of the first 10 periods on this table is approximately \$28,826 and this number represents the total depreciation taken in previous years. We will verify the depreciation of the eleventh year using the formula of (7.8). Thus, the depreciation for the eleventh year, with only 5 months is found from

$$\begin{aligned} \text{FDB}_{11\text{th year}} &= (30000 - 28826) \times \left(1 - \left(\frac{1000}{30000} \right)^{\frac{1}{10}} \right) \times \frac{12-7}{12} \\ &= 1174 \times (1 - 0.712) \times \frac{5}{12} = 140.85 \end{aligned} \quad (7.11)$$

7.1.5.4 The 125%, 150%, and 200% General Declining Balance Methods

The *125%*, *150%*, and *200%* *General Declining Balance (GDB)* methods are also accelerated depreciation methods, and are used when the productivity of an asset is expected to be greater during its early years of use. These methods use the following formula:

$$\text{GDB} = (\text{Cost} - \text{Total depreciation from prior periods}) \times \left(\frac{\text{Factor}}{\text{Life of asset}} \right) \quad (7.12)$$

where Factor is the rate at which the balance declines.

For the *125%* declining balance, Factor = 1.25

For the *150%* declining balance, Factor = 1.50

For the *200%* declining balance, Factor = 2.00

The *200%* declining balance is also referred to as *Double Declining Balance (DDB)*.

The total depreciation taken in prior years cannot exceed the salvage value.

Example 7.6

Suppose that we purchase a truck for \$30,000 and we expect to use it in our business for 10 years. Using the DDB method for determining depreciation, and assuming that the salvage value is \$1,000, for the first period (first year) the depreciation is computed using the formula of (7.11) as follows:

$$\text{DDB} = 30000 - 0 \times \left(\frac{2}{10} \right) = 6000 \quad (7.13)$$

We can use Excel's DDB function to compute the double-declining balance method of deprecia-

tion. It's syntax is

`=DDB(Cost,Salvage_Value,Life,Period,Factor)`

If the **Factor** is omitted, it is assumed to be 2

Note: Although Excel requires that **Salvage_Value** is specified, it produces the same result no matter what value we specify. Thus, with a salvage value of 0, 1000, or 2000, we get the same answer as shown below.

For this example, `=DDB(30000,0,10,1,2)` returns \$6,000

`=DDB(30000,1000,10,1,2)` returns \$6,000 also.

`=DDB(30000,2000,10,1,2)` returns \$6,000 again.

If the salvage value of an asset is relatively low, the DDB method may not fully depreciate the asset by the end of the estimated useful life. The spreadsheet below shows that at the end of the 10 year period, only \$26,779 has been depreciated.

Period	Depreciation
1	\$6,000.00
2	\$4,800.00
3	\$3,840.00
4	\$3,072.00
5	\$2,457.60
6	\$1,966.08
7	\$1,572.86
8	\$1,258.29
9	\$1,006.63
10	\$805.31
Sum	\$26,778.77

We can use the Variable Declining Balance (VDB) method which always fully depreciates the asset within the estimated life. This method is discussed in the next subsection.

7.1.5.5 The Variable Declining Balance Method

The *Variable Declining Balance* (VDB) method computes the depreciation of an asset using the DDB method, and switches to SL method when the SL depreciation is greater than the DDB method or any other variable–rate declining balance method.

We can use Excel's VDB function to compute the double–declining balance method of depreciation. It's syntax is

`=DDB(Cost,Salvage_Value,Life,Start_Period,End_Period,Factor,No_Switch)`

where:

Cost = Initial cost of the asset

Salvage_Value = The value of the asset at the end of the depreciation period

Chapter 7 Depreciation, Impairment, and Depletion

Life = The number of periods over which the asset is depreciated

Start_Period = The starting period for which we want to calculate the depreciation

End_Period = The ending period for which we want to calculate the depreciation

Factor = The rate at which the balance declines. If the **Factor** is omitted, it is assumed to be 2.

No_Switch = A logical value (**TRUE** or **FALSE**) that specifies whether to switch to straight line depreciation when SL depreciation is greater than the declining balance method. If **No_Switch** is **TRUE**, Excel does not switch to SL depreciation even when the SL depreciation is greater than the declining balance depreciation. If **No_Switch** is **FALSE** or omitted, Excel switches to SL depreciation when the SL depreciation is greater than the declining balance depreciation.

Example 7.7

Suppose that we purchase a truck for \$30,000 and we expect to use it in our business for 10 years. Using the 150% declining balance method for determining depreciation, and assuming that the salvage value is \$1,000, the Excel formulas below compute the depreciation for each year. We observe that the switch to SL depreciation occurs in the sixth period.

Period	Formula	Depreciation
First	=VDB(10000,1000,10,0,0.875,1.5)	\$1,312.50
Second	=VDB(10000,1000,10,0.875,1.875,1.5)	\$1,303.13
Third	=VDB(10000,1000,10,1.875,2.875,1.5)	\$1,107.66
Forth	=VDB(10000,1000,10,2.875,3.875,1.5)	\$941.51
Fifth	=VDB(10000,1000,10,3.875,4.875,1.5)	\$800.28
Sixth	=VDB(10000,1000,10,4.875,5.875,1.5)	\$689.74
Seventh	=VDB(10000,1000,10,5.875,6.875,1.5)	\$689.74
Eighth	=VDB(10000,1000,10,6.875,7.875,1.5)	\$689.74
Ninth	=VDB(10000,1000,10,7.875,8.875,1.5)	\$689.74
Tenth	=VDB(10000,1000,10,8.875,9.875,1.5)	\$689.74
Last	=VDB(10000,1000,10,9.875,10,1.5)	\$86.22
Total Depreciation	=Initial cost minus Salvage value	\$9,000.00

7.1.5.6 The Units of Production Method

The *Units of Production* (UOP) method allocates depreciation expenses according to actual physical usage. We can use this method for assets that have limited productive capacity. It is an appropriate method to use when the usage of a fixed asset varies from year-to-year. Unlike the SL method which treats depreciation as a fixed cost, the UOP method treats depreciation as a variable cost.

As an example, an asset that is a good candidate for this depreciation method is a large delivery truck that it can be driven for 250,000 miles before expensive maintenance costs render it useless. With such an asset, it is immaterial whether the miles are driven in a few or many years.

For the UOP depreciation method, we use the following formula:

$$\text{Annual depreciation} = \frac{\text{Initial cost} - \text{Salvage value}}{\text{Total estimated units of output}} \quad (7.14)$$

Example 7.8

Suppose that on the first day of the year we bought a machine for \$100,000 and we estimate that this machine will have a good working life of 5 years. At the end of 5 years, we expect to sell it for \$20,000. We expect that the machine will produce 180,000 units the first year, 160,000 units the second year, 140,000 units the third year, 120,000 units the fourth year, and 100,000 units the fifth year for a total of 700,000 during the 5-year period. We compute the annual depreciation with the formula of (7.14).

$$\text{Annual depreciation} = \frac{\$100000 - \$20000}{800000 \text{ units}} = \$0.1 \text{ per unit}$$

The depreciation schedule is shown on the table below.

Year Ended	Balance	Units Produced	Annual Depreciation \$0.1xunits produced	Carrying Cost
1/1/2003				\$100,000
12/31/2003	\$100,000	200,000	\$20,000	\$80,000
12/31/2004	\$80,000	180,000	\$18,000	\$62,000
12/31/2005	\$62,000	160,000	\$16,000	\$46,000
12/31/2006	\$46,000	140,000	\$14,000	\$32,000
12/31/2007	\$32,000	120,000	\$12,000	\$20,000

We observe that the carrying cost at the end of the year 2007 is exactly \$20,000 which is the salvage value of the machine.

7.1.5.7 Depreciation Methods for Income Tax Reporting

It is beyond the scope of this text to provide detailed description of the depreciation methods set forth by the Internal Revenue Service (IRS). This discussion is for instructional purposes. The reader must refer to the latest IRS publications. We will briefly discuss the following methods:

1. The Accelerated Cost Recovery System (ACRS)
2. The Modified Accelerated Cost Recovery System (MACRS)
3. Section 179

Each of these is an accelerated depreciation method, and the method used depends on the type of property and the year it was placed in service. These methods are not acceptable by the so-called Generally Accepted Accounting Principles (GAAP). For a detailed discussion of these methods,

Chapter 7 Depreciation, Impairment, and Depletion

the interested reader may refer to IRS Publication 946: How To Depreciate Property.

The Accelerated Cost Recovery System (ACRS)

The *Accelerated Cost Recovery System* (ACRS) is an accounting method for calculating the depreciation of tangible assets on the basis of the estimated-life classifications into which the assets are placed. ACRS was initiated by the Economic Recovery Act of 1981. The intent was to make investments more profitable by sheltering large amounts of income from taxation during the early years of an asset's life. The initial law established classifications of 3, 5, 10, and 15 years; these classifications were subsequently modified in order to reduce depreciation and increase the government's tax revenues. The classification into which an asset is placed determines the percentage of the cost potentially recoverable in each year.

The Modified Accelerated Cost Recovery System (MACRS)

The *Modified Accelerated Cost Recovery System* (MACRS) is a depreciation method in which assets are classified according to a prescribed life or recovery period that bears only a rough relationship to their expected economic lives. MACRS represents a 1986 change to the Accelerated Cost Recovery System (ACRS) that was instituted in 1981. The depreciation rates in MACRS are derived from the double-declining-balance method of depreciation.

MACRS consists of two depreciation systems, the *General Depreciation System* (GDS) and the *Alternative Depreciation System* (ADS). These systems provide different methods and recovery periods to use in computing depreciation deductions.

To depreciate property under MACRS, the taxpayer must be familiar with the following terms:

Class Life: A number of years that establishes the property class and recovery period for most types of property under the General Depreciation System (GDS) and the Alternative Depreciation System (ADS).

Listed Property: Passenger automobiles, or any other property used for transportation, property of a type generally used for entertainment, recreation, or amusement, computers and their peripheral equipment (unless used only at a regular business establishment and owned or leased by the person operating the establishment), and cellular telephones or similar telecommunications equipment.

Nonresidential real property: Most real property other than residential rental property.

Placed in service: Ready and available for a specific use whether in a trade or business, the production of income, a tax-exempt activity, or a personal activity.

Property class: A category for property under MACRS. It generally determines the depreciation method, recovery period, and convention.

Recovery period: The number of years over which the basis (cost) of an item of property is recovered.

Residential rental property: Real property, generally buildings or structures, if 80% or more of its annual gross rental income is from dwelling units.

Section 1245 property: Property that is or has been subject to an allowance for depreciation or amortization. Section 1245 property includes personal property, single purpose agricultural and horticultural structures, storage facilities used in connection with the distribution of petroleum or primary products of petroleum, and railroad grading or tunnel bores.

Section 1250 property: Real property (other than section 1245 property) which is or has been subject to an allowance for depreciation.

Tangible property: Property we can see or touch, such as buildings, machinery, vehicles, furniture, and equipment.

Tax-exempt: Not subject to tax.

Our use of either the General Depreciation System (GDS) or the Alternative Depreciation System (ADS) to depreciate property under MACRS determines what appreciation method and recovery period we use. Generally, we must use GDS unless we are specifically required by law to use ADS or we elect to use it.

We must use ADS for the following property:

1. Listed property used 50% or less for business.
2. Any tangible property used predominantly outside the United States during the year.
3. Any tax-exempt use property.
4. Any tax-exempt bond-financed property.
5. All property used predominantly in a farming business and placed in service in any tax year during which an election not to apply the uniform capitalization rules to certain farming costs is in effect.
6. Any imported property covered by an executive order of the President of the United States.
7. Any property for which an election to use ADS was made (discussed next).

Electing ADS

Although our property may qualify for GDS, we can elect to use ADS. The election generally must cover all property in the same property class that we placed in service during the year. However, the election for residential rental property and nonresidential real property can be made on a property-by-property basis. Once we make this election, we can never revoke it.

The following is a list of the nine property classes under GDS and examples of the types of property included in each class.

Chapter 7 Depreciation, Impairment, and Depletion

3-year property:

- a) Tractor units for over-the-road use
- b) Any race horse over 2 years old when placed in service
- c) Any other horse over 12 years old when placed in service
- d) Qualified rent-to-own property*

5-year property:

- a) Automobiles, taxis, buses, and trucks
- b) Computers and peripheral equipment
- c) Office machinery (such as typewriters, calculators, and copiers)
- d) Any property used in research and experimentation
- e) Breeding cattle and dairy cattle
- f) Appliances, carpets, furniture, etc., used in a residential rental real estate activity

7-year property:

- a. Office furniture and fixtures (such as desks, files, and safes).
- b. Agricultural machinery and equipment.
- c. Any property that does not have a class life and has not been designated by law as being in any other class.

10-year property:

- a. Vessels, barges, tugs, and similar water transportation equipment
- b. Any single purpose agricultural or horticultural structure
- c. Any tree or vine bearing fruits or nuts

15-year property:

- a. Improvements made to land or added to it (such as shrubbery, fences, roads, and bridges).

* *Qualified rent-to-own property is property held by a rent-to-own dealer for purposes of being subject to a rent-to-own contract. It is tangible personal property generally used in the home for personal use. It includes computers and peripheral equipment, televisions, videocassette recorders, stereos, camcorders, appliances, furniture, washing machines and dryers, refrigerators, and other similar consumer durable property. Consumer durable property does not include real property, aircraft, boats, motor vehicles, or trailers.*

If some of the property we rent to others under a rent-to-own agreement is of a type that may be used by the renters for either personal or business purposes, we can still treat this property as qualified property as long as it does not represent a significant portion of our leasing property. But, if this dual-use property does represent a significant portion of our leasing property, we must prove that this property is qualified rent-to-own property.

b. Any retail motor fuels outlet* such as a convenience store.

20-year property:

This class includes farm buildings (other than single purpose agricultural or horticultural structures).

25-year property:

This class is water utility property, which is either of the following.

a. Property that is an integral part of the gathering, treatment, or commercial distribution of water, and that, without regard to this provision, would be 20-year property.

b. Any municipal sewer

Residential rental property:

This is any building or structure, such as a rental home (including a mobile home), if 80% or more of its gross rental income for the tax year is from dwelling units. A dwelling unit is a house or apartment used to provide living accommodations in a building or structure. It does not include a unit in a hotel, motel, inn, or other establishment where more than half the units are used on a transient basis. If we occupy any part of the building or structure for personal use, its gross rental income includes the fair rental value of the part we occupy.

Non-Residential real property:

This is Section 1250 property, such as an office building, store, or warehouse, that is neither residential rental property nor property with a class life of less than 27.5 years.

If our property is not listed above, we can determine its property class from the Table of Class Lives and Recovery Periods in Appendix B of IRS Publication 946. The property class is generally the same as the GDS recovery period indicated in the table.

To help us compute our deduction under MACRS, the IRS has established percentage tables that incorporate the applicable convention and depreciation method. We first determine the depreciation system (GDS or ADS), property class, placed-in service date, basis amount, recovery period, convention, and depreciation method that applies to our property. Then, we can compute the depreciation using the appropriate percentage table provided by the IRS.

* *Real property is a retail motor fuels outlet if it is used to a substantial extent in the retail marketing of petroleum or petroleum products (whether or not it is also used to sell food or other convenience items) and meets any one of the following three tests:*

- *It is not larger than 1,400 square feet.*
- *50% or more of the gross revenues generated from the property are derived from petroleum sales.*
- *50% or more of the floor space in the property is devoted to petroleum marketing sales.*

A retail motor fuels outlet does not include any facility related to petroleum and natural gas trunk pipelines.

Chapter 7 Depreciation, Impairment, and Depletion

Appendix A of the IRS Publication 946* contains three charts to help us find the correct percentage table to use. Presently, there are 20 tables (Table A-1 through Table A-20). Table 7-1 on the next page is IRS Table A-1, that is used with the 3-,5-,7-,10-,15-, and 20-Year Property Half Year Convention.

TABLE 7.1 3-,5-,7-,10-,15-, and 20-Year Property Half Year Convention

Year	Depreciation rate for recovery period					
	3-year	5-year	7-year	10-year	15-year	20-year
1	33.33%	20.00%	14.29%	10.00%	5.00%	3.750%
2	44.45	32.00	24.49	18.00	9.50	7.219
3	14.81	19.20	17.49	14.40	8.55	6.677
4	7.41	11.52	12.49	11.52	7.70	6.177
5		11.52	8.93	9.22	6.93	5.713
6		5.76	8.92	7.37	6.23	5.285
7			8.93	6.55	5.90	4.888
8			4.46	6.55	5.90	4.522
9				6.56	5.91	4.462
10				6.55	5.90	4.461
11				3.28	5.91	4.462
12					5.90	4.461
13					5.91	4.462
14					5.90	4.461
15					5.91	4.462
16					2.95	4.461
17						4.462
18						4.461
19						4.462
20						4.461
21						2.231

* IRS publications undergo frequent changes. The tables provided here are for instructional purposes only and therefore the reader must obtain the latest publications from IRS.

Example 7.9

Our company is a calendar-year corporation and has been in business for more than a year. On January 15, 2006, the company purchased and placed in service a computer and peripheral equipment for a total cost of \$10,000. On December 15, 2006, the company purchased and placed in service office furniture for a total cost of \$5,000.

For this example, Table 7-1, MACRS Half-Year Convention applies. The computer and peripheral equipment is a 5-year property. For the first year the rate is 20% and thus the depreciation for the first year is

$$\$10000 \times 0.20 = \$2000$$

The office furniture is a 7-year property, and for the first year the rate is 14.9%. Therefore, the depreciation for the first year is

$$\$5000 \times 0.1429 = \$714.50$$

For the year 2007 we use the same table but we use the percentages corresponding to the second year. Thus, the rate for the computer is 32% and the depreciation for the second year is

$$\$10000 \times 0.32 = \$3200$$

The rate for the office furniture is 24.49% and the depreciation for the second year is

$$\$5000 \times 0.2449 = \$1224.50$$

The same procedure is used for later years.

Section 179

Section 179 is a deduction allowed for up to the entire cost of certain depreciable business assets, other than real estate, in the year purchased, which we may be able to use as an alternative to depreciating the asset over its useful life. The annual limit in 2005 was \$25,000. We cannot use the Section 179 deduction to the extent that it would cause us to report a loss from our business.

Typically, if property for business has a useful life of more than one year, the cost must be spread across several tax years as depreciation with a portion of the cost deducted each year.

The total amount we could elect to deduct under Section 179 for 2001 could not be more than \$25,000. If we acquired and placed in service more than one item of qualifying property during the year 2001, we could allocate the Section 179 deduction among the items in any way, as long as the total deduction did not exceed \$25,000. It is not necessary that we had to claim the entire amount of \$25,000.

Chapter 7 Depreciation, Impairment, and Depletion

Example 7.10

In 2006, we bought and placed in service a \$25,000 machine and a \$2,500 computer for our business. We elected to deduct the entire \$2,500 for the computer and \$22,500 for the machine, a total of \$25,000 which is the maximum amount we could deduct. Our \$2,500 deduction for the computer completely recovered its cost and thus the basis for depreciation for the computer is zero. For the machine, since we deducted \$22,500, the basis for depreciation for the machine is

$$\$25,000 - \$22,500 = \$2,500$$

If the cost of our qualifying Section 179 property placed in service is over \$200,000, we must reduce the dollar limit (but not below zero) by the amount of over \$200,000.

Example 7.11

In 2006, we bought and placed in service a machine costing \$207,000. Because this cost is more than \$200,000, we should have reduced the dollar limit to

$$\$25,000 - \$7,000 = \$18,000$$

7.2 Impairments

An *impairment* or *decline of value*, as defined by the accounting community, is a loss of a significant portion of the usefulness of an asset through casualty, obsolescence, or lack of demand for the asset's services. When assets suffer a permanent impairment in value, an *impairment loss* is recorded. An impairment loss is reported as part of income from continuing operations.

The following formula is used to compute the impairment loss:

$$\text{Impairment loss} = \text{Book value} - \text{Fair market value} \quad (7.15)$$

where

Book value: = net dollar value at which an asset is carried on a company's balance sheet. For example an asset that was purchased for \$500,000 but on the books has depreciated \$300,000 has a book value of \$200,000. The book value should not be confused with the fair market value which is defined below.

Fair market value = present value of an asset, that is, the value of a similar asset in its present condition.

Example 7.12

Compute the impairment loss for each of the following assets:

- a. Machine A: Book value = \$10,000 , Fair market value: \$12,500
- b. Machine B: Book value = \$15,000 , Fair market value: \$15,000
- c. Machine C: Book value = \$27,500 , Fair market value: \$25000

Solution:

a. Machine A:

$$\text{Impairment loss} = 10000 - 12500 = -2500$$

The impairment loss is negative. Therefore, impairment for Machine A has not occurred.

b. Machine B:

$$\text{Impairment loss} = 15000 - 15000 = 0$$

The impairment loss is zero. Therefore, impairment for Machine B has not occurred.

c. Machine C:

$$\text{Impairment loss} = 27500 - 25000 = 2500$$

The impairment loss is positive. Therefore, impairment for Machine C has occurred and it is \$2,500 .

7.3 Depletion

All natural resources of commercial value can be divided into two broad categories. The first group consists of those resources in which nature reproduces, or replaces, the commodity that has been extracted; the second group consists of those resources in which the extracted commodity is not replaced. Thus, in the case of agricultural land, we can assume for valuation purposes that nature will continue indefinitely to yield her annual harvest, barring any event that will render the land infertile. On the other hand, in the case of an oil well, a timber tract, or a mine, the mineral that is extracted is not replaced, and there is consequently a definite limit to the wealth obtainable from the resource. Those in the latter group are designated “wasting” or “depleting” assets, by virtue of the fact that the asset “wastes away” or becomes depleted as it is exploited. In the case of several reproductive assets, the rate of extraction of the commodity must be regulated to coincide with the rate at which nature is capable of replacing it.

The following factors are used to determine the depletion basis:

1. **Acquisition costs:** These costs include the price paid for the asset and any other costs to search for and find deposits of the natural resource.

Chapter 7 Depreciation, Impairment, and Depletion

2. **Exploration costs:** These costs include the price paid to extract the natural resource.
3. **Development costs:** These costs include drilling costs and tunnel formation. The costs for the purchase of heavy equipment to extract and ship the natural resource is not part of the depletion basis.
4. **Restoration cost:** This cost is the price paid to restore the asset after extraction of the natural resource.

7.4 Valuation of a Depleting Asset

When the purchase of a depleting asset such as an ore deposit is contemplated, its accurate appraisal requires an estimate of the total quantity of mineral present, the rate at which it can be extracted, and the cost involved. In many numerical problems, the rate of extracting the ore and the cost of this operation are considered to remain constant during the life of the asset, notwithstanding the fact that such is generally not the case. For convenience, we shall assume that the profit realized through exploitation of the asset is calculated annually.

A wasting asset may possess some residual value after it is fully exhausted. For example, when the timber has been completely removed from a tract of land, the land itself has commercial value. In problems pertaining to depletion, if no mention of salvage value appears, it is understood that none exists.

The final exhaustion of a depleting asset produces an abrupt cessation of the income derived from its exploitation, assuming the absence of salvage value. Since none of the invested capital is recovered at the conclusion of the venture, it follows that capital recovery occurred periodically while the asset was being exploited. Hence, the annual profit of a depleting asset consists of two distinct elements; the first is a return on the capital invested in the enterprise during that year, and the second is a partial restoration of the capital originally invested. The purchase and exploitation of a depleting asset therefore form a *Type B investment* and require the determination of an average investment rate rather than a constant rate pertaining to this asset alone.

To provide us with a rational means of calculating this average investment rate, the assumption is made that the owners defer the recovery of their invested capital until the conclusion of the business venture by depositing a fixed portion of the annual profit in a sinking fund, the principal in the fund to equal the original capital investment when the venture terminates. The fund is closed and this principal withdrawn by the owners at that date. Thus, the sum annually deposited in the fund represents capital recovery; the remainder of the annual profit represents a return on the original investment, and is taken by the owners in the form of dividends. In order to achieve a conservative calculation of the average investment rate, it is customary to use a rather low investment rate as the interest rate of the sinking fund.

We are in reality dealing with two distinct investments, namely, the purchase and exploitation of the depleting asset, and the reinvestment of recovered capital in a sinking fund. For simplicity, however, the two investments are considered to constitute a single composite investment, as

though it were legally binding upon the owners to defer the recovery of their capital by establishing this sinking fund. By adopting this simplifying assumption, we have transformed the actual Type B investment to an equivalent one of *Type A investment*, in which capital recovery is achieved by means of a single lump-sum payment at the termination of the investment period. Conforming to the terminology generally employed, we shall refer to the purchase of the depleting asset and the reinvestment of capital in the sinking fund as simply the “investment in the asset” and shall refer to the average investment rate pertaining to this composite investment as simply the “investment rate” of the asset.

If a depleting asset does possess salvage value, then this value represents partial capital recovery at the termination of the venture. Therefore, the sinking fund need only accumulate a principal equal to the difference between the invested capital and the salvage value.

We shall now derive an equation for calculating the *capital investment* to be made in a depleting asset, in correspondence to a given investment rate. Let

C = capital investment

L = salvage value

R = annual net profit (assumed to remain constant)

n = number of years of the business venture

r = rate of return on the investment

i = interest rate earned by the sinking fund

Then

$$\text{Annual return on the investment} = Cr$$

and

$$\text{Annual deposit on the sinking fund} = R - Cr$$

At the end of n years, the principal in the sinking fund must equal the invested capital less the salvage value. Hence,

$$C - L = (R - Cr) \frac{(1+i)^n - 1}{i}$$

and solving for capital investment C we get

$$C = \frac{R}{r + \frac{i}{(1+i)^n - 1}} + \frac{L}{r \frac{(1+i)^n - 1}{i} + 1} \quad (7.16)$$

Chapter 7 Depreciation, Impairment, and Depletion

Example 7.13

It is estimated that a timber tract will yield an annual profit of \$100,000 for 6 years, at the end of which time the timber will be exhausted. The land itself will then have an anticipated value of \$40,000. If a prospective purchaser desires a return of 8% on his investment and can deposit money in a sinking fund at 4%, what is the maximum price he should pay for the tract?

Solution:

For this example, salvage value $L = 40000$, annual net profit $R = 100000$, number of years $n = 6$, rate of return $r = 0.08$, and interest rate $i = 0.04$. By substitution of these values into (7.16), Page 7–21, we get

$$\begin{aligned} C &= \frac{100000}{0.08 + \frac{0.04}{(1 + 0.04)^6 - 1}} + \frac{40000}{0.08 \frac{(1 + 0.04)^6 - 1}{0.04} + 1} \\ &= \frac{100000}{0.08 + 0.15076} + \frac{40000}{0.53064 + 1} = \frac{100000}{0.23076} + \frac{40000}{1.53064} = \$459,480 \end{aligned}$$

Check:

Capital Invested	=	\$459,480
Salvage value	=	40,000
Principal required in sinking fund at end of 6th year	=	419,480
Annual deposit required $419,480 \times \frac{0.04}{(1 + 0.04)^6 - 1}$	=	63,241
Annual return on investment $100,000 - 63,241$	=	36,759
Rate of return $36,759/459,480$	=	8%

Although for mathematical reasons we have blended the purchase of the asset and the deposit of capital in a sinking fund into a single investment, we must not lose sight of the fact that in reality they are two distinct investments and that 8% represents the average rate of return of the two over the 6-year period. (This is sometimes referred to as the speculative rate, to distinguish it from the sinking-fund rate.) The rate of return resulting solely from purchase of the tract is greatly in excess of 8%.

This method of allowing for depletion by applying the device of a hypothetical sinking fund for the investment of recovered capital pending termination of the business venture is closely parallel, in its underlying theory, to the sinking-fund method of allocating depreciation. The deprecia-

tion charge that is periodically entered in the accounting records of a business concern serves a dual purpose. First, it records the loss represented by the disintegration of the capital invested in the assets of the enterprise, as these assets deteriorate with use. Second, by deducting this depreciation from the gross profit for that particular period, the firm is in effect reserving a portion of the gross profit for the eventual replacement of its assets, since the enterprise is intended to continue indefinitely. The sinking-fund method thus postulates that the replacement capital so accumulated should not be employed by the firm in the form of working capital, where it is surrounded by the hazards that inhere in all business activity, but rather replacement funds should be maintained intact through some conservative form of investment, such as the creation of a sinking fund. (Implicit in this reasoning is an assumed equality between the original cost of the assets and their replacement cost.)

By analogy, it is possible to regard the purchase and exploitation of a timber tract as constituting not a single isolated investment but rather one integral unit of a continuing program of investment in depleting assets. Thus, as one tract is exhausted, it is replaced in this investment program by the purchase of a new tract. If we apply this conception, it is seen that, since the invested capital vanishes as the timber is removed, it is essential that a portion of the gross profit for each period be retained to replace the liquidated capital, thereby permitting the purchase of the succeeding tract. Hence, whether we are dealing with the depreciation of an asset resulting from its use in production or with the depletion of a wasting asset resulting from the extraction of its contents, the creation of a sinking fund can be regarded as a warranty for the preservation of replacement capital.

The preceding analysis, of course, is a theoretical one, designed for the purpose of developing a simple method of evaluating a depleting asset. In actual practice, since entrepreneurs seek constantly to secure the maximum possible return on their available capital, this hypothetical sinking fund is generally not established. Where a depleting asset is owned by a corporation, the law permits dividends to be paid for the full amount of the periodic earnings. Thus, each dividend received by a stockholder comprises both a return on his investment and a partial restoration of his capital, although the stockholder usually lacks knowledge of the precise composition of his dividend.

Many problems pertaining to depletion require the calculation of the “purchase price” of an asset that corresponds to a given speculative rate. It is to be understood, however, that this term is being used in a very broad sense, as though it were synonymous with the actual investment. The total investment in a depleting asset represents, basically, the total capitalization of the firm; it encompasses the sums expended both in the purchase of the asset itself and in the purchase of the operating equipment. The periodic profit referred to in these problems denotes the actual profit realized through the sale of the commodity, prior to entering an allowance for depletion. The residual profit after this allowance has been made represents the interest earned by the invested capital. The terminology employed is therefore somewhat contradictory, for in ordinary commercial parlance the term “profit,” as applied to a business not dealing in wasting assets, denotes the profit remaining after the allowance for depreciation of the firm’s assets has been deducted.

Chapter 7 Depreciation, Impairment, and Depletion

Example 7.14

A mine is purchased for \$10,000,000, and it is anticipated that it will be exhausted at the end of 20 years. If the sinking-fund rate is 4%, what must be the annual return from the mine to realize a return of 7% on the investment?

Solution:

For this example, capital investment $C = 10,000,000$, salvage value $L = 0$, number of years $n = 20$, rate of return $r = 0.07$, and interest rate $i = 0.04$. By substitution of these values into (7.16), Page 7-21, we obtain

$$10,000,000 = \frac{R}{0.07 + \frac{0.04}{(1 + 0.04)^{20} - 1}} + 0 = \frac{R}{0.07 + 0.03358} = \frac{R}{0.10358}$$

Solving for R we get

$$R = 10,000,000 \times 0.10358 = \$1,030,580$$

7.5 Summary

- A *capital asset* is a piece of equipment, a building, or a vehicle for use in our business.
- *Depreciation* is a method that allows us to deduct the cost of business property. It is considered an expense and is listed in an income statement under expenses.
- Real property and personal property are two types of property that can be depreciated.
- Land can never be depreciated.
- To determine the annual depreciation cost for our assets, we need to know the initial cost of the assets and how many years the assets will retain some value for our business.
- Depreciable property must meet the following requirements:
 1. It must be used in business or held to produce income
 2. It must be expected to last more than one year
 3. It must be something that wears out, decays, gets used up, becomes obsolete, or loses its value from natural causes.
- We cannot depreciate the following items:
 1. Property placed in service and disposed of in the same year
 2. Inventory
 3. Land
 4. Repairs and replacements that do not increase the value of our property, make it more useful, or lengthen its useful life.
 5. Items which do not decrease in value over time, such as antiques over 100 years old, expensive solid wood furniture, and fine china.
- The annual depreciation expense depends on the depreciation method that one chooses. These are:
 1. Straight-line depreciation. It is computed using the formula

$$\text{Annual depreciation expense} = \frac{\text{Cost} - \text{Salvage value}}{\text{Estimated useful life}}$$

2. Sum of the Year's Digits. It is a method of accelerated depreciation and is computed using the formula

$$\text{SYD} = \frac{(\text{Cost} - \text{Salvage value}) \times (\text{Useful life} - \text{Period} + 1)}{(\text{Useful life}) \times (\text{Useful life} + 1) / 2}$$

Chapter 7 Depreciation, Impairment, and Depletion

3. Fixed Declining Balance. This method computes depreciation at a fixed rate and uses the formula

$$\text{FDB} = (\text{Cost} - \text{Total depreciation from prior periods}) \times \left(1 - \left(\frac{\text{Salvage value}}{\text{Cost}}\right)^{1/(\text{Life})}\right)$$

4. General Declining Balance (125%, 150%, 200%). These methods use the following formula where *Factor* is the rate at which the balance declines.

$$\text{GDB} = (\text{Cost} - \text{Total depreciation from prior periods}) \times \left(\frac{\text{Factor}}{\text{Life of asset}}\right)$$

5. Variable–Declining Balance. This method computes the depreciation of an asset using the DDB method, and switches to SL method when the SL depreciation is greater than the DDB method or any other variable–rate declining balance method.

6. Units of Production. This method allocates depreciation expenses according to actual physical usage and it is computed using the following formula:

$$\text{Annual depreciation} = \frac{\text{Initial cost} - \text{Salvage value}}{\text{Total estimated units of output}}$$

- The Accelerated Cost Recovery System (ACRS), the Modified Accelerated Cost Recovery System (MACRS), and Section 179 are depreciation methods set forth by the Internal Revenue Service (IRS).
- *Book value* is the net dollar value at which an asset is carried on a company's balance sheet.
- *Fair market value* is the present value of an asset, that is, the value of a similar asset in its present condition.
- An *impairment loss* occurs when an asset suffers a permanent impairment in value. It is computed using the formula

$$\text{Impairment loss} = \text{Book value} - \text{Fair market value}$$

- Depleting assets are resources such as oil wells, timber tracts, and mines in which the extracted commodity is not replaced. The capital investment to be made in a depleting asset is calculated using the formula

$$C = \frac{R}{r + \frac{i}{(1+i)^n - 1}} + \frac{L}{r \frac{(1+i)^n - 1}{i} + 1}$$

where C = capital investment, L = salvage value, R = annual net profit (assumed to remain constant), n = number of years of the business venture, r = rate of return on the investment, and i = interest rate earned by the sinking fund.

7.6 Exercises

1. Compute the book value at the end of 6 years of a \$100,000 machine having a useful life of 10 years and a salvage value of 10% of its first cost at that time, using straight line depreciation.
2. Compute the book value at the end of 4 years of a \$20,000 machine having a useful life of 8 years and a salvage value of \$2,000 at the end of that time, using the Sum of the Years Digits (SYD) method
3. Compute the depreciation for each year using the fixed declining balance (FDB) method of an asset costing \$1,000,000, having life expectancy of 5 years, and a salvage value of \$400,000.
4. Repeat Exercise 3 above using the
 - a. 125% General Declining Balance Method.
 - b. 150% General Declining Balance Method.
 - c. 200% General Declining Balance Method.

Verify your answers with the Microsoft Excel **DDB** function or the MATLAB **deppendb** function and tabulate the results for comparison. Include the results of Exercise 3 in your table.

5. It is estimated that a mine, now operating, can be expected to make a net profit after all taxes are paid of \$150,000,000 per year for 35 years, at which time it will be exhausted and have no salvage value. What should a group of investors pay for the mine now, so that the group will have an annual income of 12% on its investment after it has made an annual deposit into a fund which, at 3% interest, will accumulate to the amount of his investment (return of investment) in 35 years, when the mine will be exhausted?
6. A syndicate wishes to purchase an oil well which, estimates indicate, will produce a net income of \$200,000,000 per year for 30 years, with a salvage value of \$1,000,000. What should the syndicate pay for the well if, out of this net income, a return of 10% on the investment is desired and a sinking fund is to be established at 3% interest to recover this investment?
7. An office building, exclusive of its elevators and mechanical equipment, costs \$100,000,000 has an estimated useful life of 40 years, and has no salvage value. Its elevators and mechanical equipment cost \$32,000,000, have an estimated useful life of 20 years and a salvage value of \$2,000,000.
 - a. Using the straight line method of depreciation determine the value of the building and its equipment at the end of 10 years' service.
 - b. What is the annual cost of amortizing the initial purchase of equipment and of providing a sinking fund to replace the equipment at the end of 20 years if the annual interest rate is 4%?

Chapter 7 Depreciation, Impairment, and Depletion

7.7 Solutions to End-of-Chapter Exercises

1. Salvage value is $\$100,000 \times 0.10 = \$10,000$ and the depreciation at the end of the 6th year is

$$\text{Depreciation}|_{6\text{th year}} = \frac{100000 - 10000}{10} \times 6 = 9000 \times 6 = \$54,000$$

Therefore, the book value at the end of the 6th year is

$$\text{Bookvalue}|_{6\text{th year}} = 100000 - 54000 = \$46,000$$

2. To compute the total depreciation to the end of the n th year, we use equation (7.4), Page 7-5.

$$\text{SYD}|_{\text{Total to } n\text{th year}} = \frac{\text{Period} \times (\text{Cost} - \text{Salvage value}) \times (2 \times \text{Useful life} - \text{Period} + 1)}{(\text{Useful life}) \times (\text{Useful life} + 1)}$$

or

$$\text{SYD}|_{\text{Total to 4th year}} = \frac{4 \times (20000 - 2000) \times (2 \times 8 - 4 + 1)}{8 \times (8 + 1)} = \frac{4 \times 18000 \times 13}{72} = \$13,000$$

Therefore, the book value at the end of the 4th year is

$$\text{Bookvalue}|_{4\text{th year}} = 20000 - 13000 = \$7000$$

- 3.

$$\text{FDB}|_{1\text{st year}} = \text{Cost} \times \left(1 - \left(\frac{\text{Salvage value}}{\text{Cost}} \right)^{\frac{1}{\text{Life}}} \right) \times \frac{\text{Month}}{12}$$

$$\begin{aligned} \text{FDB}|_{1\text{st year}} &= 1,000,000 \times \left(1 - \left(\frac{400,000}{1,000,000} \right)^{1/5} \right) \times \frac{12}{12} \\ &= 1,000,000 \times (1 - 0.83255) = \$167,450 \end{aligned}$$

Since the depreciation is taken on a yearly basis, for the 2nd through the last year, FDB uses the following formula:

$$\text{FDB} = (\text{Cost} - \text{Total depreciation from prior periods}) \times \left(1 - \left(\frac{\text{Salvage value}}{\text{Cost}} \right)^{1/\text{Life}} \right)$$

For the 2nd year

$$\begin{aligned} \text{FDB}|_{2\text{st year}} &= (1,000,000 - 167,450) \times \left(1 - \left(\frac{400,000}{1,000,000} \right)^{1/5} \right) \times \frac{12}{12} \\ &= 832,550 \times (1 - 0.83255) = \$139,410 \end{aligned}$$

For the 3rd year

$$\begin{aligned} \text{FDB}|_{3\text{rd year}} &= (1,000,000 - 167,450 - 139,410) \times (1 - 0.83255) \\ &= 693,140 \times (1 - 0.83255) = \$116,060 \end{aligned}$$

For the 4th year

$$\begin{aligned} \text{FDB}|_{4\text{th year}} &= (1,000,000 - 167,450 - 139,410 - 116,060) \times (1 - 0.83255) \\ &= 577,080 \times (1 - 0.83255) = \$96,630 \end{aligned}$$

and for the 5th year

$$\begin{aligned} \text{FDB}|_{5\text{th year}} &= (1,000,000 - 167,450 - 139,410 - 116,060 - 96,630) \times (1 - 0.83255) \\ &= 480,450 \times (1 - 0.83255) = \$80,450 \end{aligned}$$

4.

a.

$$125\% \text{ GDB} = (\text{Cost} - \text{Total depreciation from prior periods}) \times \left(\frac{1.25}{\text{Life of asset}} \right)$$

$$125\% \text{ GDB}|_{1\text{st year}} = (1,000,000 - 0) \times \left(\frac{1.25}{5} \right) = \$250,000$$

$$125\% \text{ GDB}|_{2\text{nd year}} = (1,000,000 - 250,000) \times \left(\frac{1.25}{5} \right) = \$187,500$$

$$125\% \text{ GDB}|_{3\text{rd year}} = (1,000,000 - 250,000 - 187,500) \times \left(\frac{1.25}{5} \right) = \$140,630$$

We now pose and observe that the total depreciation up to the 3rd year is

$$250,000 + 187,500 + 140,630 = \$578,130$$

and since the salvage value is \$400,000, we can only depreciate

$$1,000,000 - 400,000 = \$600,000$$

Therefore, the depreciation for the 4th year is $600,000 - 578,130 = \$21,870$ and thus there is no depreciation for the 5th year.

Check with Microsoft Excel:

$$=\text{DDB}(1000000,400000,5,1,1.25) \text{ returns } \$250,000$$

$$=\text{DDB}(1000000,400000,5,2,1.25) \text{ returns } \$187,500$$

$$=\text{DDB}(1000000,400000,5,3,1.25) \text{ returns } \$140,630$$

Chapter 7 Depreciation, Impairment, and Depletion

=DDB(1000000,400000,5,4,1.25) returns \$21, 870

=DDB(1000000,400000,5,5,1.25) returns \$0

Check with MATLAB:

GDB125=depgendb(1000000, 400000, 5, 1.25)

returns

GDB125 =

250000 187500 140625 21875 0

b.

$$150\% \text{ GDB} = (\text{Cost} - \text{Total depreciation from prior periods}) \times \left(\frac{1.50}{\text{Life of asset}} \right)$$

$$150\% \text{ GDB}|_{1\text{st year}} = (1,000,000 - 0) \times \left(\frac{1.50}{5} \right) = \$300,000$$

$$150\% \text{ GDB}|_{2\text{nd year}} = (1,000,000 - 300,000) \times \left(\frac{1.50}{5} \right) = \$210,000$$

$$150\% \text{ GDB}|_{3\text{rd year}} = 1,000,000 - 400,000 - (300,000 + 210,000) = \$90,000$$

Check with Microsoft Excel:

=DDB(1000000,400000,5,1,1.50) returns \$300, 000

=DDB(1000000,400000,5,2,1.50) returns \$210, 000

=DDB(1000000,400000,5,3,1.50) returns \$90, 000

=DDB(1000000,400000,5,4,1.50) returns \$0

Check with MATLAB:

GDB150=depgendb(1000000, 400000, 5, 1.50)

returns

GDB150 =

300000 210000 90000 0 0

c.

$$200\% \text{ GDB} = (\text{Cost} - \text{Total depreciation from prior periods}) \times \left(\frac{2.00}{\text{Life of asset}} \right)$$

$$200\% \text{ GDB}|_{1\text{st year}} = (1,000,000 - 0) \times \left(\frac{2.00}{5}\right) = \$400,000$$

$$200\% \text{ GDB}|_{2\text{nd year}} = 1,000,000 - 400,000 - 400,000 = \$200,000$$

$$200\% \text{ GDB}|_{3\text{rd year}} = 1,000,000 - 400,000 - (400,000 + 200,000) = \$0$$

Check with Microsoft Excel:

$$=DDB(1000000,400000,5,1,2.00) \text{ returns } \$400,000$$

$$=DDB(1000000,400000,5,2,2.00) \text{ returns } \$200,000$$

$$=DDB(1000000,400000,5,3,2.00) \text{ returns } \$0$$

Check with MATLAB:

$$\text{GDB200} = \text{depgendb}(1000000, 400000, 5, 2.00)$$

returns

$$\text{GDB200} =$$

400000 200000 0 0 0

Period	FDB	125% GDB	150% GDB	200% GDB
1	\$167,450	\$250,000	\$300,000	\$400,000
2	139,410	187,500	210,000	200,000
3	116,060	140,630	90,000	
4	96,630	21,870		
5	80,450			
Total	\$600,000	\$600,000	\$600,000	\$600,000

5.

The capital investment C is computed with relation (7.16), Page 7–21, where salvage value $L = 0$, annual net profit $R = 150000$, number of years $n = 35$, rate of return $r = 0.12$, and interest rate $i = 0.03$. Then,

$$C = \frac{R}{r + \frac{i}{(1+i)^n - 1}} + \frac{L}{r \frac{(1+i)^n - 1}{i} + 1} = \frac{150,000,000}{0.12 + \frac{0.03}{(1+0.03)^{35} - 1}} + 0 = \$1,098,585,000$$

Chapter 7 Depreciation, Impairment, and Depletion

6.

Relation (7.16), Page 7–21, applies also here. With annual net profit $R = \$200,000,000$, salvage value $L = \$1,000,000$, number of years $n = 30$, rate of return $r = 0.10$, and interest rate $i = 0.03$, the required capital investment C is

$$C = \frac{R}{r + \frac{i}{(1+i)^n - 1}} + \frac{L}{r \frac{(1+i)^n - 1}{i} + 1} = \frac{200,000,000}{0.10 + \frac{0.03}{(1+0.03)^{30} - 1}} + \frac{1,000,000}{0.10 \times \frac{(1+0.03)^{30} - 1}{0.03} + 1}$$

$$= \frac{200,000,000}{0.12102} + \frac{1,000,000}{5.7575} = 1,652,619,000 + 173,690 = \$1,652,792,690$$

7.

a.

At the end of the 10-year period, the building and equipment depreciation is shown on the table below.

	Building	Equipment
Annual straight line depreciation: $SL_B = (100,000,000 - 0)/40$ $SL_E = (32,000,000 - 2,000,000)/20$	\$2,500,000	\$1,500,000
Total depreciation at the end of 10 years: $2,500,000 \times 10$ $1,500,000 \times 10$	25,000,000	15,000,000
Book value at the end of 10 years: $100,000,000 - 25,000,000$ $32,000,000 - 15,000,000$	75,000,000	17,000,000
Total book value of building and equipment at the end of 10 years: $75,000,000 + 17,000,000 = \mathbf{\$92,000,000}$		

b.

For this example, Initial investment = 32,000,000, $i = 0.04$, and $n = 20$. Then,

$$\text{Annual cost} = \text{Initial investment} \times \frac{\text{Interest}}{1 - (1 + \text{Interest})^{-n}}$$

$$= 32,000,000 \times \frac{0.04}{1 - (1 + 0.04)^{-20}} = 32,000,000 \times 0.07358 = \$2,354,616$$

Check with Microsoft Excel:

The function

$$=\text{PMT}(0.04,20,-320000) \text{ returns } \$2,354,616$$

Next, we must find the sinking fund to replace equipment at the end of 20 years. Since the equipment have a salvage value of \$2,000,000, the uniform series of deposits is based on an investment of $32,000,000 - 2,000,000 = \$30,000,000$. Then,

$$R = S_n \frac{i}{(1+i)^n - 1} = 30,000,000 \times \frac{0.04}{(1+0.04)^n - 1} = 30,000,000 \times 0.03358 = \$1,007,453$$

Finally, the annual cost of amortizing the initial purchase of equipment and of providing a sinking fund to replace equipment at the end of 20 years is:

$$2,354,616 + 1,007,453 = \$3,362,069$$

Chapter 8

Introduction to Probability and Statistics

This chapter discusses the basics of probability and statistics. It is intended for readers who have a need for an introduction to this topic, or an accelerated review of it. Readers with a strong background in probability and statistics may skip this chapter. Others will find it useful, as well as a convenient source for review.

8.1 Introduction

Probability and statistics are two closely related branches of mathematics. They apply to all science and engineering fields, sociology, business, education, insurance, and many others. For this reason, we devote all subsequent chapters to this topic.

In some cases, measurements of various parameters are deterministic; this means that the results can be predicted exactly. However, in most cases there are many unpredictable variables involved, and it becomes an impossible task to predict the result exactly. Consider, for example, the case of a radar system detection, where the radar system detects the presence of a target and extracts useful information such as *range* (distance) and *range rate* (velocity) of the target. Here, we are confronted with probabilities such as:

- a. *the probability of false alarm*, that is, the probability that the radar receiver decides that a target is present when in fact it is not present.
- b. *the probability of detection*, that is, the probability that the radar receiver decides a target is present when in fact it is present.

Obviously, we need to use a mathematical procedure to analyze such situations. This mathematical procedure is the field of probability and statistics.

8.2 Probability and Random Experiments

Probability is a measure of the likelihood of the occurrence of some event. The probability of an event A occurring is denoted as $P(A)$; this is a real number between zero and one, that is,

$$0 \leq P(A) \leq 1$$

The sum of the probabilities of all possible events A_1, A_2, \dots, A_n is equal to one, that is,

$$P(A_1) + P(A_2) + P(A_3) + \dots + P(A_n) = \sum_{i=1}^n P(A_i) = 1 \quad (8.1)$$

where the symbol Σ denotes summation.

Relation (8.1) states that the total probability is equal to 1 ; it is similar to the geometric axiom that the whole is equal to the sum of its parts.

We define a *random experiment* as one that is repeated many times under similar circumstances, but it yields unpredictable results in each trial. For example, when we roll a die, we conduct a random experiment where the outcome (result or event)^{*} is one of the numbers 1 through 6. As another example, a security (stock in a certain public company) may yield a predictable dividend value, but the price of the security varies in an unpredictable manner over a period of time.

The term *population* refers to a set of all possible outcomes of a random experiment. A specific number or value which is a result of a random experiment, is called a *random variable*. A *discrete* random variable can assume only certain values; for instance, the tossing of a die or the flipping of a coin, can only result in certain values. A *continuous* random variable, on the other hand, can assume any value within a given interval. For instance, temperature or pressure can assume an infinite number of values within a certain interval. We will discuss random variables, in detail, on the next chapter.

As stated earlier, the outcomes of a random experiment fluctuate because the causes which determine the outcomes cannot be controlled. However, if the conditions of the experiment are somewhat uniform – for example, if the same die is used in successive throws, or if the same resistor is used in an experiment and held at a constant temperature – we will observe some *statistical regularities* when we consider all results of a very large number of the random experiment. These statistical regularities can be expressed by laws, which yield the probability of the occurrence of an event. An example of statistical regularity will be given in the next section.

8.3 Relative Frequency

Let A denote one of the possible outcomes of a random experiment as a result of N independent trials. Now, let $N(A)$ represent the number of times that result A occurs. Then, the ratio

$$f_N(A) = \frac{N(A)}{N} \quad (8.2)$$

is called the *relative frequency* of the result A. Obviously, the relative frequency must be a number between zero and one, that is,

$$0 \leq f_N(A) \leq 1 \quad (8.3)$$

The plot of Figure 8.1 below shows $f_N(A)$ versus N for a typical sequence of trials in a coin-tossing experiment, where A denotes either “heads” or “tails”. We observe that the relative frequency

^{*} Sometimes outcomes and results (events) are used interchangeably; in a strict sense the set of outcomes in a random experiment is infinite, and if we choose those sets which are of interest to us, we call them events or results.

$f_N(A)$ varies wildly for small values of N independent trials, but as $N \rightarrow \infty$, it stabilizes in the vicinity of the $1/2$ value. This is a simple example of statistical regularity.

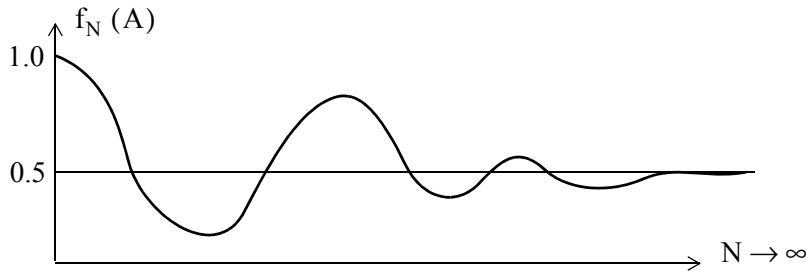


Figure 8.1. Relative Frequency in N Independent Trials

Different events from a random experiment are denoted as $A_1, A_2, A_3, \dots, A_N$. Events that cannot occur simultaneously in a given trial, are called *mutually exclusive*. For instance, if A_1 represents “heads” and A_2 “tails”, then A_1 and A_2 are mutually exclusive events since they cannot occur simultaneously if only a single coin is tossed just once.

The occurrence of the event “either A_i or A_j ” satisfies the equality

$$N(A_i \text{ or } A_j) = N(A_i) + N(A_j) \tag{8.4}$$

or

$$f_N(A_i \text{ or } A_j) = f_N(A_i) + f_N(A_j) \tag{8.5}$$

As another example, consider the throwing of a die where A_i denotes the result that the i th face shows. Then, the result “even face shows up” represents the result A_2 or A_4 or A_6 and therefore,

$$f_N(A_2 \text{ or } A_4 \text{ or } A_6) = f_N(A_2) + f_N(A_4) + f_N(A_6) \tag{8.6}$$

Thus, for a fair die, we expect the relative frequency of each A_i to settle about the $1/6$ value, and the relative frequency of “even face shows up” to stabilize at $1/2$.

To illustrate the basic concepts of probability, we find it convenient to use probabilities that relate to games of chance. The following example illustrates the validity of Equation (8.5).

Example 8.1

A box contains two blue and eight red balls. Compute the probability of drawing a blue ball (without looking inside the box, of course).

Solution:

Here, we have two possible events: B is a blue ball and R is a red ball. Since there are two blue balls in a box containing 10 balls total, we would expect the blue balls to be drawn 20% of the time and the red balls 80% of the time. Thus, the probabilities are $P(B) = 0.2$ and $P(R) = 0.8$.

Now, suppose that these 10 balls, besides being colored, are also numbered 1 through 10. If we keep drawing a ball from the box a very large number of times while we are placing the ball back in the box before drawing the next one, we would expect that any of these balls would appear $1/10$ or 10% of the time since each ball is equally likely to be drawn. Then, if we add together the chances of drawing any of the two blue balls we get the probability of

$$\frac{1}{10} + \frac{1}{10} = \frac{2}{10} = 0.2 = 20\%$$

This demonstrates Equation (8.5) which states that *we can add the probabilities of each group of mutually exclusively events to determine the probability of occurrence of the overall group.*

Example 8.2

Two dice are thrown. What is the probability of getting a six?

Solution:

Each die has six faces; then the possible outcomes are $6 \times 6 = 36$, that is, the probability of drawing any combination of two faces is $1/36$. The five different combinations that add up to 6 are $5 + 1, 4 + 2, 3 + 3, 2 + 4, 1 + 5$. Therefore, the probability of getting a six is

$$\frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{5}{36}$$

We can also use the following approach:

There are 36 possible outcomes but only 5 are the desired ones which will yield a 6. Therefore, if we choose the drawing of a 6 as the desired event, this will occur $5/36$ of the time in many repeated tosses.

8.4 Combinations and Permutations

Sometimes we use *combinations* and *permutations* to compute probabilities. Consider the experiment where we select two numbers from the four number set [1, 2, 3, and 4]. The possible pairs,

without regard to the order, are (1, 2), (1, 3), (1, 4), (2, 3), (2, 4), and (3, 4). Thus, there are six possible pairs. In this case, we say that when selecting two items from a set of four distinct items, there are six possible combinations.

In general, suppose we have a collection of n objects. The number of combinations of these n objects taken r at a time, is denoted as $C(n, r)$; It is computed from the expression

$$C(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!} \quad (8.7)$$

Example 8.3

A deck of 52 cards consists of the numbers 2 through 10, jack, queen, king and ace where each of these can be any of the four suits, i.e., heart, spade, diamond or club. Compute the number of combinations in a five card poker game that can be dealt from this deck of cards.

Solution:

We are seeking the number of combinations. Therefore, from (8.7),

$$C(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!} = C(52, 5) = \binom{52}{5} = \frac{52!}{5!(52-5)!} = 2598960$$

Next, let us consider the case where we pick r items from n possible events, but we are interested in the *order* of selection of the r items. Now, the number of *permutations* of n possible events taken r at a time is denoted as $P(n, r)$. It is computed from the expression

$$P(n, r) = \frac{n!}{(n-r)!} \quad (8.8)$$

Example 8.4

Compute the number of permutations if we select 2 numbers from the set of four numbers [1, 2, 3, and 4].

* A number m followed by an exclamation mark (!), i.e., $m!$, is the product of all positive integers from m down to 1. In other words, $m! = m \cdot (m-1) \cdot (m-2) \cdot (m-3) \cdot \dots \cdot 3 \cdot 2 \cdot 1$. Thus, $10! = 10 \times 9 \times \dots \times 2 \times 1 = 3,628,000$. We call $m!$ m factorial. It is also discussed in Appendix C.

Solution:

The permutations are (1, 2), (2, 1), (1, 3), (3, 1), (1, 4), (4, 1), (2, 3), (3, 2), (2, 4), (4, 2), (3, 4), and (4, 3). Therefore, for this example, there are 12 permutations. The number of permutations can also be obtained from (8.8). For this example, $n = 4$ and $r = 2$. Thus,

$$P(n, r) = \frac{n!}{(n-r)!} = P(4, 2) = \frac{4!}{(4-2)!} = 12$$

Example 8.5

In a lotto contest, six numbers are selected from the set of numbers [1, 2, 3, ..., 55, 56]. Compute the probabilities of selecting the correct six numbers to win the lotto when:

- the order of selection of the six numbers is immaterial
- the order of selection is important

Solution:

- a. Since the order of selection is immaterial, we need to compute the number of possible combinations. For this example, $n = 56$ and $r = 6$. Then, using (8.7), Page 8–5, we obtain

$$C(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!} = C(56, 6) = \binom{56}{6} = \frac{56!}{6!(56-6)!} = 3.2468436 \times 10^7$$

Check with Excel: =FACT(56)/(FACT(6)*FACT(56-6)) returns the same number.

Therefore, the probability of winning the lotto when combinations are considered, is

$$(\text{Probability}_{\text{combinations}}) = \frac{1}{3.2468436 \times 10^7} = 3.0799143 \times 10^{-8}$$

- b. When the order of selection is important, we need to compute the number of possible permutations. Then, using (8.8), Page 8–5, we obtain

$$P(n, r) = \frac{n!}{(n-r)!} = P(56, 6) = \frac{56!}{(56-6)!} = 2.3377274 \times 10^{10}$$

Thus, in this case, the probability of winning the lotto when permutations are considered, is

$$(\text{Probability}_{\text{permutations}}) = \frac{1}{2.3377274 \times 10^{10}} = 4.2776587 \times 10^{-11}$$

Example 8.6

In a 52-card game, a *royal flush* is a selection of five cards consisting of an ace, a king, a queen, a jack and a ten, all of the same suit. Compute the probability that a player will draw a royal flush when the order of the five cards drawn is unimportant.

Solution:

In Example 8.3, Page 8–5, we found that the number of combinations of a five card poker game that can be dealt from this deck of cards is

$$C(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!} = C(52, 5) = \binom{52}{5} = \frac{52!}{5!(52-5)!} = 2598960$$

and since there are four possible royal flushes (one for each suit), the probability of a royal flush is

$$(\text{Probability}_{\text{royal flush}}) = 4 \times \frac{1}{2598960} = 1.5390772 \times 10^{-6}$$

8.5 Joint and Conditional Probabilities

In the preceding discussion, we have considered the probability of one particular event occurring, i.e., the tossing of a coin or the throwing of a die. Quite often, however, we need to consider the *joint probability* representing the *joint occurrence* of two or more events. Consider, for example, that in the drawing of two cards from a 52 deck the first card is diamond and the second is also a diamond. This is an example of a joint probability and we wish to find the probability that the second card will be diamond when the first was also a diamond.

In our subsequent discussion, we will consider the joint occurrence of only two events A and B and we will denote the joint probability as P(AB).

In the above example, the second event is *conditioned* (dependent) upon the first. We measure this dependence by the *conditional probability* of B occurring given that A has occurred; this conditional probability is denoted as P(B|A). It is computed from

$$P(B|A) = \frac{P(AB)}{P(A)} \tag{8.9}$$

Relation (8.9) is derived on Page 8–9, relation (8.13).

Example 8.7

Compute the probability that, in the drawing of two cards from a 52 deck, the first card is red (heart or diamond) and the second is black (club or spade), provided that the first card drawn is

not returned back to the deck of the cards before the second is drawn.

Solution:

We let R represent the drawing of the first card (red) and B the drawing of the second (black). The desired joint probability of drawing first R and then B is denoted as $P(RB)$. Since there are 26 red and 26 black cards, the probability of drawing first a red card is $P(R) = 26/52 = 1/2$. After the first card is drawn, there are 51 cards left, 25 red and 26 black. Therefore, the probability of drawing a black card is $26/51$. Thus, the second drawing becomes *conditional* (dependent) upon the first drawing, and the conditional probability is $P(B|R) = 26/51$. The joint probability $P(RB)$ of drawing a red card first and a black second is computed from (8.9), which is rearranged as

$$P(RB) = P(R)P(B|R) = \frac{1}{2} \times \frac{26}{51} = \frac{26}{102} = \frac{13}{51}$$

The concepts of joint occurrences and conditional probability can be generalized as follows:

Suppose we perform a random experiment, and we observe the occurrence of events A and B in that order. We repeat this experiment many times while we measure the relative frequency of occurrence *separately* as A and B and also as a *pair* AB. Here, we wish to find if event B is, or is not dependent on event A, and if it is, to determine the relative dependence, that is, to compute $P(B|A)$ from the relative frequency measurements.

We let n_{AB} denote the number of times that the combination AB occurs in n repetitions of the experiment. Then, using the definition of the relative frequency in (8.2), and assuming that n represents a very large number, the joint probability $P(AB)$ of A occurring first and then B is

$$P(AB) = \frac{n_{AB}}{n} \quad (8.10)$$

However, in the same n trials, the event A occurred n_A times and the event B occurred n_B times. Since some of the times in which event A occurred it was followed by event B, n_A includes n_{AB} . Then, the ratio

$$\frac{n_{AB}}{n_A} \leq 1 \quad (8.11)$$

represents the relative frequency of occurrence of event B occurring after event A has occurred. This is the conditional probability $P(B|A)$, that is,

$$P(B|A) = \frac{n_{AB}}{n_A} \leq 1 \quad (8.12)$$

In the previous example, n represented a very large number of drawings of two successive cards, n_A the number of times that a red card shows up, n_B the number of times a black card shows up, and n_{AB} the number of times a red is followed by a black card.

If, in (8.12), we divide both n_{AB} and n_A by n we get

$$P(B|A) = \frac{n_{AB}/n}{n_A/n} = \frac{P(AB)}{P(A)} \quad (8.13)$$

and multiplying (8.13) by $P(A)$ we get

$$P(AB) = P(A)P(B|A) \quad (8.14)$$

We observe that (8.14) is the same as relation (8.9), Page 8–7, except that it is re-arranged.

Now, suppose that the probability of event B occurring is independent of event A . Then (8.14) reduces to $P(B|A) = P(B)$. This situation would occur if, in the two-card example, the first card drawn is returned back to the deck of the cards before the second is drawn. Then, substitution of $P(B|A) = P(B)$ into (8.14) yields

$$P(AB) = P(A)P(B) \quad (8.15)$$

Relation (8.15) states that if event B is independent of event A , we can compute the joint probability $P(AB)$ by multiplying together the separate probabilities $P(A)$ and $P(B)$. In this case, the order of occurrence of events A and B is immaterial and we can also express (8.15) as

$$P(BA) = P(B)P(A) \quad (8.16)$$

These simplifications have led us to the condition of *statistical independence* which states that two events are said to be *statistically independent* if their probabilities satisfy the following relations.

$$\begin{aligned} P(AB) &= P(BA) = P(A)P(B) \\ P(B|A) &= P(B) \\ P(A|B) &= P(A) \end{aligned} \quad (8.17)$$

Example 8.8

A box contains two red and three white balls. Two balls are drawn from this box. Compute the joint probabilities of drawing two red balls in succession if

- a. we do not place the first ball back into the box before drawing the second.
- b. we do place the first ball back into the box before drawing the second.

Solution:

- a. Let event R_1 be a red ball drawn on the first draw, and event R_2 another red ball on the second draw. Then, the probability of drawing a red ball on the first draw is $P(R_1) = 2/5$. Since only one red ball remains in the box which now contains only four balls, the conditional probability of drawing the second red ball on the second draw is $P(R_2|R_1) = 1/4$. Therefore, the joint probability $P(R_1R_2)$ is

$$P(R_1R_2) = P(R_1)P(R_2|R_1) = \left(\frac{2}{5}\right)\left(\frac{1}{4}\right) = \frac{1}{10}$$

We can obtain the same result using combinations (no need to use permutations since the red balls are assumed identical). Then,

$$C(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!} = C(5, 2) = \binom{5}{2} = \frac{5!}{2!(5-2)!} = 10$$

Thus, we have 10 possible combinations of five balls arranged in groups of two. Since we are only interested in the combination which contains two red balls, the probability of this occurring is $P(R_1R_2) = 1/10$. This is the same answer as before.

- b. If the first ball is placed back into the box before the second is drawn, the events R_1 and R_2 are statistically independent and thus

$$P(R_1R_2) = P(R_1)P(R_2) = \frac{2}{5} \times \frac{2}{5} = \frac{4}{25}$$

Example 8.9

Box A contains two red balls and one white, and Box B contains three red balls and two white. One of the boxes is selected at random, and one ball is drawn from it. Compute the probability that the ball drawn is red.

Solution:

Here, there are two possibilities of drawing a red ball, so we let R denote a red ball, $P(R)$ denote the probability of drawing a red ball, $P(AR)$ denote the joint probability that a red ball is drawn from Box A, and $P(BR)$ denote the joint probability that a red ball is drawn from Box B. The possibilities of drawing a red ball are:

a. Choose Box A and draw R or

b. Choose box B and draw R

Since events (a) and (b) are mutually exclusive,

$$P(R) = P(AR) + P(BR)$$

Now, the probability of choosing Box A is $1/2$ and drawing a red ball from it is $2/3$. Then,

$$P(AR) = P(A)P(R|A) = \frac{1}{2} \times \frac{2}{3} = \frac{2}{6} = \frac{1}{3}$$

Similarly, the probability of choosing Box B is $1/2$ and drawing a red ball from it is $3/5$. Then,

$$P(BR) = P(B)P(R|B) = \frac{1}{2} \times \frac{3}{5} = \frac{3}{10}$$

Therefore, the probability that ball drawn is red is

$$P(R) = \frac{1}{3} + \frac{3}{10} = \frac{19}{30}$$

8.6 Bayes' Rule

Let C be any event, and D_1, D_2, \dots, D_n be mutually exclusive events one of which *will definitely* occur. We use Bayes' Rule to compute the probabilities of the various events D_1, D_2, \dots, D_n which can cause event C to occur. Mathematically, *Bayes' rule* states that

$$P(D_k|C) = \frac{P(D_k)P(C|D_k)}{\sum_{k=1}^n P(D_k)P(C|D_k)} \quad (8.18)$$

In (8.18), $P(D_k)$ represents the probabilities of the events before the experiment is performed; these are referred to as the *a priori probabilities*. After the experiment is performed and event C has occurred, the probabilities $P(C|D_k)$ are referred to as the *posteriori probabilities*.

We can think of Bayes' rule as one which evaluates the probability of the "cause" D_k given the "effect" C .

Example 8.10

In Example 8.9 what is the probability that the red ball was drawn from Box A?

Solution:

We will apply Bayes' rule to obtain the solution. Since there were two boxes, we use $n = 2$ in (8.18). For our example, C is the drawn red ball and so we replace it with R . Also, D_k represents either Box A or B, and since we are seeking the probability that the red ball was drawn from Box A, we replace it with A. Because box A contained 3 balls, one of which was red, while Box B contained 5 balls, three of which were red, (8.18) is written as

$$P(A|R) = \frac{P(A)P(R|A)}{P(A)P(R|A) + P(B)P(R|AB)} = \frac{\left(\frac{1}{2}\right)\left(\frac{2}{2+1}\right)}{\left(\frac{1}{2}\right)\left(\frac{2}{2+1}\right) + \left(\frac{1}{2}\right)\left(\frac{3}{3+2}\right)} = \frac{10}{19}$$

Therefore, the probability that the red ball was drawn from Box A is

$$P(A|R) = 10/19$$

8.7 Summary

- Probability is a measure of the likelihood of the occurrence of some event.
- The probability of an event A occurring is denoted as $P(A)$; this is a real number between zero and one.
- The sum of the probabilities of all possible events is equal to one.
- A random experiment is an experiment that is repeated many times under similar circumstances, but it yields unpredictable results in each trial.
- The term *population* refers to a set of all possible outcomes of a random experiment.
- A random variable is a specific number or value which is a result of a random experiment.
- A discrete random variable can assume only certain values.
- A continuous random variable can assume any value within a given interval.
- The relative frequency $f_N(A)$ is defined as $f_N(A) = N(A)/N$ where $N(A)$ is the number of times that result A occurs in N independent trials.
- Statistical regularity refers to the condition where the relative frequency stabilizes in a certain value as the number of independent trials approaches infinity.
- Events that cannot occur simultaneously in a given trial are said to be mutually exclusive.
- Combinations are arrangements of objects where the order of arrangement does not matter. In a collection of n objects the number of combinations of these n objects taken r at a time, is denoted as $C(n, r)$ and it is computed from the expression

$$C(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

- Permutations are arrangements of objects where different ordering are considered different results. In a collection of n objects the number of permutations of n possible events taken r at a time is denoted as $P(n, r)$. It is computed from the expression

$$P(n, r) = \frac{n!}{(n-r)!}$$

- A joint probability represents the joint occurrence of two or more events. The joint probability of event A occurring first and event B occurring second is denoted as $P(AB)$.
- If a second event B is conditioned (dependent) on an event A occurring first, we measure this dependence by the conditional probability of B occurring given that A has occurred; this

Chapter 8 Introduction to Probability and Statistics

conditional probability is denoted as $P(B|A)$ and it is computed from

$$P(B|A) = \frac{P(AB)}{P(A)}$$

- Two events are said to be statistically independent if their probabilities satisfy the following relations.

$$P(AB) = P(BA) = P(A)P(B)$$

$$P(B|A) = P(B)$$

$$P(A|B) = P(A)$$

- Bayes' rule states that if C is any event, and D_1, D_2, \dots, D_n be mutually exclusive events one of which will definitely occur, we can compute the probabilities of the various events D_1, D_2, \dots, D_n which can cause event C to occur using the relation

$$P(D_k|C) = \frac{P(D_k)P(C|D_k)}{\sum_{k=1}^n P(D_k)P(C|D_k)}$$

8.8 Exercises

1. A box contains 9 white and 14 black balls. Compute the probability that a white ball is drawn on the first trial.
2. Two dice are thrown. What is the probability of getting
 - a. 7
 - b. 11
3. Compute the number of combinations and permutations in a 7 card poker game that can be dealt from a 52-card deck.
4. In a lotto contest, 5 numbers are selected from the set of numbers $[1, 2, 3, \dots, 45]$. Compute the probability of selecting the correct five numbers to win the lotto if the order of the selection of the five numbers is unimportant.
5. Compute the probability that in the drawing of 2 cards from a 52-card deck the first card is a club and the second is a diamond.
6. A box contains 7 blue and 9 red balls. Compute the joint probability of drawing two blue balls in succession if we do not place the first ball back into the box before drawing the second ball.
7. Box A contains 5 red and 4 blue balls and Box B contains 6 red and 9 blue balls. One of the boxes was selected at random and a red ball was drawn from it. What is the probability that the drawn red ball was picked from Box B?

8.9 Solutions to End-of-Chapter Exercises

1. There are 23 balls total and the probability of getting any ball is $1/23$. Since there are 9 white balls, $P(\text{white}) = 9/23$

2. The probability of drawing any combination is $1/6 \times 1/6 = 1/36$. Then,

a. The combinations which add up to 7 are $1 + 6$, $2 + 5$, $3 + 4$, $4 + 3$, $5 + 2$, and $6 + 1$, that is, 6 combinations. Therefore, $P(7) = 6/36 = 1/6$

b. The combinations which add up to 11 are $5 + 6$ and $6 + 5$, that is, 2 combinations. Therefore, $P(11) = 2/36 = 1/18$

3. Combinations:

$$C(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!} = C(52, 7) = \binom{52}{7} = \frac{52!}{7!(52-7)!} = 133, 784, 560$$

Check with Excel: =FACT(52)/(FACT(7)*FACT(52-7)) returns the same number.

Permutations:

$$P(n, r) = \frac{n!}{(n-r)!} = P(52, 7) = \frac{52!}{(52-7)!} = 674, 274, 180, 000$$

4. Because it is stated that the order is unimportant, we need to find the combinations. Thus,

$$C(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!} = C(45, 5) = \binom{45}{5} = \frac{45!}{5!(45-5)!} = 1, 221, 759$$

and

$$(\text{Probability}_{\text{combinations}}) = \frac{1}{1, 221, 759} = 8.18492 \times 10^{-7}$$

5. Let $P(C)$ = Probability of drawing a club, $P(D)$ = Probability of drawing a diamond. Then,

$$P(AB) = P(A)P(B|A) = 13/52 \times 13/51 = 169/2652 = 13/204$$

6. $B = 7$, $R = 9$, $B + R = 16$, $P(B_1) = 7/16$, $P(B_2|B_1) = 6/15$ and thus

$$P(B_1B_2) = P(B_1)P(B_2|B_1) = 7/16 \times 6/15 = 42/240 = 7/40$$

7.



Box A



Box B

and by Bayes' rule

$$P(\text{Red}|\text{Box B}) = \frac{\left(\frac{1}{2}\right)\left(\frac{6}{6+9}\right)}{\left(\frac{1}{2}\right)\left(\frac{6}{6+9}\right) + \left(\frac{1}{2}\right)\left(\frac{5}{5+4}\right)} = \frac{6/30}{6/30 + 5/18} = \frac{90}{215} = \frac{18}{43}$$

This chapter defines discrete and continuous random variables, the cumulative distribution function, the probability density function, and their properties. Statistical averages, moments, and characteristic functions are also discussed. Some sections presume knowledge of advanced mathematics; these may be skipped without loss of continuity.

9.1 Definition of Random Variables

We may associate an experiment and its possible outcomes with a space and its points. Then, we may associate a point called the *sample point* s_i with each possible outcome of the experiment. The total number of sample points $\{s\}$, corresponding to the aggregate of all possible outcomes of the experiment, is called the *sample space*.

An *event* may correspond to a *single sample point* or a *set of sample points*. For example, in an experiment involving the throw of a die, there are six possible outcomes $\{1, 2, \dots, 6\}$. By assigning a sample point to each of these possible outcomes we get a sample space that consists of six sample points. Thus a “five” corresponds to a single sample point. Also, if we choose an even number to represent a sample point, then the sample space consists of three sample points, that is, $\{2, 4, 6\}$.

Now, suppose that the outcome is a *variable* that can wander over the set of sample points, and whose value is determined by the experiment. Then, a function whose domain is a sample space and whose range is some set of real numbers, is called a *random variable* of the experiment. In our subsequent discussion we will abbreviate a random variable as *rv*.

If the outcome of an experiment is s , we will represent the *rv* as $X(s)$ or simply X . For example, the sample space representing the outcomes of the throw of a die is a set of six sample points $\{1, 2, \dots, 6\}$. Thus, if we identify the sample point k with the event that k dots are showing (for example number 3 has three dots) when the die is thrown, we say that $X(k)$ is a *rv* where k represents the number of dots that show up when the die is thrown. This is an example of a discrete *rv*.

In general, an *rv* X is a *discrete rv* if X can take on only a finite number of values in any finite observation interval. If the *rv* can take on any value in the entire observation interval, then X is a *continuous rv*. For example, the value of the temperature at a particular instant of time is a continuous *rv*.

9.2 Probability Function

The *probability function* (or *probability distribution*) is defined as

$$P(X = x) = f(x) \quad (9.1)$$

In general, $f(x)$ is a probability function if

$$f(x) \geq 0 \quad (9.2)$$

and

$$\sum_x f(x) = 1 \quad (9.3)$$

9.3 Cumulative Distribution Function

Let the *rv* X be defined in the axis of real numbers as $-\infty \leq X \leq \infty$, and consider any point x on this axis. Then, the probability of an event $X \leq x$ is referred to as *cumulative distribution function* or *probability distribution function* of the *rv* X . It is abbreviated as *cdf*, and it is denoted as

$$P_X(x) = P(X \leq x) \quad (9.4)$$

In other words, $P_X(x)$ is the probability that the *rv* X will be equal to or less than some value x . Since all values of x are mutually exclusive (i.e., temperature can have only one value at any instant, or a pointer may stop at only one point), then $P_X(x)$ must be the sum of all probabilities from $-\infty$ to x .

Now, we will introduce the concept of the *cumulative frequency distribution*; this will help us understand the meaning of the cumulative distribution function.

Consider the ages of 25 adult people picked at random from a population. Their ages and the frequency at which they occur is shown in Table 9.1.

TABLE 9.1 Frequency distribution of the ages of 25 adults.

Age	Frequency
23 to less than 30	4
30 to less than 40	9
40 to less than 50	8
50 to less than 60	2
60 to less than 75	2

Using the frequencies of occurrence shown on the right column of Table 9.1, we can compute the cumulative relative frequency by dividing each frequency by the total number of people, i.e., 25.

The *cumulative frequency percentages* are shown in Table 9.2.

TABLE 9.2 Cumulative frequency percentages for 25 adults.

Age (less than)	Cumulative Frequency	Cumulative Relative Frequency	Cumulative Frequency Percentage
23	0	$0/25=0.00$	0
30	4	$4/25=0.16$	16
40	$4+9=13$	$13/25=0.52$	52
50	$13+8=21$	$21/25=0.84$	84
60	$21+2=23$	$23/25=0.92$	92
75	$23+2=25$	$25/25=1.00$	100

In the second column (cumulative frequency) of Table 9.2, the successive numbers are obtained by addition. For example, on the third row, we added the values 4 and 9 yielding the value 13, since there are thirteen people whose age is less than 40.

We can use Excel's **FREQUENCY** function to compute how often values occur within a range of values such as those in Table 9.1. The syntax is

FREQUENCY(data_array,bins_array)

where

data_array = an array or a reference to a set of values for which we want to count frequencies of occurrence.

bins_array = an array or a reference to intervals into which we want to group the values in **data_array**.

For example, suppose that the test scores of an examination are as shown in Column A of the spreadsheet of Figure 9.1. These are the **data_array**. Column B shows the **bins_array**; these values indicate that we want to count the number of scores corresponding to ranges

$0 \rightarrow 9, 10 \rightarrow 19, 20 \rightarrow 29, 30 \rightarrow 39, 40 \rightarrow 49, 50 \rightarrow 59, 60 \rightarrow 69, 70 \rightarrow 79, 80 \rightarrow 89$ and $90 \rightarrow 100$.

We want the frequencies of occurrence to be displayed in Column C; therefore, we highlight Cells C1:C10, in C1 we type **=FREQUENCY(A1:A20,B1:B9)**, and we press the keys Ctrl → Shift → Enter simultaneously. Cell C1 indicates that there are zero scores below 10, C2 indicates that there is only one score below 19 and so on. Cell C10 displays 2; this indicates that there are 2 values (95 and 97) that are greater than the highest interval which we specified, that is, 89.

	A	B	C
1	69	10	0
2	45	19	1
3	81	29	1
4	86	39	1
5	32	49	2
6	73	59	2
7	29	69	1
8	56	79	3
9	18	89	7
10	79		2
11	85		
12	78		
13	85		
14	83		
15	81		
16	95		
17	88		
18	97		
19	45		
20	53		

Figure 9.1. Test scores

Properties of the cdf $P_X(x)$

- $P_X(x)$ is bounded between zero and one, that is,

$$P_X(-\infty) = 0 \quad \text{and} \quad P_X(\infty) = 1 \tag{9.5}$$

Proof:

Since $P_X(x)$ is a probability, it lies between 0 and 1. Moreover, the condition $X = -\infty$ excludes all outcomes, and thus it is zero; however, $X = \infty$ includes all outcomes and therefore it is equal to one.

- $P_X(x)$ is a *monotonic*,* non-decreasing function of x , that is,

$$\text{if } x_1 < x_2 \quad \text{then} \quad P_X(x_1) \leq P_X(x_2) \tag{9.6}$$

Proof:

If $x_1 < x_2$, the events $X \leq x_1$ and $x_1 < X \leq x_2$ are mutually exclusive. Also, the event $X \leq x_2$ implies $X < x_1$ and $x_1 < X \leq x_2$. Therefore,

$$P(X \leq x_2) = P(X \leq x_1) + P(x_1 < X \leq x_2)$$

* These are sequences in which the successive values either consistently increase or decrease but do not oscillate in relative value. Each value of a monotonic increasing sequence is greater than, or equal to the preceding value; likewise, each value of a monotonic decreasing sequence is less than, or equal to the preceding value.

or

$$P(x_1 < X \leq x_2) = P(X \leq x_2) - P(X \leq x_1) = P_X(x_2) - P_X(x_1) \quad (9.7)$$

Moreover, since $P(x_1 < X \leq x_2)$ is a non-negative real number, then,

$$P_X(x_2) \geq P_X(x_1) \text{ or } P_X(x_1) \leq P_X(x_2) \quad (9.8)$$

3. $P_X(x)$ is continuous on the right.

Proof:

Let the rv X assume some value a with probability $P(X = a)$, and consider the probability $P(X \leq x)$. If $x < a$, the event $X = a$ is excluded but if $x = a$, the event $X = a$ is included in $P(X \leq x)$. Since the events $X \leq x < a$ and $x = a$ are mutually exclusive, it follows that when $x = a$, the probability $P(X \leq x)$ must jump by the amount $P(X = a)$ as shown in Figure 9.2.

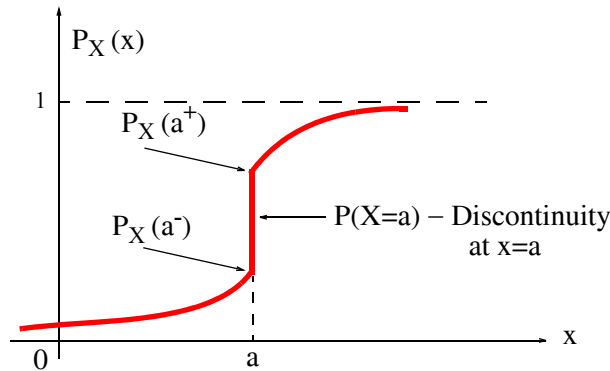


Figure 9.2. Plot of $P_X(x)$

Now, let δ and ϵ be two *infinitesimal** values that continuously approach zero as a limit but never become exactly zero. The notations $P_X(a^-)$ and $P_X(a^+)$ are used to denote that

$$P_X(a^-) = \lim_{\delta \rightarrow 0} P_X(a - \delta)$$

and

$$P_X(a^+) = \lim_{\epsilon \rightarrow 0} P_X(a + \epsilon)$$

Therefore, $P_X(x)$ is continuous on the right.

* Immeasurably or incalculably minute. In other words, very small.

In general, $P_X(x)$ can have stepwise discontinuities, as in Figure 9.2, can be everywhere continuous, or be mixed, that is, be continuous in some ranges, and have discontinuities in others.

Example 9.1

A coin is tossed twice so the sample space is $S = \{HH, HT, TH, TT\}$ where H = Heads and T = Tails. Let X be a discrete *rv* where arbitrarily we have assigned $X = 0$ for TT (2 Tails to represent the number of tails which can come up, $X = 1$ for HT and TH, and $X = 2$ for HH (2 Heads) as shown in Table 9.3.

TABLE 9.3 Table for Example 9.1

Sample Point	TT	HT	TH	HH
X	0	1	1	2

- Find the probability function corresponding to the *rv* X .
- Construct a probability graph.
- Find the cdf $P_X(x)$ of the *rv* X .
- Construct the graph of $P_X(x)$.

Solution:

- a. The probability function was defined in (9.1), Page 9–2, as $P(X = x) = f(x)$. The probabilities that we will get two tails, or one head and one tail etc. are

$$P(TT) = 1/4 \quad P(HT) = 1/4 \quad P(TH) = 1/4 \quad P(HH) = 1/4$$

Then, the probability function is obtained from Table 9.3 above as

$$P(X = 0) = P(TT) = 1/4$$

$$P(X = 1) = P(HT) + P(TH) = 1/4 + 1/4 = 1/2$$

$$P(X = 2) = P(HH) = 1/4$$

The probability function is often shown in table form as in Table 9.4.

TABLE 9.4 Probability Function of Example 9.1

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

b. A probability graph can be either in the form of a bar chart, or a histogram as shown in Figure 9.3.

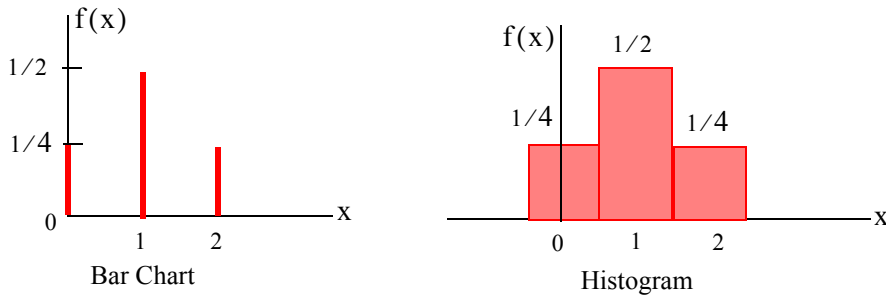


Figure 9.3. Bar Chart and Histogram for Example 9.1

We observe that in the bar chart the sum of the *ordinates* is 1, while in the histogram the sum of the *areas of the rectangles* is 1.

c. The cdf $P_X(x) = P(X \leq x)$ function is

$$P_X(x) = \begin{cases} 0 & -\infty < x < 0 \\ 1/4 & 0 \leq x < 1 \\ 1/4 + 1/2 = 3/4 & 1 \leq x < 2 \\ 3/4 + 1/4 = 1 & 2 \leq x < \infty \end{cases}$$

d. The plot of the cumulative distribution $P_X(x)$ function is shown in Figure 9.4.

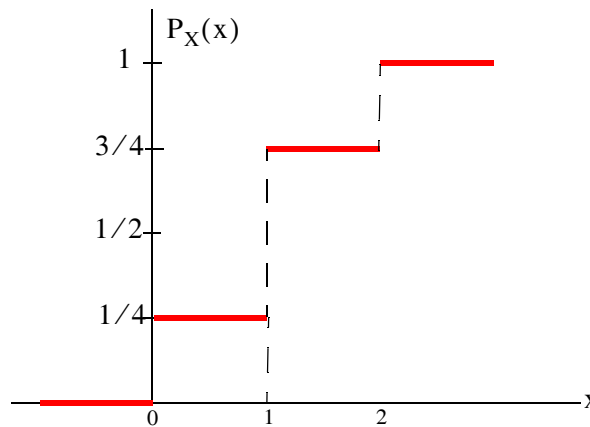


Figure 9.4. Cumulative Distribution Function for Example 9.1

We observe that the magnitudes of the jumps at 0, 1 and 2 are the ordinates 1/4, 1/2 and 1/4 of the bar chart of part (b). Therefore, it is possible to construct the bar chart representing a probability function from the cdf graph. The plot of Figure 9.4 is often referred to as a

Chapter 9 Random Variables

staircase or step function, and the value of the function at an integer is assigned from the higher step. Thus, the value at 1 is $3/4$, not $1/4$. For this reason, we say that the cdf is continuous from the right at 0, 1 and 2. We also observe that the function either remains the same, or jumps to a higher value, and so it is a monotonically increasing function.

Example 9.2

Let the *rv* X represent the transmission times x of messages in a communications system where it is known that

$$P_X(X > x) = e^{-\lambda x} \quad x > 0 \quad \lambda > 0 \quad (9.9)$$

- Find and sketch the cdf $P_X(x) = P(X \leq x)$ for (9.9).
- Find the probability $P(1/\lambda < X \leq 2/\lambda)$

Solution:

- Since $P(X \leq x) = 1 - P_X(X > x)$, the cdf for this example is

$$P_X(x) = 1 - P_X(X > x) = 1 - e^{-\lambda x} \quad (9.10)$$

and the graph of this cdf is shown in Figure 9.5.

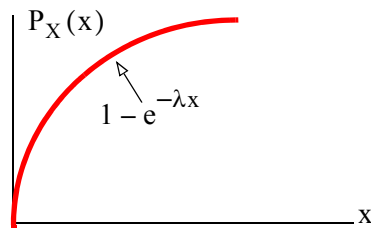


Figure 9.5. Cdf for Example 9.2

We observe that the cdf $P_X(x)$ is continuous for all $x > 0$, and it is monotonically increasing.

- From (9.7), Page 9-5,

$$P(x_1 < X \leq x_2) = P_X(x_2) - P_X(x_1)$$

Then, with (9.10) above,

$$P\left(\frac{1}{\lambda} < X \leq \frac{2}{\lambda}\right) = (1 - e^{-2}) - (1 - e^{-1}) = e^{-1} - e^{-2} = 0.233$$

9.4 Probability Density Function

While the cdf is particularly useful when computing probabilities, the *probability density function* (*pdf*) is more useful when we need to compute statistical averages. We will discuss statistical averages in Section 9.6.

We denote the pdf as $f_X(x)$ and by definition,

$$f_X(x) = \frac{d}{dx}P_X(x) \quad (9.11)$$

From the definition of cdf, i.e.,

$$P_X(x) = P(X \leq x)$$

and the definition of the derivative, we obtain

$$f_X(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X \leq x + \Delta x) - P(X \leq x)}{\Delta x} \quad (9.12)$$

Next, we recall (9.7), Page 9–5,

$$P(x_1 < X \leq x_2) = P(x_2) - P(x_1) = P(X \leq x_2) - P(X \leq x_1) \quad (9.13)$$

and we replace x_2 with $x + \Delta x$ and x_1 with x . Then, (9.13) becomes

$$P(x < X \leq x + \Delta x) = P(X \leq x + \Delta x) - P(X \leq x) \quad (9.14)$$

By substitution of (9.14) into (9.12), we obtain

$$f_X(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x < X \leq x + \Delta x)}{\Delta x} \quad (9.15)$$

and in the limit as $\Delta x \rightarrow 0$, we express (9.15) as

$$f_X(x)dx = P(x < X \leq x + dx) \quad (9.16)$$

We call $f_X(x)dx$ the *probability differential*; it represents the probability that the *rv* X lies between x and $x + dx$.

Since

$$P_X(-\infty) = 0 \quad \text{and} \quad f_X(x) = \frac{d}{dx}P_X(x) \quad (9.17)$$

it follows that the area under the pdf $f_X(x)$ curve from $-\infty$ to x , is the probability that *rv* X is less than or equal to x , that is,

$$P_X(x) = \int_{-\infty}^x f_X(\sigma) d\sigma \quad (9.18)$$

where σ is a *dummy variable** of integration.

Properties of the pdf $f_X(x)$

1. $f_X(x) \geq 0$, i.e., the pdf of any rv is always non-negative.

Proof:

From the definition

$$f_X(x) = \frac{d}{dx} P_X(x) \quad (9.19)$$

and recalling that $P_X(x)$ is a monotonic, non-decreasing function of x , its slope is never negative and thus $f_X(x) \geq 0$.

2. The area under any pdf is always unity. In other words,

$$\int_{-\infty}^{\infty} f_X(x) dx = 1 \quad (9.20)$$

Proof:

From (9.18) above,

$$P_X(x) = \int_{-\infty}^x f_X(\sigma) d\sigma$$

and letting the upper limit of integration $x \rightarrow \infty$, we get

$$P_X(\infty) = \int_{-\infty}^{\infty} f_X(x) dx = 1 \quad (9.21)$$

3. The probability that the rv X lies in the range $x_1 < X \leq x_2$ is obtained from the integral

$$\int_{x_1}^{x_2} f_X(x) dx, \text{ that is,}$$

* Since a definite integral depends only on the limits of integration, we can use any variable as the symbol for integration. For instance, $\int_a^b f(x) dx = \int_a^b f(t) dt = \int_a^b f(\tau) d\tau$ etc. For this reason, the variable x, t, τ , etc. is referred to as “dummy variable”.

$$P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f_X(x) dx \quad (9.22)$$

Proof:

From (9.7), Page 9–5,

$$P(x_1 < X \leq x_2) = P(x_2) - P(x_1) = P(X \leq x_2) - P(X \leq x_1)$$

and from (9.18), Page 9–10,

$$P_X(x) = \int_{-\infty}^x f_X(\sigma) d\sigma \quad (9.23)$$

Therefore,

$$P(x_1 < X \leq x_2) = P(x_2) - P(x_1) = \int_{-\infty}^{x_2} f_X(x) dx - \int_{-\infty}^{x_1} f_X(x) dx = \int_{x_1}^{x_2} f_X(x) dx$$

9.5 Two Random Variables

This section requires knowledge of partial differentiation* and double integration†. It may be skipped without loss of continuity.

Let X and Y be two random variables. By definition the *joint probability distribution function (joint cdf)*, denoted as $P_{XY}(x, y)$, is the probability that rv X is less or equal to a specified value x and rv Y is less or equal to a specified value y . The joint sample space in this case is the $x - y$ plane.

The joint cdf $P_{XY}(x, y)$ represents the probability that the outcome of an experiment will result in a sample point lying inside the quadrant

$$-\infty < X \leq x, -\infty < Y \leq y$$

that is,

$$P_{X,Y} = P(X \leq x, Y \leq y) \quad (9.24)$$

If the joint cdf P_{XY} is continuous everywhere, and its partial derivative

* This is differentiation of a variable of a function of two or more variables such as $z = f(x, y)$. Partial differentiation is performed by differentiation with respect to a single variable regarding other variables as constants.

† We use double integration to find the area of a region of an xy -plane. For instance, the double integral $\int_0^1 \int_2^3 dy dx$ represents the area of a region of the xy -plane.

$$f_{X,Y} = \frac{\partial^2}{\partial x \partial y} P_{X,Y}(x,y) \quad (9.25)$$

exists and is continuous everywhere except possibly on a finite set of curves, then, we call $f_{X,Y}$ the *joint probability density function*, or *joint pdf* of the random variables X and Y .

From the definition of the partial derivative and (9.22), we may write:

$$\int_x^{x+\Delta x} \int_y^{y+\Delta y} f_{X,Y}(x,y) dx dy = P(x < X \leq x + dx, y < Y \leq y + dy) \quad (9.26)$$

Relation (9.26) states that $f_{X,Y}(x,y) dx dy$ may be thought of as the probability that the result of an experiment will correspond to a sample point falling in the incremental area $dx dy$ about point x, y in the joint sample space.

The joint cdf $P_{X,Y}(x,y)$ is also a monotonic, non-decreasing function of both x and y . Thus, its derivative is $f_{X,Y}(x,y) \geq 0$ for all x and y . If the joint pdf $f_{X,Y}(x,y)$ is continuous everywhere, then the joint cdf is

$$P_{X,Y}(x,y) = \int_{-\infty}^y \int_{-\infty}^x f_{X,Y}(\sigma, \tau) (d\sigma) d\tau \quad (9.27)$$

The joint cdf $P_{X,Y}(x,y)$ assumes its maximum value at $x = \infty$ and $y = \infty$. Therefore,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(\sigma, \tau) (d\sigma) d\tau = 1 \quad (9.28)$$

If two random variables X and Y are continuous and their probability density functions exist, then X and Y are said to be *statistically independent* if and only if

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) \quad (9.29)$$

9.6 Statistical Averages

By definition, the *mean* m_X or *expected value* $E(X)$ of a continuous rv X is

$$m_X = E(X) = \int_{-\infty}^{\infty} x f_X(x) dx \quad (9.30)$$

Similarly, the expected value of a function of X , say $g(X)$ is defined as

$$E\{g(X)\} = \int_{-\infty}^{\infty} g(x) f_X(x) dx \quad (9.31)$$

For the special case where $g(X) = X^n$, we obtain the n th moment of the rv X which is defined as

$$E(X^n) = \int_{-\infty}^{\infty} x^n f_X(x) dx \quad (9.32)$$

We observe that when $n = 1$, (9.32) gives the expected value $E(X)$, i.e., (9.30).

For $n = 2$, we get the *mean-square* value of X , defined as

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx \quad (9.33)$$

The expected value of a positive integer power n of a rv X is referred to as the *n th moment about the mean** and it is computed from

$$E\{(X - m_X)^n\} = \int_{-\infty}^{\infty} (x - m_X)^n f_X(x) dx \quad (9.34)$$

We observe that the left side of (9.34) is the expected value of the difference between the rv X and its mean m_X . For $n = 1$ we obtain the *first moment*, that is,

$$E\{(X - m_X)\} = \int_{-\infty}^{\infty} (x - m_X) f_X(x) dx = \int_{-\infty}^{\infty} x f_X(x) dx - m_X \int_{-\infty}^{\infty} f_X(x) dx = m_X - m_X(1) = 0 \quad (9.35)$$

Thus, the first moment of an rv X is zero. If, in (9.34) we let $n = 2$, we get the second moment which is known as the *variance* of the rv X . It is denoted as $\text{Var}(X)$ or σ_X^2 . In this case,

$$\begin{aligned} \text{Var}(X) &= \sigma_X^2 = E\{(X - m_X)^2\} = \int_{-\infty}^{\infty} (x - m_X)^2 f_X(x) dx = E(X^2 - 2m_X X + m_X^2) \\ &= E(X^2) - 2m_X E(X) + m_X^2 = E(X^2) - 2m_X m_X + m_X^2 = E(X^2) - 2m_X^2 + m_X^2 \end{aligned}$$

or

$$\text{Var}(X) = \sigma_X^2 = E(X^2) - m_X^2 \quad (9.36)$$

In the special case where $m_X = 0$, the variance σ_X^2 and the mean-square value of X , i.e., $E(X^2)$ are equal as it can be seen from the right side of (9.36).

* In some probability and statistics books, it is also referred to as the r th central moment.

Chapter 9 Random Variables

The variance* provides a measure of the spread or dispersion of the pdf of rv X and its mean m_X . The square root of the variance is called the *standard deviation* of the rv X , and it is denoted as σ_X , that is,

$$\sigma_X = \sqrt{\text{Var}(X)} \quad (9.37)$$

Example 9.3

A continuous rv X has the pdf

$$f_X(x) = \begin{cases} \frac{1}{16}x + k & 0 \leq x \leq 4 \\ 0 & \text{otherwise} \end{cases} \quad (9.38)$$

- Sketch this distribution, and find the value of k .
- Compute the probability that $1 \leq x \leq 3$, that is x lies between 1 and 3.
- Compute the expected value $E[X]$.
- Compute the variance $\text{Var}(X)$ and standard deviation σ_X .

Solution:

- The expression $\frac{1}{16}x + k$ represents a straight line with slope $m = 1/16$ and y -intercept k . It is shown in Figure 9.6.

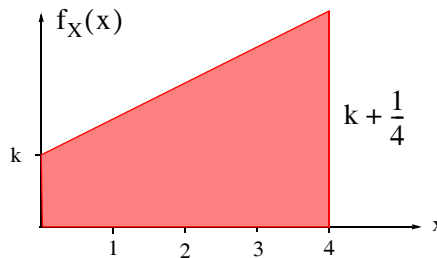


Figure 9.6. Pdf for Example 9.3.

* In analogy with physics, we can think of the mean as the center of mass. This is the point in a system of bodies at which the mass of the system may be considered to be concentrated. It is also called the center of gravity, or centroid. We can also think of the variance as the moment of inertia, which is a measure of a body's resistance to angular acceleration. In physics, the moment of inertia is equal the product of the mass of a particle and the square of its distance from the center of gravity. In probability, the square root of the variance, that is, the standard deviation is a measure of the distance from the mean.

The ordinate $k + \frac{1}{4}$ is found by the substitution of $x = 4$ in (9.38).

Now, if $f_X(x)$ is to be a pdf, the shaded area, which is a trapezoid, must be equal to 1. This area can be found from (3.26), Chapter 3, Page 3–13, that is, the area of a trapezoid. Then, we must have

$$\frac{1}{2} \times 4 \times \left[k + \left(k + \frac{1}{4} \right) \right] = 1 \tag{9.39}$$

Solution of (9.39) yields $k = \frac{1}{8}$ and thus (9.38), Page 9–14, reduces to

$$f_X(x) = \begin{cases} \frac{1}{16}x + \frac{1}{8} & 0 \leq x \leq 4 \\ 0 & \text{otherwise} \end{cases}$$

b. The probability that $P[1 \leq x \leq 3]$ is represented by the shaded area of Figure 9.7. As before, we find this area, with the formula that gives the area for a trapezoid.

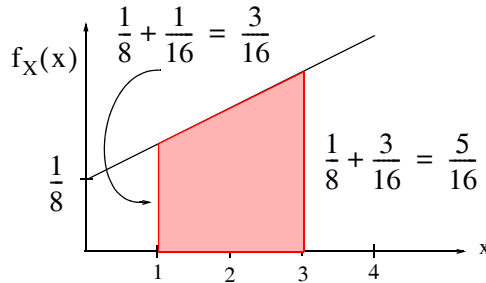


Figure 9.7. Computation of the probability $P[1 \leq x \leq 3]$ for Example 9.3.

The area of the shaded trapezoid is

$$\frac{1}{2} \times 2 \times \left(\frac{3}{16} + \frac{5}{16} \right) = \frac{1}{2} = 0.5$$

Therefore,

$$P[1 \leq x \leq 3] = 50\%$$

c. The expected value is found from (9.30), Page 9–12. For this example,

$$\begin{aligned} m_X = E(X) &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^4 x \left(\frac{1}{16}x + \frac{1}{8} \right) dx \\ &= \int_0^4 \frac{x^2}{16} dx + \int_0^4 \frac{x}{8} dx = \frac{x^3}{48} \Big|_0^4 + \frac{x^2}{16} \Big|_0^4 = \frac{64}{48} + \frac{16}{16} = \frac{7}{3} = 2.33 \end{aligned} \tag{9.40}$$

d. The variance is found from (9.36), Page 9–13, that is,

$$\text{Var}(X) = \sigma_X^2 = E(X^2) - m_X^2 \quad (9.41)$$

We know the value of m_X from part (c), and thus we need only to compute $E(X^2)$. Then, using (9.33), Page 9–13, we obtain

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^4 x^2 \left(\frac{1}{16}x + \frac{1}{8} \right) dx \\ &= \int_0^4 \frac{x^3}{16} dx + \int_0^4 \frac{x^2}{8} dx = \frac{x^4}{64} \Big|_0^4 + \frac{x^3}{24} \Big|_0^4 = \frac{256}{64} + \frac{64}{24} = \frac{20}{3} \end{aligned} \quad (9.42)$$

Next, by substitution of the values of (9.40) and (9.42) into (9.41), we get

$$\text{Var}(X) = \sigma_X^2 = E(X^2) - m_X^2 = \frac{20}{3} - \left(\frac{7}{3} \right)^2 = \frac{11}{9}$$

The standard deviation is

$$\sigma_X = \sqrt{\text{Var}(X)} = \sqrt{\frac{11}{9}} = \frac{\sqrt{11}}{3} = 1.106$$

For a discrete *rv* X that has possible values x_1, x_2, \dots, x_n with probabilities P_1, P_2, \dots, P_n respectively, the expected value $E(x)$ is defined as

$$E(x) = x_1 P_1 + x_2 P_2 + \dots + x_n P_n = \sum_{i=1}^n x_i P_i \quad (9.43)$$

and if x_1, x_2, \dots, x_n have equal probability of occurrence, then

$$E(x) = \mu = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (9.44)$$

The expected value is referred to as the mean or average of x_1, x_2, \dots, x_n , and it is denoted as μ .

The mean square $E(X^2)$ of a discrete *rv* X is

$$E(X^2) = x_1^2 P_1 + x_2^2 P_2 + \dots + x_n^2 P_n = \sum_{i=1}^n x_i^2 P_i \quad (9.45)$$

The variance of a discrete *rv* X having probability function $f(x)$ is given by

$$\sigma_X^2 = E[(X - \mu)^2] = \sum_{i=1}^n (x_i - \mu)^2 f(x_i) = \sum (x - \mu)^2 f(x)$$

and using the same procedure as in the derivation of (9.36), we get

$$\sigma_X^2 = E(X^2) - \mu^2 \tag{9.46}$$

Note 9.1

In a gambling game, the expected value can be used to determine whether the game is favorable to the player or not. If the expected value is positive, the game is considered to be favorable to the player; if negative, the game is unfavorable. If the expected value is zero, the game is considered fair.

Example 9.4

Suppose that a player tosses a fair die. If an odd number occurs, he wins that number of dollars and if an even number occurs, he loses that number of dollars.

- a. Determine whether this game is fair, favorable, or unfavorable to the player.
- b. Compute the standard deviation.

Solution:

- a. Let us denote the odd numbers 1, 3, and 5, as positive since the player wins when these occur, and the even numbers as negative, that is -2, -4, and -6, since he loses when these numbers occur. We show the possible outcomes x_i of the game and their probabilities of occurrence $P[X = x_i]$ in Table 9.5.

TABLE 9.5 Table for Example 9.4

x_i	1	3	5	-2	-4	-6
$P[X = x_i]$	1/6	1/6	1/6	1/6	1/6	1/6

The expected value is found from (9.43), Page 9-16. For this example,

$$\begin{aligned} E(x) &= x_1P_1 + x_2P_2 + \dots + x_nP_n \\ &= 1 \times \frac{1}{6} + 3 \times \frac{1}{6} + 5 \times \frac{1}{6} - 2 \times \frac{1}{6} - 4 \times \frac{1}{6} - 6 \times \frac{1}{6} = -\frac{1}{2} \end{aligned}$$

and since the expected value is negative, we conclude that the game is unfavorable to the player.

- b. The mean square of a discrete rv X is obtained from (9.45), Page 9-16, that is,

$$\begin{aligned} E(X^2) &= x_1^2 P_1 + x_2^2 P_2 + \dots + x_n^2 P_n \\ &= 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} = \frac{91}{6} \end{aligned}$$

and from (9.46), Page 9–17, the variance is

$$\sigma_X^2 = E(X^2) - m_X^2 = \frac{91}{6} - \left(-\frac{1}{2}\right)^2 = \frac{91}{6} - \frac{1}{4} = \frac{179}{12} = 14.92$$

Therefore, the standard deviation is

$$\sigma_X = \sqrt{14.92} = 3.86$$

Another statistical average is the *characteristic function** $\Phi_X(v)$ of rv X which is defined as the expectation of e^{jvX} , that is,

$$\Phi_X(v) = E\{e^{jvX}\} = \int_{-\infty}^{\infty} f_X(x)e^{jvX} dx \quad (9.47)$$

Here, we observe that except for a sign change in the exponent, the characteristic function $\Phi_X(v)$ is the Fourier transform of the pdf $f_X(x)$. Also, in analogy to the Inverse Fourier transform, we have the relation

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_X(v)e^{-jvx} dv \quad (9.48)$$

Example 9.5

Find the pdf of the rv Z defined as the sum of two statistically independent random variables X and Y , that is, $Z = X + Y$.

Solution:

From (9.47) above,

$$\Phi_Z(v) = E\{e^{jvZ}\} = E\{e^{jv(X+Y)}\} = E\{e^{jvX}e^{jvY}\} \quad (9.49)$$

and since X and Y are statistically independent, it follows that

$$\Phi_Z(v) = E\{e^{jvX}\}E\{e^{jvY}\} = \Phi_X(v)\Phi_Y(v) \quad (9.50)$$

* This is an advanced topic of probability theory and may be skipped without loss of continuity.

and by analogy with Fourier analysis that the convolution of two time functions corresponds to multiplication of their Fourier transforms and vice versa, * the pdf $f_Z(z)$ of the rv Z is given by the convolution of the pdf of X and Y , i.e.,

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx \quad (9.51)$$

* This is an advanced topic in probability and may be skipped without loss of continuity.

9.7 Summary

- A sample point s_i is associated with each possible outcome of the experiment.
- The sample space is the total number of sample points $\{s\}$, corresponding to the aggregate of all possible outcomes of an experiment.
- An event may correspond to a single sample point or a set of sample points.
- A random variable (*rv*) is a variable that can wander over the set of sample points, and whose value is determined by an experiment.
- A discrete *rv* can take on only a finite number of values in any finite observation interval.
- A continuous *rv* can take on any value in the entire observation interval.
- A probability function (or probability distribution) is defined as $P(X = x) = f(x)$ where $f(x) \geq 0$ and $\sum_x f(x) = 1$
- The cumulative distribution function (cdf) or probability distribution function $P_X(x)$ is the probability that a *rv* X will be equal to or less than some value x . In other words, $P_X(x) = P(X \leq x)$
- A relative frequency table shows how often values occur within a range of values.
- $P_X(x)$ is bounded between zero and one, that is, $P_X(-\infty) = 0$ and $P_X(\infty) = 1$
- $P_X(x)$ is a monotonic, non-decreasing function of x , i.e., if $x_1 < x_2$, then $P_X(x_1) \leq P_X(x_2)$
- $P_X(x)$ is continuous on the right.
- The probability density function (pdf) is more useful when we need to compute statistical averages. It is denoted as $f_X(x)$ and defined as $f_X(x) = \frac{d}{dx}P_X(x)$
- $f_X(x) \geq 0$, i.e., the pdf of any *rv* is always non-negative.
- The area under any pdf is always unity. In other words, $\int_{-\infty}^{\infty} f_X(x)dx = 1$
- The probability that the *rv* X lies in the range $x_1 < X \leq x_2$ is obtained from the integral

$$\int_{x_1}^{x_2} f_X(x)dx, \text{ that is, } P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f_X(x)dx$$

- The mean m_X or expected value $E(X)$ of a continuous *rv* X is $m_X = E(X) = \int_{-\infty}^{\infty} xf_X(x)dx$
- The mean-square value of X , defined as $E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)dx$
- The expected value of a positive integer power n of a *rv* X is referred to as the n th moment about the mean and it is computed from $E\{(X - m_X)^n\} = \int_{-\infty}^{\infty} (x - m_X)^n f_X(x)dx$
- The first moment of an *rv* X is zero.
- The second moment which is known as the *variance* of the *rv* X is denoted as $\text{Var}(X)$ or σ_X^2 and is related to the mean square value and the mean as $\text{Var}(X) = \sigma_X^2 = E(X^2) - m_X^2$
- The variance provides a measure of the spread or dispersion of the pdf of *rv* X and its mean m_X .
- The square root of the variance is called the standard deviation of the *rv* X , and it is denoted as σ_X , that is, $\sigma_X = \sqrt{\text{Var}(X)}$
- For a discrete *rv* X that has possible values x_1, x_2, \dots, x_n with probabilities P_1, P_2, \dots, P_n respectively, the expected value $E(x)$ is defined as $E(x) = x_1P_1 + x_2P_2 + \dots + x_nP_n = \sum_{i=1}^n x_iP_i$
and if x_1, x_2, \dots, x_n have equal probability of occurrence, then $E(x) = \mu = \frac{x_1 + x_2 + \dots + x_n}{n}$
- The mean square $E(X^2)$ of a discrete *rv* X is $E(X^2) = x_1^2P_1 + x_2^2P_2 + \dots + x_n^2P_n = \sum_{i=1}^n x_i^2P_i$
- The variance of a discrete *rv* X having probability function $f(x)$ is given by

$$\sigma_X^2 = E[(X - \mu)^2] = \sum_{i=1}^n (x_i - \mu)^2 f(x_i) = \sum (x - \mu)^2 f(x)$$

and

$$\sigma_X^2 = E(X^2) - \mu^2$$

Chapter 9 Random Variables

- In a gambling game, the expected value can be used to determine whether the game is favorable to the player or not. If the expected value is positive, the game is considered to be favorable to the player; if negative, the game is unfavorable. If the expected value is zero, the game is considered fair.
- Another statistical average is the *characteristic function* $\Phi_X(v)$ of rv X which is defined as the expectation of e^{jvX} , that is,

$$\Phi_X(v) = E\{e^{jvX}\} = \int_{-\infty}^{\infty} f_X(x)e^{jvX}dx$$

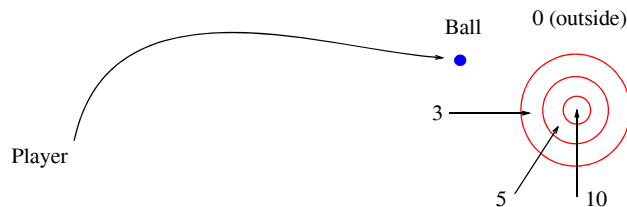
9.8 Exercises

- A fair coin is flipped four times.
 - List the possible outcomes
 - Find the probability of occurrence of each of the possible outcomes.
 - Find the probability that at most two heads occur.
 - Find the probability that at least two heads occur.
- In a gambling game, a player tosses a fair die and wins if 2, 3 and 5 occurs and his winnings are that number of dollars. He loses if 1, 4 or 6 occurs and his losses are that number of dollars. Is this game favorable to the player?
- An rv X has the distribution shown below.

x_i	-2	1	2	4
$P[x = x_i]$	1/4	1/8	1/2	1/8

Show the cdf in a graph form.

- A fair coin is tossed until a head or five tails occur. Compute the expected value $E(X)$ of the tosses of the coin for this to occur.
- The concentric circles of the figure below represent a game in a theme park where a ball, from some distance, is thrown into a box with three holes represented by the three concentric circles. The numbers indicate the number of points won in each case. Thus, a player wins 10, 5, 3, or 0 points if he throws the ball inside the small, middle, large or outside the large circle respectively. The probability of winning any of these four points is $1/2$ and it is just as likely to throw the ball inside any of these holes or outside. Compute the expected value $E(X)$ of points earned each time a player throws a ball.



- A player tosses two fair coins and wins \$5 if two heads occur, \$2 if one head occurs, and \$1 if no heads occur.
 - Compute his expected winnings.
 - What is the maximum dollar amount he should pay to play the game if the casino requires a fee to be paid in order to play?

9.9 Solutions to End-of-Chapter Exercises

- 1.
- a.
- | | | | |
|------|------|------|------|
| HHHH | HTHH | THHH | TTHH |
| HHHT | HTHT | THHT | TTHT |
| HHTH | HTTH | THTH | TTTH |
| HHTT | HTTT | THTT | TTTT |
- b. The experiments (flips of the coin) are independent of each other and the coin is fair. Therefore, the probability of each occurrence is $1/16$
- c.

$$\begin{aligned} P[\# \text{ of Heads} \leq 2] &= P[\text{zero Heads}] + P[1 \text{ Head}] + P[2 \text{ Heads}] \\ &= \frac{1}{16} + \frac{4}{16} + \frac{6}{16} = \frac{11}{16} \end{aligned}$$

d.

$$\begin{aligned} P[\# \text{ of Heads} \geq 2] &= P[2 \text{ Heads}] + P[3 \text{ Heads}] + P[4 \text{ Heads}] \\ &= \frac{6}{16} + \frac{4}{16} + \frac{1}{16} = \frac{11}{16} \end{aligned}$$

- 2.
- The possible outcomes x_i of the game and their respective probabilities are shown below where the negative numbers indicate that the player loses if a non-prime number occurs.

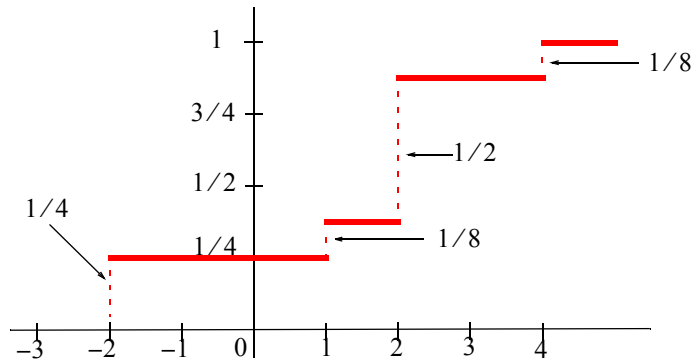
x_i	2	3	5	-1	-4	-6
$P[X = x_i]$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

The expected value is

$$E[X] = 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} - 1 \cdot \frac{1}{6} - 4 \cdot \frac{1}{6} - 6 \cdot \frac{1}{6} = -\frac{1}{6}$$

and since the result is negative, the game is unfavorable to the player.

3.



4.

Only one toss occurs if heads occurs the first time, i.e., the event H. Two tosses occur if the first is tails and the second is heads, i.e., the event TH, and so on. Therefore,

$$P[X = x_1] = P[H] = P[X = 1] = 1/2$$

$$P[X = x_2] = P[TH] = P[X = 2] = P[T] \cdot P[H] = 1/2 \times 1/2 = 1/4$$

$$P[X = x_3] = P[TTH] = P[X = 3] = 1/2 \times 1/2 \times 1/2 = 1/8$$

$$P[X = x_4] = P[TTTH] = P[X = 4] = 1/16$$

$$P[X = x_5] = P[TTTTH] = P[X = 5] = 1/32$$

$$\begin{aligned} \text{or } & P[TTTTT] = P[X = 5] = 1/32 \\ & = 1/32 + 1/32 = 1/16 \end{aligned}$$

Then,

$$E[X] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + 5 \cdot \frac{1}{16} = \frac{31}{16}$$

5.

$$P[X = 10] = \frac{1}{2} \cdot \frac{\text{Area of 10 points}}{\text{Area of target}} = \frac{1}{2} \cdot \frac{\pi(1)^2}{\pi(5)^2} = \frac{1}{50}$$

$$P[X = 5] = \frac{1}{2} \cdot \frac{\text{Area of 5 points}}{\text{Area of target}} = \frac{1}{2} \cdot \frac{\pi(3)^2 - \pi(1)^2}{\pi(5)^2} = \frac{8}{50}$$

$$P[X = 3] = \frac{1}{2} \cdot \frac{\text{Area of 3 points}}{\text{Area of target}} = \frac{1}{2} \cdot \frac{\pi(5)^2 - \pi(3)^2}{\pi(5)^2} = \frac{16}{50}$$

$$P[X = 0] = \text{Area outside the outer circle} = 0$$

and

$$E[X] = 10 \cdot \frac{1}{50} + 5 \cdot \frac{1}{50} + 3 \cdot \frac{16}{50} + 0 = \frac{98}{50} = \frac{49}{25}$$

6.

a.

The probability of winning \$5 = $1/4$

The probability of winning \$2 = $1/2$

The probability of winning \$1 = $1/4$

and thus the expected value is

$$E[X] = 5 \cdot \frac{1}{4} + 2 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = \frac{10}{4} = \frac{5}{2}$$

or

$$E[X] = \$2.50$$

b.

If the player pays less than \$2.50, the game is favorable to the player.

If the player pays exactly \$2.50, the game is fair.

If the player pays more than \$2.50, the game is unfavorable to the player.

Chapter 10

Common Probability Distributions and Tests of Significance

This chapter is a continuation of the previous chapter. It discusses several probability distributions, percentiles, Chebyshev's inequality, law of large numbers, sampling distribution of means, tests of hypotheses, levels of significance, and the z , t , F , and χ^2 (chi-square) tests.

10.1 Properties of Binomial Coefficients

The notation

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} \quad (10.1)$$

where n represents the number of trials in an experiment, and k the probability that an event will occur, is used extensively in probability theory.

For $k = 0, 1, 2,$ and 3 and making use of $0! = 1^*$, we obtain

$$\binom{n}{0} = \frac{n!}{(n-0)!0!} = \frac{n!}{n!1} = 1 \quad (10.2)$$

$$\binom{n}{1} = \frac{n!}{(n-1)!1!} = n \quad (10.3)$$

$$\binom{n}{2} = \frac{n!}{(n-2)!2!} = \frac{n(n-1)}{2!} \quad (10.4)$$

$$\binom{n}{3} = \frac{n!}{(n-3)!3!} = \frac{n(n-1)(n-2)}{3!} \quad (10.5)$$

$$\binom{n}{n} = \frac{n!}{0!n!} = 1 \quad (10.6)$$

$$\binom{n}{n-1} = \frac{n!}{(n-n+1)!(n-1)!} = \frac{n!}{1!(n-1)!} = n \quad (10.7)$$

* By definition of the factorial, $n! = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 1$. Therefore, $(n+1)! = (n+1) \cdot n!$ and for $n = 0$, the last expression reduces to $1! = 1 \cdot 0!$ and thus $0! = 1$.

and if $a + b = n$

$$\binom{n}{a} = \binom{a+b}{a} = \frac{(a+b)!}{(a+b-a)!a!} = \frac{(a+b)!}{b!a!} \quad (10.8)$$

$$\binom{n}{b} = \binom{a+b}{b} = \frac{(a+b)!}{(a+b-b)!a!} = \frac{(a+b)!}{a!b!} \quad (10.9)$$

and therefore,

$$\binom{n}{a} = \binom{n}{b} \quad (10.10)$$

if $a + b = n$

The relation

$$\begin{aligned} (a+b)^n &= a^n + na^{n-1}b + \frac{n(n-1)}{2!}a^{n-2}b^2 + \frac{n(n-1)(n-2)}{3!}a^{n-3}b^3 + \dots + ab^{n-1} + b^n \\ &= \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \end{aligned} \quad (10.11)$$

is known as the *binomial expansion*^{*}, *binomial theorem*, or *binomial series*.

10.2 The Binomial (Bernoulli) Distribution

The *Binomial (Bernoulli) distribution* is defined as

$$b(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} \quad (10.12)$$

for $k = 0, 1, 2, \dots, n$ and $0 < p < 1$. Therefore, $b(k;n;p)$ is non-negative.

From (10.12) and (10.11),

$$\sum_{k=0}^n b(k;n,p) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = [p + (1-p)]^n = 1 \quad (10.13)$$

and thus the binomial distribution $b(k;n;p)$ is a pdf of the *discrete type*.

In (10.13), we call p the probability of a success. Likewise, we call q the probability of a failure.

^{*} The expansion of $(a+b)^n$ contains $n+1$ terms. Formulated in medieval times, the binomial theorem was developed (about 1676) for fractional exponents by the English scientist Sir Isaac Newton, enabling him to apply his newly discovered methods of calculus to many difficult problems. The binomial theorem is useful in various branches of mathematics, particularly in the theory of probability.

Then,

$$q = 1 - p \tag{10.14}$$

For $n = 2$ and $p = 1/2$, (10.12) becomes

$$b(k;2,\frac{1}{2}) = \binom{2}{k} \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{2-k} = \frac{2!}{(2-k)!k!} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{2-k} = \frac{2!}{(2-k)!k!} \left(\frac{1}{2}\right)^2 \tag{10.15}$$

Moreover, for $k = 0, 1$ and 2 , (10.15) reduces respectively to

$$b(0;2,\frac{1}{2}) = \frac{2!}{2!0!} \left(\frac{1}{2}\right)^2 = 0.25$$

$$b(1;2,\frac{1}{2}) = \frac{2!}{1!1!} \left(\frac{1}{2}\right)^2 = 0.50$$

$$b(2;2,\frac{1}{2}) = \frac{2!}{0!2!} \left(\frac{1}{2}\right)^2 = 0.25$$

The pdf and cdf of the binomial distribution for $n = 2$, $p = 1/2$, and $k = 0, 1$, and 2 are shown in Figures 10.1 and 10.2. We observe that they are the same as in Example 9.1, Chapter 9, Page 9–7.

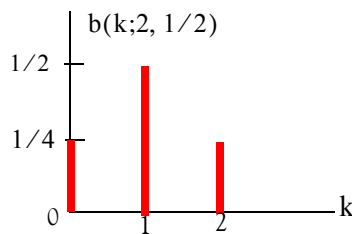


Figure 10.1. Probability density function for the binomial distribution

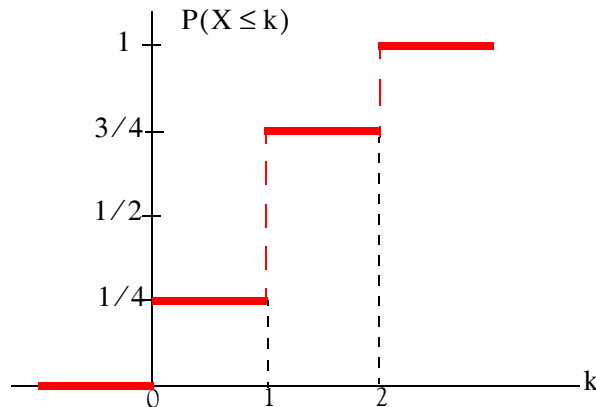


Figure 10.2. Cumulative distribution function for the binomial distribution.

Example 10.1

A coin is tossed six times. Compute:

- the probability that *exactly* (no more, no less) two heads show up
- the probability of getting at *least* four heads
- the probability of *no* heads.
- the probability of at *least* one head.

Solution:

Here, $n = 6$. Let us denote p as a success if heads show up, and q as a failure if tails show up. Obviously, $q = 1 - p$ and $p = q = 1/2$.

- a. We use the binomial distribution with $k = 2$ successes, that is,

$$b(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} p^k q^{n-k} = \binom{6}{2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{6-2}$$

or

$$\frac{6!}{(6-2)!2!} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{6-2} = \frac{6!}{4!2!} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^4 = \frac{6 \times 5}{2 \times 1} \left(\frac{1}{4}\right) \left(\frac{1}{16}\right) = \frac{15}{64}$$

- b. The probability of getting at least four heads is obtained by adding the probabilities with $k = 4$, $k = 5$, and $k = 6$. Then,

$$\begin{aligned} \binom{n}{k} p^k (1-p)^{n-k} &= \binom{6}{4} \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^2 + \binom{6}{5} \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^1 + \binom{6}{6} \left(\frac{1}{2}\right)^6 \\ &= \frac{15}{64} + \frac{6}{64} + \frac{1}{64} = \frac{22}{64} = \frac{11}{32} \end{aligned}$$

- c. The probability of no heads (all tails), is the probability of all failures. Thus,

$$q^6 = \left(\frac{1}{2}\right)^6 = \frac{1}{64}$$

- d. The probability of at least one head is

$$1 - q^6 = 1 - \frac{1}{64} = \frac{63}{64}$$

Example 10.2

A die is tossed seven times. Compute the probability that:

- a. a 5 or a 6 occurs exactly three times
- b. a 5 or a 6 never occurs
- c. a 5 or 6 occurs at least once

Solution:

If a 5 or 6 shows up, we will call it a success p .

- a. If a 5 or 6 occurs three times, then $p = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$ and $k = 3$. Therefore,

$$\begin{aligned} \binom{n}{k} p^k (1-p)^{n-k} &= \binom{7}{3} \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^4 = \frac{7!}{(7-3)! \times 3!} \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^4 \\ &= \frac{7 \times 6 \times 5}{3 \times 2} \left(\frac{1}{27}\right) \left(\frac{16}{81}\right) = \frac{560}{2187} = 0.256 = 25.6\% \end{aligned}$$

- b. The probability that 5 or 6 never occurs in seven trials, is equivalent to seven failures. A failure q is

$$q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3}$$

and thus the probability of seven failures is

$$q^7 = \left(\frac{2}{3}\right)^7 = \frac{128}{2187} = 0.059 = 5.9\%$$

- c. The probability that a 5 or 6 occurs at least once is 1 minus the probability that 5 or 6 never occurs. Therefore,

$$1 - q^7 = 1 - \frac{128}{2187} = \frac{2059}{2187} = 0.942 = 94.2\%$$

The mean μ and the variance σ^2 for the binomial distribution are:

$$\mu = np \tag{10.16}$$

and

$$\sigma^2 = npq \tag{10.17}$$

where

n = number of samples, p = probability of success, and q = probability of failure.

Example 10.3

It is estimated that 80% of PC users buy a new computer every three years. In a group of 4000 PC users, compute

- the mean number of users using the same PC after three years
- the standard deviation

Solution:

Here, $n = 4000$ and $p = 0.8$. Then,

- Using (10.16), Page 10–5, we find that the mean μ is

$$\mu = np = 4000 \times 0.8 = 3200$$

- The standard deviation σ is the square root of the variance. Then, from (10.17), Page 10–5,

$$\sigma = \sqrt{npq} = \sqrt{4000 \times 0.8 \times 0.2} = 25.3$$

10.3 The Uniform Distribution

Let X be an *rv* with pdf $f_X(x)$. The *rv* X is said to be *uniformly distributed* in the interval $a \leq x \leq b$ if

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases} \quad (10.18)$$

The pdf of the uniform distribution is shown in Figure 10.3.

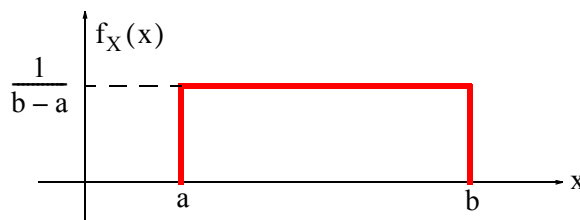


Figure 10.3. The pdf of the uniform distribution.

From Figure 10.3 we observe that

$$f_X(x) \geq 0$$

and

$$\text{Area} = \int_a^b f_X(x) dx = 1$$

The cdf of the uniform distribution is found from

$$P_X(x) = \int_a^b f_X(x) dx \tag{10.19}$$

and it is shown in Figure 10.4.

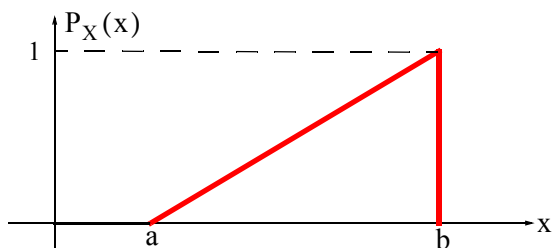


Figure 10.4. The cdf of the uniform distribution

We observe that in Figure (10.4),

$$\text{slope} \Big|_a^b = \frac{\text{rise}}{\text{run}} = \frac{1-0}{b-a} = \frac{1}{b-a}$$

and thus the equation of the straight line of the cdf, in terms of the slope and the vertical axis intercept c (not shown), is

$$P_X(x) = (\text{slope}) \times x + c = \frac{x}{b-a} + c \tag{10.20}$$

To find the intercept c , we evaluate $P_X(x)$ at $x = a$. Then,

$$P_X(x) \Big|_{x=a} = 0 = \frac{a}{b-a} + c$$

Therefore,

$$c = -\frac{a}{b-a} \tag{10.21}$$

and by substitution (10.21) into (10.20), we obtain

$$P_X(x) = \frac{x}{b-a} - \frac{a}{b-a} = \frac{x-a}{b-a}$$

Therefore, the cdf of the uniform distribution for the entire interval $0 \leq x \leq \infty$ is

$$P_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases} \quad (10.22)$$

Example 10.4

A Silicon Valley commuter has determined that the commuting time from home to work varies between 55 and 75 minutes. Compute the probability that in any day, the commuting time will be between 60 and 70 minutes.

Solution:

The pdf of the uniform distribution for this example is shown in Figure 10.5 where $b = 75$ and $a = 55$.

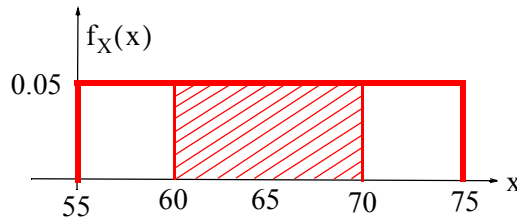


Figure 10.5. The pdf of Example 10.4

Then, $b - a = 20$, and $1/(b - a) = 1/20 = 0.05$. The probability that the commuting time will be between 60 and 70 minutes is shown by the cross-hatched area whose value (probability) is

$$\text{Probability} \Big|_{60}^{70} = \text{Area} \Big|_{60}^{70} = (70 - 60) \times 0.05 = 0.5 \text{ or } 50\%$$

Example 10.5

Prove that the mean μ and variance σ^2 of the uniform distribution are given by

$$\mu = \frac{a + b}{2} \quad (10.23)$$

and

$$\sigma^2 = \frac{(b - a)^2}{12} \quad (10.24)$$

respectively.

Proof:

From Chapter 9, equations (9.30) and (9.36), Pages 9–12 and 9–13 respectively,

$$\mu = m_X = E(X) = \int_{-\infty}^{\infty} x f_X(x) dx \quad (10.25)$$

and

$$\text{Var}(X) = \sigma_X^2 = E\{(X - m_X)^2\} = E(X^2) - m_X^2 \quad (10.26)$$

Then, for the uniform distribution, the mean or expected value is

$$\mu = m_X = E(X) = \int_a^b x \frac{1}{b-a} dx = \frac{x^2}{2(b-a)} \Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}^*$$

and the variance is

$$\text{Var}(X) = \sigma_X^2 = E\{(X - m_X)^2\} = E(X^2) - m_X^2 = E(X^2) - \mu^2$$

where

$$E(X^2) = \int_a^b x^2 \frac{1}{b-a} dx = \frac{x^3}{3(b-a)} \Big|_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}^\dagger$$

and thus

$$\text{Var}(X) = \sigma_X^2 = E(X^2) - \mu^2 = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

Example 10.6

A food distributor sells 10 lb. bags of potatoes at a certain price. It is known that the actual weight of each bag varies between 10.75 and 11.75 lbs. Compute:

- a. the mean (average) weight of each bag
- b. the variance and the standard deviation
- c. the probability that a bag weighs more than 10 lbs.

Solution:

- a. The weight of each bag is uniformly distributed and thus from (10.23), Page 10–8,

$$\mu = \frac{a+b}{2} = \frac{9.75 + 10.75}{2} = 10.25$$

- b. The variance and standard deviation are found from (10.24), Page 10–8. Then,

* We have used the identity $(x^2 - y^2) = (x + y)(x - y)$

† We have used the identity $x^3 - y^3 = (x - y)(x^2 + xy + y^2)$

$$\text{Var}(X) = \sigma_X^2 = \frac{(b-a)^2}{12} = \frac{(10.75-9.75)^2}{12} = \frac{1}{12}$$

and

$$\sigma = \sqrt{\frac{1}{12}} = 0.29$$

- c. The probability that a bag weighs more than 10 lbs. is denoted by $P(X > 10)$. It is shown in Figure 10.6. The probability is represented by the cross-hatched area whose value is $1 \times 0.75 = 0.75$. Thus, $P(X > 10) = 0.75$, that is, there is a probability that 75% of the bags weigh more than 10 lbs.

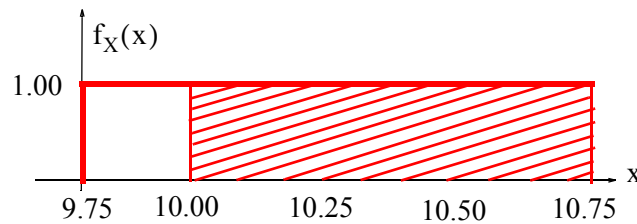


Figure 10.6. The pdf of Example 10.6

10.4 The Exponential Distribution

The pdf of the *exponential distribution* is defined as

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x} & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (10.27)$$

It can be shown that the mean of the exponential distribution is

$$\mu = 1/\alpha \quad (10.28)$$

and the variance is

$$\sigma^2 = 1/\alpha^2 \quad (10.29)$$

The exponential distribution is also expressed as

$$f_X(x) = \begin{cases} \frac{1}{\mu} e^{-(x/\mu)} & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (10.30)$$

By integrating the pdf, we find that the cdf of the exponential distribution is

$$P(X \leq a) = 1 - e^{-x/\mu} \quad (10.31)$$

The pdf of the exponential distribution is shown in Figure 10.7 and the cdf in Figure 10.8.

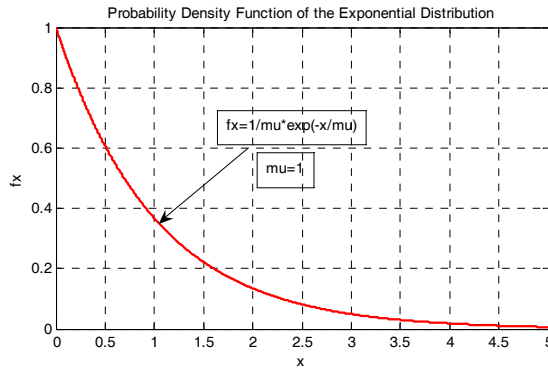


Figure 10.7. The pdf of the exponential distribution

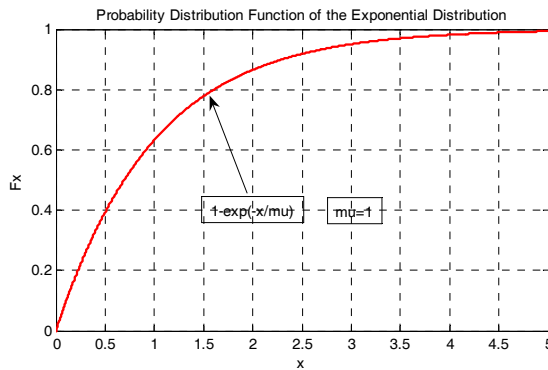


Figure 10.8. The cdf of the exponential distribution

Example 10.7

The lifetimes for a certain brand of computer peripherals are exponentially distributed with a mean $\mu = 4$ years. Compute the probability that these peripherals have lifetimes between 4 and 8 years.

Solution:

The probability $P(4 < x < 8)$ is computed from the cross-hatched area in Figure 10.9. For this example,

$$P(4 < x < 8) = P(x \leq 8) - P(x \leq 4) \quad (10.32)$$

From (10.31), Page 10–10,

$$P(x \leq 8) = 1 - e^{-(8/4)}$$

and

$$P(x \leq 4) = 1 - e^{-(4/4)}$$

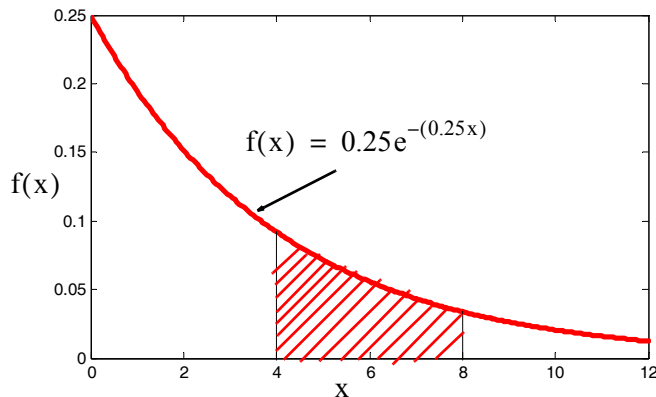


Figure 10.9. Computation of the pdf for Example 10.7

By substitution into (10.32), Page 10–11,

$$P(4 < x < 8) = (1 - e^{-(8/4)}) - (1 - e^{-(4/4)}) = 1 - e^{-2} - 1 + e^{-1} = 0.233 \quad (10.33)$$

Therefore, we can say that there is a probability of 23.3% that the peripherals have lifetimes between 4 and 8 years.

The result can also be found with Excel's EXPONDIST function whose syntax is

=EXPONDIST(x,lambda,cumulative)

where

x = value of the function

lambda = the parameter value

cumulative = exponential function. If TRUE, the cdf is returned, and if FALSE, the pdf is returned.

With the data given in Example 10.7, $\lambda = 1/\mu = 1/4$. Then,

=EXPONDIST(8,1/4,TRUE) returns 0.865, while

=EXPONDIST(4,1/4,TRUE) returns 0.632.

The area in the interval $4 \leq x \leq 8$ is $0.865 - 0.632 = 0.233$. This is the same answer as the one we found using (10.32), Page 10–11.

10.5 The Normal (Gaussian) Distribution

The *normal distribution* is the most widely used distribution in scientific and engineering applications, demographic studies, poll results, and business applications.

Suppose we gather and plot the heights (in feet and inches) of several people versus the number of these people. If many samples are collected, we obtain the curve shown in Figure 10.10. This is the so-called *bell curve* of the normal distribution.

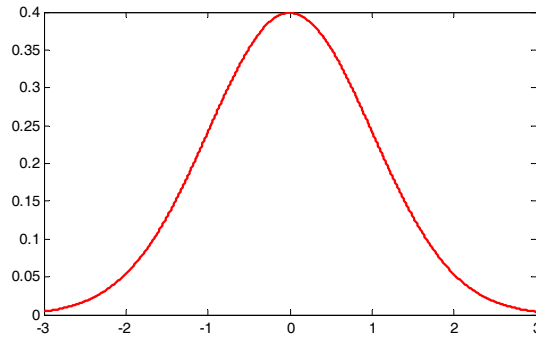


Figure 10.10. The shape of the normal distribution.

The pdf of the normal distribution is defined as

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-(x-\mu)^2/2\sigma_X} \quad (10.34)$$

where μ = mean and σ_X = standard deviation.

If the mean is zero, i.e., if $\mu = 0$, (10.34) reduces to

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-x^2/2\sigma_X} \quad (10.35)$$

The cdf of the normal distribution is

$$P_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^x e^{-(x-\mu)^2/2\sigma_X} dx \quad (10.36)$$

The normal distribution is usually expressed in *standardized form*, where the standardized *rv* Z is referred to as the standardized *rv* X and, in this case, Z is related to X as

$$Z = \frac{X - \mu}{\sigma_X} \quad (10.37)$$

Then, the mean of the *rv* Z is $\mu = 0$ and the variance is one, that is, $\sigma_X^2 = 1$. The pdf of the standardized normal distribution is then simplified to

$$f_z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad (10.38)$$

and the corresponding cdf becomes

$$P_z(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-u^2/2} du = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^z e^{-u^2/2} du \quad (10.39)$$

where u is a dummy variable of integration.

Figure 10.11 shows the *standard normal curve*. It also shows the areas within 1, 2 and 3 standard deviations from the mean taken as zero.

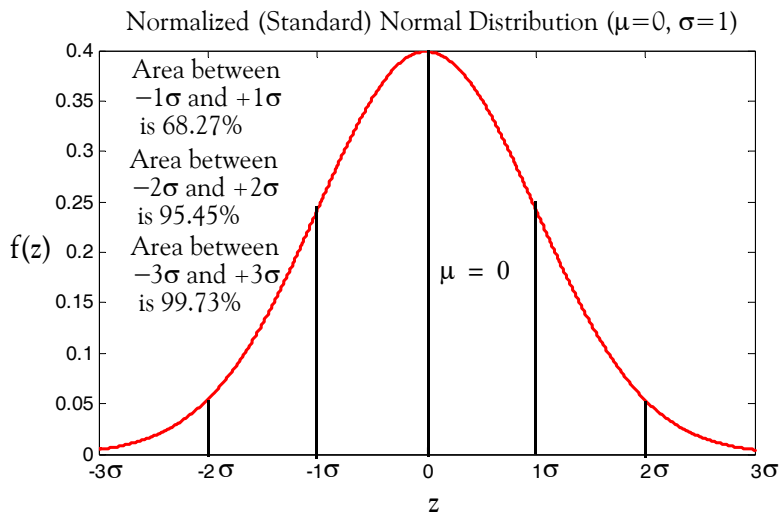


Figure 10.11. The standardized normal curve

As shown in Figure 10.11, the area for the interval $-1\sigma \leq z \leq +1\sigma$ (one standard deviation) is equal to approximately 68.27% of the total area, and the areas for the intervals $-2\sigma \leq z \leq +2\sigma$ (two standard deviations) and $-3\sigma \leq z \leq +3\sigma$ (three standard deviations), are 95.45% and 99.73% of the total area respectively.

Example 10.8

In a college physics examination, the mean and standard deviation are 74 and 12 respectively. Compute the scores in standard units of students receiving the grades

- a. 65 b. 74 c. 86 d. 92

Solution:

Using (10.37), Page 10–13, we obtain:

a.
$$Z_1 = \frac{x_1 - \mu}{\sigma} = \frac{65 - 74}{12} = -0.75$$

b.
$$Z_2 = \frac{x_2 - \mu}{\sigma} = \frac{74 - 74}{12} = 0.00$$

c.
$$Z_3 = \frac{x_3 - \mu}{\sigma} = \frac{86 - 74}{12} = 1.00$$

d.
$$Z_4 = \frac{x_4 - \mu}{\sigma} = \frac{92 - 74}{12} = 1.50$$

Example 10.9

The mean and standard deviation on a college mathematics examination are 74 and 12 respectively. Compute the grades corresponding to the standard scores

- a. -1.00 b. 0.50 c. 1.25 d. 1.75

Solution:

Solving (10.37), Page 10–13, for x , we obtain

$$x = \sigma Z + \mu \tag{10.40}$$

Then,

a.
$$x_1 = \sigma Z_1 + \mu = 12(-1) + 74 = 62$$

b.
$$x_2 = \sigma Z_2 + \mu = 12(0.50) + 74 = 80$$

c.
$$x_3 = \sigma Z_3 + \mu = 12(1.25) + 74 = 89$$

d.
$$x_4 = \sigma Z_4 + \mu = 12(1.75) + 74 = 95$$

We can use Excel's **STANDARDIZE** function to normalize a value from a distribution with a known mean and known standard deviation. The syntax is

STANDARDIZE(x,mean,standard_dev)

where

x = value to be normalized

mean = mean of the distribution

standard_dev = standard deviation of the distribution

For example, the normalized value of 50 from a distribution with $\mu = 25$ and $\sigma = 2.8$ is found with **=STANDARDIZE(50,25,2.8)** which returns $z = 8.929$.

Chapter 10 Common Probability Distributions and Tests of Significance

We can obtain probability values from the normalized normal curve from published tables; however, we must be careful how to use these tables. For special and repeated applications, it is preferable to construct a spreadsheet, from which we can obtain quick and accurate answers to a variety of questions. We will refer to the spreadsheet of Figure 10.12 to answer the questions posed in Example 10.10 that follows. We will provide instructions to construct this spreadsheet at the conclusion of this example.

Example 10.10

Suppose we are employed by a manufacturing company that makes computer memory chips. Industry standards, which take into consideration life, size, tolerance, and reliability, have established ratings that vary from 3 to 9 where 3 is the lowest and 9 the highest. For the past ten years, our facility has received the following ratings:

3, 4, 3, 5, 7, 6, 7, 8, 9, 8

We can compute the mean μ and standard deviation σ with the Excel functions **AVERAGE(number1,number2,...)** and **STDEV(number1,number2,...)** respectively. Then,

=AVERAGE(3,4,3,5,7,6,7,8,9,8) returns 6. Also,

=STDEV(3,4,3,5,7,6,7,8,9,8) returns 2.16 which, for convenience, we round to 2.

Therefore, for this example $\mu = 6$ and $\sigma = 2$

The management needs answers to the following questions assuming that our competitors' ratings have the same mean and standard deviation.

Question 1:

If one of our competitors has a rating of 8 points in the current year, what is the probability that he ranks among the top 5% of memory chip manufacturers?

We can answer this question by referring to the table (spreadsheet) of Figure 10.12.

The X values in Column A represent memory chip ratings, and we find that the mean $\mu = 6$ appears in A16. To find the probability that chips have a rating less than 6, that is, less than the mean, we look at E16 and we see that there is 49.99% probability that the chips have a rating less than 6. In other words, there is a probability of almost 50% that the chips have a rating below the mean. We also find, from E21, that there is a probability of 84.04% that the chips have a rating less than 8. To find the probability that our competitor is in the top 5%, we look for a value of 95% in Column E. This percentage is between 94.45% and 96.35%, and Column A shows that the 95% probability corresponds to a rating between 11.2 and 11.6 as indicated in A24 and A25. Therefore, we conclude that the probability that our competitor is in the top 5% is zero.

	X	Z	Curve	Bar	Left	Right	Z to	Both	Between	
	value	value	height	area	side	side	mean	ends	ends	
1	0.00	-3.00	0.00443	0.13%	0.13%	99.87%	49.87%	0.26%	99.74%	
2	0.40	-2.80	0.00792	0.12%	0.25%	99.75%	49.75%	0.51%	99.49%	
3	0.80	-2.60	0.01358	0.21%	0.47%	99.53%	49.53%	0.94%	99.06%	
4	1.20	-2.40	0.02239	0.36%	0.83%	99.17%	49.17%	1.66%	98.34%	b. exceeds a certain value (right side - Column F)
5	1.60	-2.20	0.03547	0.58%	1.41%	98.59%	48.59%	2.81%	97.19%	
6	2.00	-2.00	0.0540	0.89%	2.30%	97.70%	47.70%	4.60%	95.40%	
7	2.40	-1.80	0.0790	1.33%	3.63%	96.37%	46.37%	7.26%	92.74%	
8	2.80	-1.60	0.1092	1.90%	5.53%	94.47%	44.47%	11.06%	88.94%	
9	3.20	-1.40	0.1497	2.61%	8.14%	91.86%	41.86%	16.27%	83.73%	
10	3.60	-1.20	0.1942	3.44%	11.58%	88.42%	38.42%	23.15%	76.85%	
11	4.00	-1.00	0.2420	4.36%	15.94%	84.06%	34.06%	31.87%	68.13%	c. is between mean and z (z to mean - Column G)
12	4.40	-0.80	0.2897	5.32%	21.25%	78.75%	28.75%	42.51%	57.49%	
13	4.80	-0.60	0.3332	6.23%	27.48%	72.52%	22.52%	54.97%	45.03%	
14	5.20	-0.40	0.3683	7.01%	34.50%	65.50%	15.50%	69.00%	31.00%	
15	5.60	-0.20	0.3910	7.59%	42.09%	57.91%	7.91%	84.18%	15.82%	
16	6.00	0.00	0.3989	7.90%	49.99%	50.01%	0.01%	99.98%	0.02%	
17	6.40	0.20	0.3910	7.90%	57.89%	42.11%	7.89%	84.22%	15.78%	
18	6.80	0.40	0.3683	7.59%	65.48%	34.52%	15.48%	69.03%	30.97%	
19	7.20	0.60	0.3332	7.01%	72.50%	27.50%	22.50%	55.00%	45.00%	
20	7.60	0.80	0.2897	6.23%	78.73%	21.27%	28.73%	42.54%	57.46%	
21	8.00	1.00	0.2420	5.32%	84.04%	15.96%	34.04%	31.91%	68.09%	
22	8.40	1.20	0.1942	4.36%	88.41%	11.59%	38.41%	23.19%	76.81%	
23	8.80	1.40	0.1497	3.44%	91.85%	8.15%	41.85%	16.31%	83.69%	
24	9.20	1.60	0.1109	2.61%	94.45%	5.55%	44.45%	11.10%	88.90%	
25	9.60	1.80	0.0790	1.90%	96.35%	3.65%	46.35%	7.30%	92.70%	
26	10.00	2.00	0.0540	1.33%	97.68%	2.32%	47.68%	4.64%	95.36%	
27	10.40	2.20	0.0355	0.89%	98.57%	1.43%	48.57%	2.85%	97.15%	
28	10.80	2.40	0.0224	0.58%	99.15%	0.85%	49.15%	1.69%	98.31%	
29	11.20	2.60	0.0136	0.36%	99.51%	0.49%	49.51%	0.97%	99.03%	
30	11.60	2.80	0.0079	0.21%	99.73%	0.27%	49.73%	0.54%	99.46%	
31	12.00	3.00	0.0044	0.12%	99.85%	0.15%	49.85%	0.30%	99.70%	

Figure 10.12. Spreadsheet for finding areas under the normal curve.

Question 2:

Let us assume that the highest rating awarded to any manufacturing facility is 12. Based on our previous ratings, what is the probability that our chips will be rated above 10?

We look in Column F which lists the probabilities that a random sample exceeds a certain value.

Chapter 10 Common Probability Distributions and Tests of Significance

F16 indicates that there is only a 2.32% probability that the chips will be rated above 10.

Question 3:

Suppose our policy dictates that manufactured chips with a rating of 2 or less must be rejected. What is the probability that we find chips which will have this rating.

We look in Column E which lists the probabilities that a random sample is less than a certain value. E6 indicates that there is only a 2.30% probability that the chips will be rated 2 or below.

Question 4:

Suppose our rating system includes decimal values between integers. What is the probability that a shipment contains chips which have a rating between 4.8 and 7.2?

We look in Column E which lists the probabilities that a random sample is less than a certain value. Here, we subtract the left side area for the lower value from the left side area for the higher value. That is, we subtract $4.8 \rightarrow 27.48\%$ from $7.2 \rightarrow 72.50\%$. Thus, $72.50\% - 27.48\% = 45.02\%$. Therefore, the probability that this shipment consists only of chips rated between 4.8 and 7.2 is 45.02%.

Question 5:

What is the probability that the rating is more than 4 points above or below the mean?

Column H lists the probabilities for both ends. For a rating of less than 2 ($6 - 4$), H6 gives 4.60% probability. Likewise, for a rating of more than 10 ($6 + 4$), H26 gives 4.64% probability. These probabilities must be equal to each other since the standardized normal distribution is symmetrical about the mean.

Question 6:

What is the probability that the ratings are within 4 points of the mean?

We look in Column I which lists the probabilities between ends. The answer can be found from either I6 or I26. We observe that I6 indicates a probability of 95.40% and I26 shows 95.36%. Of course, we could have obtained the answer by subtracting the result of Question 5 from unity (100%), that is, $100\% - 4.60\% = 95.40\%$.

Question 7:

What answer should we give to a customer of ours, if he wants to know what the probability is that the chips he buys have ratings from 6 to 8?

We look in Column G which lists the probabilities between the mean and a given value. G21 yields a probability of 34.06% and this is the answer to our customer's question.

Instructions for constructing the table of Figure 10.12.

1. Begin with a blank spreadsheet and in rows **A1:J6** of the spreadsheet type the information shown. List the data 5, 4, 7, 8, 3, 9, 6, 3, 8, and 7 of the known observations in range **N1:N10** of the spreadsheet.

In **B2**, type the formula `=ROUND(AVERAGE(N1:N10),0)`

and in **B3**, the formula `=ROUND(STDEV(N1:N10),0)`

These formulas compute and round, to the nearest integer, the average and standard deviation of the data in range **N1: N10**.

2. Using the autofill feature, we enter the numbers 1 through 31 in **A7:A37** of the spreadsheet. We will use these as references to indicate the appropriate rows of our table.
3. In Column **B** of the table (Column **C** of the spreadsheet), we enter the Z values starting with -3.00 up to $+3.00$ in 0.2 increments. Again, we use the autofill feature to enter these values. In Row 1, Column **A** of the table (cell **B7** of the spreadsheet), type the formula `=C7*B3+B2`. This formula represents the equation

$$X = Z \times \text{std} \times \text{mean} = Z\sigma + \mu \quad (10.41)$$

This is the same equation as (10.37), Page 10–13, that is,

$$Z = \frac{X - \mu}{\sigma_x} \quad (10.42)$$

Therefore, if the observations data in Column **L** of the spreadsheet change, the mean and standard deviation will change accordingly, but it will not be necessary to change our table. We copy **B7** to **B8:B37**.

4. In Column **C** of the table (cell **D7** of the spreadsheet) we enter the formula `=EXP(C7^2*(-1/2))/SQRT(2*PI())`

This represents the height of the curve; this is, $f(z)$ in (10.38), Page 10–13, that is,

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad (10.43)$$

5. The normalized cdf of the normal distribution is, from (10.39), Page 10–14, that is,

$$P_Z(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-u^2/2} du \quad (10.44)$$

This gives us the area of the curve from $-\infty$ to some point z . Obviously, the area from $-\infty$ to

- $+\infty$ is 1. Now, we recall that we have divided the range $-3 \leq z \leq +3$ into 0.2 increments (Z values column.) We want our starting point to be at $z = -3$ and therefore, we need to find the area under the curve for the interval $-\infty \leq z \leq -3$. This can be found by integrating (10.44). That is, in the upper limit of integration, we could let $z = -3$. However, this is a very difficult integration. Instead, we refer to math tables, such as *CRC Standard Mathematical Tables*, where we find that this area is 0.0013 and this establishes our first value for Column D of the table (cell E7 of the spreadsheet). This value is entered as a percentage, that is, 0.13%. The second value in Column D (cell E8 of the spreadsheet), is found by taking the average of the first two values in Column C (cells D7 and D8 of the spreadsheet), so that this value will be approximately equal to the value obtained from math tables. Accordingly, we enter the second value in Column D as `=AVERAGE(D7:D8)*0.2` and this represents the area of the bar for the $0.2 \times z$ units width. Then, we copy E8 to E9:E37.
6. We want Column E of the table (F on the spreadsheet) to represent the sum of the areas of Column D of the table from a given point. The first value is the same as in Column D of the table, that is, 0.13%. Therefore, we copy 0.13% from cell E7 to F7. The next value in Column E of the table (F on the spreadsheet) must be the sum of the areas from -3 to -2.8 so we type `=E8+F7` in cell F8. Then, we copy F8 to F9:F37.
 7. Column F of the table (G on the spreadsheet) is used to show the total area from any point z to $+\infty$. This area is obviously 1 minus the area indicated by Column E of the table (F on the spreadsheet). Therefore, in cell G7, we type `=1-F7` and we copy this to G8:G37.
 8. Column G of the table (H on the spreadsheet) is used to show the area between a point z and the mean. Since the mean divides the curve into two equal areas, we need be concerned with only half the area under the curve. Therefore, from the leftmost tail of the curve, Row 1 of the table (cell H7 of the spreadsheet) to the middle of the curve, Row 16 of the table (cell H22 of the spreadsheet), we subtract the left side values of Column E of the table (F of the spreadsheet) from 0.5. Thus, in cell H7 we type `=0.5-F7` and we copy this to cells H8:H22. For the remaining cells, we subtract the right side values, Column E of the table (F on the spreadsheet). Therefore, in cell H23 we type `=0.5-G23` and we copy this to cells H23:H37.
 9. Column H of the table (I on the spreadsheet) shows the area at both ends. It is found by subtracting twice the area between point z and the mean from 1. Therefore, in cell I7 we type `=1-2*H7` and we copy it to cells I8:I37.
 10. Column I of the table (J on the spreadsheet) shows the area which remains between the ends after we subtract both left and right ends. Therefore, in cell J7 we type `=1-I7` and we copy it to cells J8:J37.

We can use the MATLAB* script below to plot the standardized normal distribution.

```
% This file produces a plot of the standard normal distribution;
z=linspace(-3,3); fz=1./sqrt(2*pi).*exp(-z.^2/2); plot(z,fz); grid;
title('Standard Normal Distribution'); xlabel('Independent Variable Z');
ylabel('Dependent Variable f(Z');
```

Example 10.11

Prove that the Gaussian (normal) pdf is truly a probability density function.†

Proof:

We must prove that:

- a. the pdf is non-negative, that is, $f_X(x) \geq 0$
- b. the area under the curve is equal to 1, that is, $\int_{-\infty}^{\infty} f_X(x)dx = 1$

The proof that $f_X(x) \geq 0$ can be seen from the bell curve of Figure 10.11, Page 10–14.

To prove that the area under the curve is equal to unity, we begin with

$$\int_{-\infty}^{\infty} f_X(x)dx = \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma_X^2}(x-\mu)^2} dx \tag{10.45}$$

and we let

$$\frac{x-\mu}{\sigma_X} = t \tag{10.46}$$

Then,

$$dx = \sigma_X dt \tag{10.47}$$

and by substitution of (10.46) and (10.47) into (10.45), we obtain

$$\int_{-\infty}^{\infty} f_X(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2} dt \tag{10.48}$$

Next, we let

* MATLAB (MATrix LABoratory) is a very powerful tool for making mathematical calculations, from basic to advanced computations. It also produces impressive plots. It is introduced in Appendix B.

† This proof is lengthy and requires knowledge of advanced mathematics. It may be skipped without loss of continuity.

$$A = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2} dt$$

then,

$$A^2 = AA = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy$$

or

$$A^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x+y)^2} dx dy \tag{10.49}$$

Now, we convert (10.49) to polar coordinates by letting $x = r\cos\theta$, $y = r\sin\theta$, and $dx dy = dA = r dr d\theta$. We note that as x and y vary from $-\infty$ to $+\infty$, r varies from 0 to ∞ whereas θ varies from 0 to 2π . Therefore,

$$A^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} e^{-(1/2)r^2} r dr d\theta = \frac{1}{2\pi} \int_0^{2\pi} d\theta \int_0^{\infty} e^{-(1/2)r^2} r dr \tag{10.50}$$

or

$$A^2 = \frac{2\pi}{2\pi} \int_0^{\infty} e^{-(1/2)r^2} r dr \tag{10.51}$$

Letting $r^2/2 = \tau$, it follows that $d\tau = r dr$ and thus

$$A^2 = \int_0^{\infty} e^{-\tau} d\tau = -e^{-\tau} \Big|_0^{\infty} = 0 - (-1) = 1$$

or

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

Example 10.12

For the normal pdf, i.e.,

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-(x-\mu)^2/(2\sigma_X^2)} \tag{10.52}$$

prove that $E[X] = \mu^*$ and $\text{Var}[X] = \sigma_X^2$.

* This proof also requires knowledge of advanced mathematics. Thus, it may be skipped.

Proof:

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma_X} e^{-(x-\mu)^2/(2\sigma_X^2)} dx \quad (10.53)$$

Letting

$$\frac{x-\mu}{\sigma_X} = t \quad (10.54)$$

it follows that

$$x = \sigma_X t + \mu \quad (10.55)$$

and

$$dx = \sigma_X dt \quad (10.56)$$

Substitution of (10.55) and (10.56) into (10.53) yields

$$E[X] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma_X t + \mu) e^{-t^2/2} dt$$

or

$$E[X] = \frac{\sigma_X}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{-t^2/2} dt + \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} dt \quad (10.57)$$

From the plot of Figure 10.13 below, we observe that

$$t e^{-t^2/2} \quad (10.58)$$

is an odd function, i.e., $-f(-t) = f(t)$.

$$u = f(x) = t e^{-(t^2/2)}$$

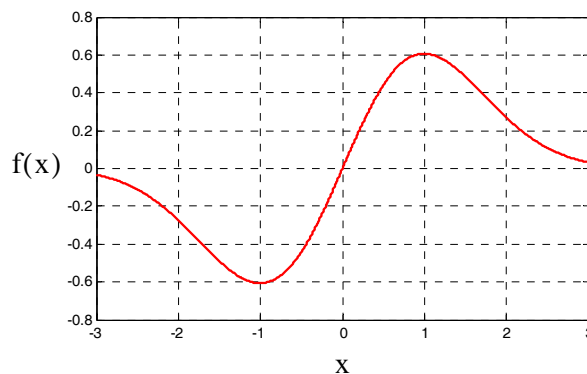


Figure 10.13. Plot of the function $u = f(t) = t e^{-t^2/2}$

Chapter 10 Common Probability Distributions and Tests of Significance

Then, the first integral of (10.57) is zero* whereas the second is $\mu \times 1 = \mu$ since it is the integral of the pdf of the normal distribution of (10.52) with $\mu = 0$ and $\sigma_X = 1$. Therefore, $E(X) = \mu$.

Next, for the variance, we first compute the mean square value of X , $E[X^2]$, that is,

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^{\infty} x^2 e^{-(x-\mu)^2/(2\sigma_X^2)} dx \quad (10.59)$$

As before, we let

$$\frac{x-\mu}{\sigma_X} = t$$

Then,

$$x = \sigma_X t + \mu$$

and

$$dx = \sigma_X dt$$

By substitution into (10.59), we obtain

$$E[X^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma_X t + \mu)^2 e^{-t^2/2} dt$$

or

$$E[X^2] = \frac{\sigma_X^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2/2} dt + \frac{2\mu\sigma_X}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{-t^2/2} dt + \frac{\mu^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} dt \quad (10.60)$$

and as we found in part (a), the second integral of the above expression is zero and the third is $\mu^2 \times 1 = \mu^2$. Therefore, (10.60) is reduced to

$$E[X^2] = \frac{\sigma_X^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2/2} dt + \mu^2$$

Next, we compute the variance $\text{Var}[X]$ from

$$\text{Var}[X] = E[X^2] - \mu^2 = \frac{\sigma_X^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2/2} dt + \mu^2 - \mu^2 = \sigma_X^2 \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2/2} dt \right)$$

The proof will be complete if we show that the integral

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2/2} dt \quad (10.61)$$

is equal to unity.

* Odd functions have the property that the net area from $-\infty$ to $+\infty$ or, in general, from $-x$ to $+x$ is zero.

We will integrate (10.61) by parts* recalling that

$$\int u dv = uv - \int v du \quad (10.62)$$

Here, we let

$$u = t \quad (10.63)$$

and

$$dv = te^{-t^2/2} dt \quad (10.64)$$

Differentiating (10.63) and integrating (10.64) we obtain

$$du = dt$$

and

$$v = -e^{-t^2/2}$$

We also recall that if

$$y = e^{x^2}$$

then

$$\frac{dy}{dx} = e^{x^2} \frac{d}{dx}(x^2) = 2xe^{x^2} \quad (10.65)$$

Therefore, using (10.62) we obtain

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \left(-te^{-\frac{t^2}{2}} \right) \Bigg|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} dt = 1$$

since the first term on the right side is zero and the second is unity.

Related to the cfd $P_X(x)$ of the normal distribution are the error function[†] $\text{erf}(u)$ defined as

$$\text{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-\tau^2} d\tau \quad (10.66)$$

where τ is a dummy variable.

The complementary error function $\text{erfc}(u)$ is defined as

* This is a method of integration, where the given integral is separated into two parts to simplify the integration. It is discussed in calculus textbooks.

† The error function is also known as the “probability integral”.

$$\operatorname{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^{\infty} e^{-\tau^2} d\tau = 1 - \operatorname{erf}(u) \quad (10.67)$$

The error function can also be expressed in series form as

$$\operatorname{erf}(u) = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n u^{2n+1}}{n!(2n+1)} \quad (10.68)$$

The error function is an odd function, that is,

$$-\operatorname{erf}(-u) = \operatorname{erf}(u) \quad (10.69)$$

The cdf $P_X(x)$ of the normal distribution is often expressed in terms of the error function. Their relation is derived as follows:

We know that

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

and that the pdf of the normal distribution is symmetrical about the vertical axis. Therefore,

$$\int_{-\infty}^0 f_X(x) dx = \frac{1}{2}$$

From (10.36), Page 10–13,

$$P_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^x e^{-(x-\mu)^2/(2\sigma_X^2)} dx = \frac{1}{2} + \frac{1}{\sqrt{2\pi}\sigma_X} \int_0^x e^{-(x-\mu)^2/(2\sigma_X^2)} dx \quad (10.70)$$

Next, we make a variable change in (10.70), that is, we let

$$\tau = \frac{x-\mu}{\sqrt{2}\sigma_X} \quad (10.71)$$

then,

$$d\tau = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sigma_X} dx \quad (10.72)$$

By substitution of (10.71) and (10.72) into (10.66), Page 10–25, we obtain:

$$\operatorname{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma_X}\right) = \frac{2}{\sqrt{2\pi}\sigma_X} \int_0^x e^{-\left(\frac{x-\mu}{\sqrt{2}\sigma_X}\right)^2} dx \quad (10.73)$$

or

$$\frac{1}{2} \operatorname{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma_X}\right) = \frac{1}{\sqrt{2\pi}\sigma_X} \int_0^x e^{-\left(\frac{x-\mu}{\sqrt{2}\sigma_X}\right)^2} dx \quad (10.74)$$

Finally, substitution of (10.74) into (10.70), Page 10–26, yields

$$P_X(x) = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma_X}\right) \right] \quad (10.75)$$

Example 10.13

Determine the probability that the Gaussian rv X lies in the interval

$$\mu - k\sigma_X < X \leq \mu + k\sigma_X$$

where k is a non-zero positive integer.

Solution:

By definition,

$$P_X(x) = P(X \leq x)$$

and from (9.7), Chapter 9, Page 9–5,

$$P_X(x_1 < X \leq x_2) = P_X(x_2) - P_X(x_1) \quad (10.76)$$

Then,

$$P_X(\mu - k\sigma_X < X \leq \mu + k\sigma_X) = P_X(\mu + k\sigma_X) - P_X(\mu - k\sigma_X)$$

and using (10.75) above, we obtain

$$P_X(\mu + k\sigma_X) = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{\mu + k\sigma_X - \mu}{\sqrt{2}\sigma_X}\right) \right] = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{k}{\sqrt{2}}\right) \right]$$

Similarly,

$$P_X(\mu - k\sigma_X) = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{\mu - k\sigma_X - \mu}{\sqrt{2}\sigma_X}\right) \right] = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{-k}{\sqrt{2}}\right) \right]$$

Therefore, with (10.76) above,

$$P_X(\mu - k\sigma_X < X \leq \mu + k\sigma_X) = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{k}{\sqrt{2}}\right) \right] - \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{-k}{\sqrt{2}}\right) \right]$$

or

$$P_X(\mu - k\sigma_X < X \leq \mu + k\sigma_X) = \frac{1}{2} \left[\operatorname{erf}\left(\frac{k}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{-k}{\sqrt{2}}\right) \right]$$

Chapter 10 Common Probability Distributions and Tests of Significance

Since the error function is an odd function, that is, $-\text{erf}(-u) = \text{erf}(u)$, the above relation reduces to

$$P_X (\mu - k\sigma_X < X \leq \mu + k\sigma_X) = \text{erf}\left(\frac{k}{\sqrt{2}}\right) \quad (10.77)$$

The error function is tabulated in math tables such as *Standard Mathematical Tables* by CRC Press and *Handbook of Mathematical Functions* by Dover Publications. It can also be computed with the MATLAB **erf(1/sqrt(2))** function.

When $k = 1$, (10.77) above yields

$$\text{erf}\left(\frac{1}{\sqrt{2}}\right) = \text{erf}(0.707) = 0.6827 \quad (10.78)$$

Likewise, when $k = 2$,

$$\text{erf}\left(\frac{2}{\sqrt{2}}\right) = \text{erf}(1.414) = 0.9545 \quad (10.79)$$

and when $k = 3$,

$$\text{erf}\left(\frac{3}{\sqrt{2}}\right) = \text{erf}(2.1213) = 0.9973 \quad (10.80)$$

The values of (10.78), (10.79), and (10.80) are in agreement with the values shown on the normalized Gaussian distribution of Figure 10.11, Page 10–14. In particular, we see that when $k = 3$, the probability that a Gaussian *rv* X lies between $\pm 3\sigma_X$ of its mean μ , is very close to unity.

The error and complementary error functions are also available in Excel. The syntax for the error function is

`=ERF(lower_limit,upper_limit)`

where

`lower_limit` = lower bound for integrating the error function

`upper_limit` = upper bound for integrating the error function. If omitted, the integration is between zero and the lower limit.

Excel uses the formulas

$$\text{ERF}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt \quad (10.81)$$

and

$$\text{ERF}(a, b) = \frac{2}{\sqrt{\pi}} \int_a^b e^{-t^2} dt = \text{ERF}(b) - \text{ERF}(a) \quad (10.82)$$

For example,

=ERF(0.75) returns 0.711

=ERF(0.5, 1) returns 0.322

The syntax for the complementary error function in Excel is

=ERFC(x)

where

x = lower bound for integrating the *error* function, not the complementary error function.

Excel uses the formula

$$\text{ERFC}(z) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt = 1 - \text{ERF}(x) \quad (10.83)$$

For example, =ERFC(0.25) returns 0.724

Example 10.14

A satellite communications system sends out two possible messages to a ground station; let these be represented by 0 and 5 volt signals. However, the transmitted messages are corrupted by additive atmospheric noise. The ground receiver station has been designed so that any signal equal to or below 2.5 volts, is interpreted as 0 volts, while any signal above 2.5 volts is interpreted as 5 volts. From past experience, it is known that if a 5 volt signal is transmitted, the received signal is random whose Gaussian pdf has mean value $\mu = 3$ and standard deviation $\sigma_x = \sqrt{2}$.

Compute the probability that a 5 volt transmitted message will be interpreted as 0 volts message by the ground receiver station.

Solution:

The Gaussian pdf is

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2\sigma_x^2}(x-\mu)^2}$$

The probability that the rv X is equal to or less than x volts is

$$P[X \leq x] = \int_{-\infty}^x f_X(x) dx = \int_{-\infty}^0 f_X(x) dx + \int_0^x f_X(x) dx$$

or

$$P[X \leq x] = \frac{1}{2} + \frac{1}{\sqrt{2\pi}\sigma_X} \int_0^x e^{-\frac{1}{2\sigma_X^2}(x-1)^2} dx \quad (10.84)$$

and using (10.75), Page 10–27, we express (10.84) as

$$P[X \leq x] = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x - \mu}{\sqrt{2}\sigma_X} \right) \right] \quad (10.85)$$


Next, with $x = 2.5$, $\mu = 3$, and $\sigma_X = \sqrt{2}$, we obtain

$$P[X \leq 2.5] = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{2.5 - 3}{\sqrt{2} \times \sqrt{2}} \right) \right] = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{-0.5}{2} \right) \right]$$

Since $-\operatorname{erf}(-x) = \operatorname{erf}(x)$, it follows that

$$P[X \leq 2.5] = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{0.5}{2} \right) \right] = \frac{1}{2} [1 - 0.2763] = 0.3618$$

Therefore, there is a probability of 36.18% that the 5 volt signals transmitted by the satellite, will be misinterpreted by the ground receiver station.



A third useful function, which is related to the standard Gaussian cdf, is the $Q(\alpha)$ function defined as

$$Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{-\gamma^2/2} d\gamma \quad (10.86)$$

that is, the $Q(\alpha)$ function gives the probability of the right tail of the normal distribution. Recalling that

$$P_Z(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-u^2/2} du \quad (10.87)$$

and comparing (10.86) with (10.87), we see that

$$Q(z) = 1 - P_Z(z) \quad (10.88)$$

It can also be shown that

$$Q(\alpha) = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{\alpha}{\sqrt{2}} \right) \right] \quad (10.89)$$

The $Q(\alpha)$ function is used in electronic communications theory to define error probability bounds.

The MATLAB script below, tabulates several values of α and the corresponding values of $Q(\alpha)$.

```
alpha=0: 0.25: 6; qalpha=zeros(25,2); qalpha(:,1)=alpha';  
qalpha(:,2)=0.5 .* (1-erf(alpha ./ sqrt(2)));  
fprintf('alpha    qalpha \n'); fprintf('%3.2f \t %E \n', qalpha);
```

alpha	qalpha
0.00	5.000000E-001
0.25	4.012937E-001
0.50	3.085375E-001
0.75	2.266274E-001
1.00	1.586553E-001
1.25	1.056498E-001
1.50	6.680720E-002
1.75	4.005916E-002
2.00	2.275013E-002
2.25	1.222447E-002
2.50	6.209665E-003
2.75	2.979763E-003
3.00	1.349898E-003
3.25	5.770250E-004
3.50	2.326291E-004
3.75	8.841729E-005
4.00	3.167124E-005
4.25	1.068853E-005
4.50	3.397673E-006

4.75	1.017083E-006
5.00	2.866516E-007
5.25	7.604961E-008
5.50	1.898956E-008
5.75	4.462172E-009
6.00	11.865876E-010

Plots of $\text{erf}(x)$ and $Q(x)$ using MATLAB are shown in Figure 10.14.

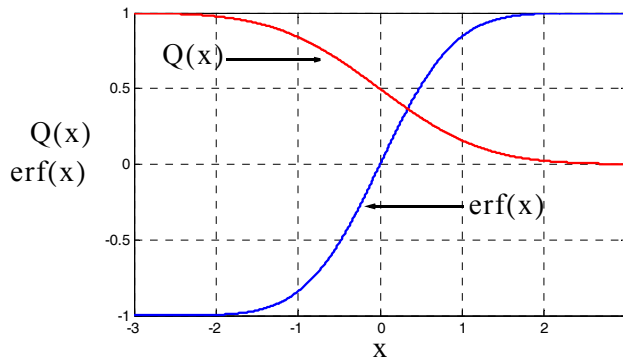


Figure 10.14. The Error and the Q functions

10.6 Percentiles

Percentiles split the data into 100 parts. Let us consider the pdf of Figure 10.15 below.

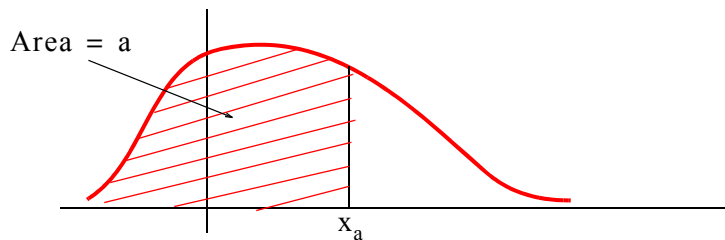


Figure 10.15. Sketch to illustrate the concept of percentile

The area to the left of point x_a is a ; we denote this area as a *percentile* p_a . Typical percentile values are 0.05 or 5%, 0.10 or 10%, 0.95 or 95%, and 0.99 or 99% and these are denoted as $p_{0.05}$, $p_{0.10}$, ..., $p_{0.95}$, or $p_{0.99}$ respectively.

For continuous functions, the j th percentile p_j is defined as

$$p_j = P[X \leq x_p] = \int_{-\infty}^{x_p} p(x) dx = \frac{p}{100} \quad (10.90)$$

Example 10.15

Compute $p_{0.90}$, $p_{0.95}$, and $p_{0.99}$ for the cdf of the exponential distribution of Figure 10.16 if $\mu = 1/2$.

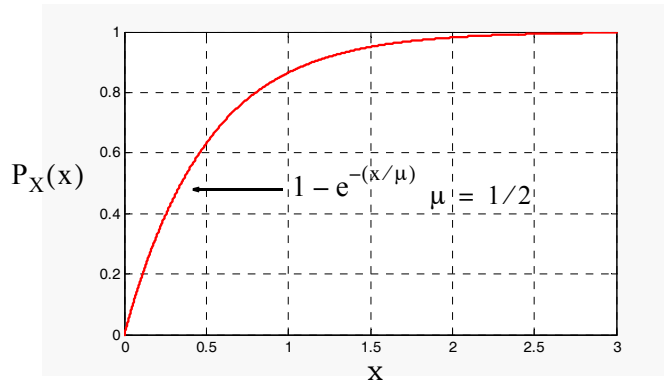


Figure 10.16. Cdf for Example 10.15

Solution:

From (10.90) above,

$$p_j = P[X \leq x_p] = \int_{-\infty}^{x_p} p(x) dx = 1 - e^{-(x/\mu)} = 1 - e^{-2x} = \frac{p}{100}$$

or

$$e^{-2x} = 1 - \frac{p}{100}$$

Taking the natural log of both sides of the above expression, we obtain

$$-2x = \ln\left(1 - \frac{p}{100}\right)$$

or

$$x = p_j = -\frac{1}{2} \ln\left(\frac{100-p}{100}\right) = \frac{1}{2} \ln\left(\frac{100}{100-p}\right)$$

Then,

$$p_{0.90} = \frac{1}{2} \ln\left(\frac{100}{100-90}\right) = \frac{2.3}{2} = 1.15$$

$$p_{0.95} = \frac{1}{2} \ln\left(\frac{100}{100-95}\right) = \frac{3}{2} = 1.50$$

$$p_{0.99} = \frac{1}{2} \ln\left(\frac{100}{100-99}\right) = \frac{4.6}{2} = 2.3$$

For a pdf of the discrete type, we can obtain *approximate values* of percentiles, if we first arrange the sample in ascending order of magnitude; then, we compute the j th percentile p_j , $j = 1, 2, \dots, 99$ from the relation

$$i = \frac{jn}{100} \tag{10.91}$$

where n = number of samples .

If i in (10.91) above turns out to be an integer, the j th percentile is the average of the observations in positions i and $i + 1$ in the sampled data set starting at the bottom (lowest value) of the distribution. If i is not an integer, we select the next integer which is greater than i .

Example 10.16

Given the test scores 87, 85, 90, 80, 82, 92, 79, 85, 95, 86, find the score at the 90th percentile.

Solution:

The first step is to arrange the sample in ascending order. This is done in Table 10.1.

TABLE 10.1 Table for Example 10.16

i position	1	2	3	4	5	6	7	8	9	10
Scores	79	80	82	85	85	86	87	90	92	95

Next, using (10.91), with $j = 90$ and $n = 10$, we obtain

$$i = \frac{jn}{100} = \frac{90 \times 10}{100} = 9$$

and since $i = 9$ is an integer, we compute the 90th percentile by taking the average of the values in the 9th and 10th position, that is,

$$\frac{92 + 95}{2} = 93.5 \tag{10.92}$$

Therefore, the score at the 90th percentile is 93.5 .

We can also use Excel's **PERCENTILE** function which returns the j th percentile values in a range. The syntax is

=PERCENTILE(array,j)

where

array = range of data that defines the relative standing

j = percentile value in the range $0 \leq j \leq 1$.

For the above example,

=PERCENTILE({87,85,90,80,82,92,79,85,95,86},0.9) returns 92.3 .

This differs somewhat from the value of 93.5 which we found in (10.92). This is due to the fact that this formula of (10.91) is an approximation, and Excel also returns an approximation.

We observe that when using Excel's **PERCENTILE** function, we need not arrange the sample in ascending order of magnitude before we compute the j th percentile.

Example 10.17

The sampled data below, represent the averages for the same test given at 10 different schools.

87.3, 85.8, 90.7, 80.2, 82.4, 92.6, 71.1, 85.9, 95.1, 86.7

Find the average score at the 95th percentile.

Solution:

The first step is to arrange the sample in ascending order. This is done in Table 10.2.

TABLE 10.2 Table for Example 10.17

i position	1	2	3	4	5	6	7	8	9	10
Scores	71.1	80.2	82.4	85.8	85.9	86.7	87.3	90.7	92.6	95.1

Next, using (10.91), Page 10–34, with $j = 95$ and $n = 10$, we obtain

$$i = \frac{jn}{100} = \frac{95 \times 10}{100} = 9.5$$

and since $i = 9.5$ is not an integer, we compute the 95th percentile by taking the next integer greater than 9.5 which is 10 . From Table 10.2 we find that the average score at the 95% percentile is 95.1 .

With the Excel function

=PERCENTILE({87.3,85.8,90.7,80.2,82.4,92.6,711.1,85.9,95.1,86.7},0.95)

we obtain the value 93.98.

10.7 The Student's t -Distribution

Approximately 95% of the area under the standard normal curve is between $z = -1.96$ and $z = +1.96$. This interval is called *95% confidence interval* for the population mean μ . Here, the word *population* refers to a particular situation where our estimates are based on, and the word *sample* implies a small selection from that population. For instance, population may refer to a group of students passing a certain examination, and a sample may be one or more students from this population.

When the sample size is less than 30, the *Student's t -distribution*^{*} or simply *t -distribution* is used instead of the standard normal distribution. The t -distribution is actually a family of probability distributions, and each separate t -distribution is determined by a parameter called the *degrees of freedom*, which we will denote as df . If n is the sample size for $n < 30$, then the degrees of freedom can be computed from

$$df = n - 1 \quad (10.93)$$

The t -distribution curves, like the standard normal distribution curve, are centered at zero. Thus, the mean is zero, that is,

$$\mu = 0 \quad (10.94)$$

The variance of each t -distribution depends upon the value of df . It is found from

$$\sigma^2 = \frac{df}{df - 2} \quad df \geq 3 \quad (10.95)$$

Figure 10.17 below shows the t -distribution for $df = 1$, $df = 3$, and $df = 20$ where we observe that, for higher values of df , this distribution approximates the normal distribution.

The pdf of the t -distribution is given in terms of the Γ (gamma) function. This function is discussed in Appendix C. Thus, for the t -distribution,

* The name "student" has nothing to do with a student. It is a pen name for William Gossett who published his work under that name.

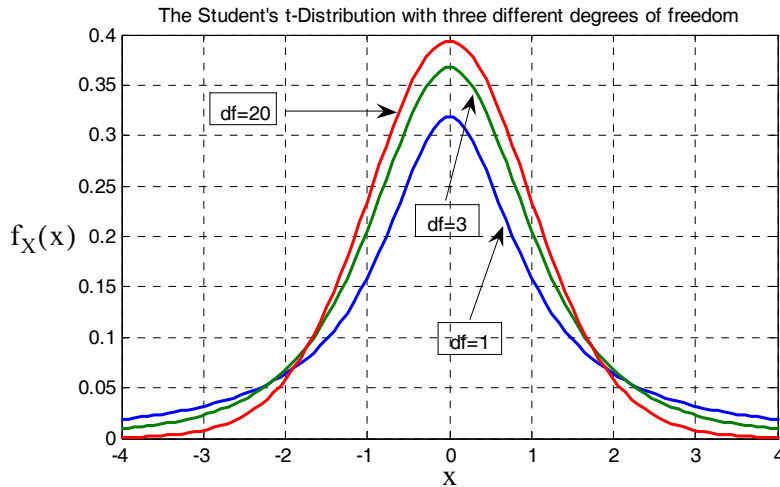


Figure 10.17. The Student's t-distribution for different degrees of freedom

$$f_X(x) = \frac{\Gamma\left(\frac{df+1}{2}\right)}{\sqrt{df \cdot \pi} \Gamma\left(\frac{df}{2}\right)} \left(1 + \frac{x^2}{df}\right)^{-\frac{(df+1)}{2}} \quad (10.96)$$

for $-\infty < x < +\infty$

The MATLAB script to plot the pdf of the t-distribution is listed below.

```
% MATLAB Program for Student's t-Distribution;
t=linspace(-4,4);
stud_t1=gamma((1+1)/2)./(sqrt(1.*pi).*gamma(1./2)).*(1+t.^2./1).^(-(1+1)/2);
stud_t3=gamma((3+1)/2)./(sqrt(3.*pi).*gamma(3./2)).*(1+t.^2./3).^(-(3+1)/2);
stud_t20=gamma((20+1)/2)./(sqrt(20.*pi).*gamma(20./2)).*(1+t.^2./20).^(-(20+1)/2);
plot(t,stud_t1, t,stud_t3, t,stud_t20); grid
title('The Student's t-Distribution with three different degrees of freedom')
```

Percentile values of the t-distribution for different degrees of freedom are denoted as t_p . These can be found in math tables such as the *CRC Standard Mathematical Tables*. The tables give values of x for the cdf

$$P_X(x) = \int_{-\infty}^x \frac{\Gamma\left(\frac{df+1}{2}\right)}{\sqrt{df \cdot \pi} \Gamma\left(\frac{df}{2}\right)} \left(1 + \frac{x^2}{df}\right)^{-\frac{(df+1)}{2}} dx \quad (10.97)$$

Chapter 10 Common Probability Distributions and Tests of Significance

These tables give values for the number of degrees of freedom equal to 1, 2, ..., 30, 40, 60, 120, ∞ and for $P_X(x) = 0.60, 0.75, 0.90, 0.95, 0.975, 0.99, 0.995$ and 0.9995 .

The t -distribution is symmetrical; therefore,

$$t_{1-p} = -t_p \quad (10.98)$$

For instance,

$$t_{0.10} = -t_{0.90}$$

In Figure 10.18, the shaded area represents the area p from $-\infty$ to t_p . Therefore, the area to the right of t_p is $1 - t_p$.

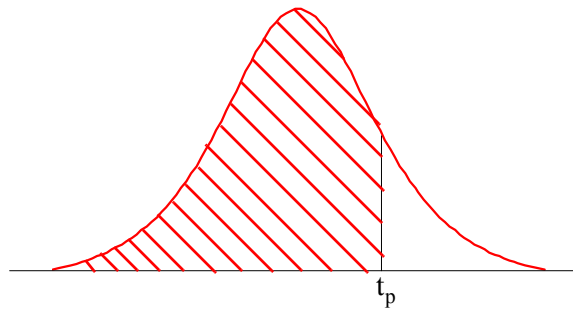


Figure 10.18. Percentile values t_p for Student's t -distribution.

We will illustrate the use of the t -distribution with the following example.

Example 10.18

Use the percentile values t_p for the t -distribution with $df = 7$ to find:

- the area under the curve to the right of $t_p = 2.998$.
- the area under the curve to the left of $t_p = -2.998$
- the area under the curve between $t_p = -2.998$ and $t_p = 2.998$.

Solution:

- From a Student's t -distribution table we find that

$$\text{Area}_{df=7}^{\text{left of } t_p = 2.998} = 0.99$$

Therefore, the area to the right of $t_p = 2.998$ is $1 - 0.99 = 0.01$

b. Since the t -distribution is symmetrical,

$$\text{Area}_{df=7}^{\text{left of } t_p = -2.998} = \text{Area}_{df=7}^{\text{right of } t_p = 2.998} = 0.99$$

Therefore, the area to the left of $t_p = -2.998$ is the same as in part (a), that is, $1 - 0.99 = 0.01$

c. Since the total area under the curve is unity, the area under the curve between $t_p = -2.998$ and $t_p = 2.998$ is $1 - 0.01 - 0.01 = 0.98$

The *one-tail test* and *two-tail test* are discussed in Section 10.21, Page 10–65. Briefly, we choose the one-tail test if we are interested on some extreme values on one side of the mean, that is, in one “tail” of the distribution. But if we are interested on some extreme values on both sides of the mean, that is, both “tails” of the distribution, we choose a two-tail test.

Microsoft Excel has the build-in function **TDIST** which returns the probability associated with the Student's t -distribution. Its syntax is

=TDIST(x,degrees_freedom,tails)

where

x = the numeric value at which to evaluate the distribution.

degrees_freedom = an integer indicating the number of degrees of freedom.

tails = number of distribution tails to return. If **tails** = 1, **TDIST** returns the one-tail distribution. If **tails** = 2, **TDIST** returns the two-tail distribution.

In Example 10.18, we were given $t_p = x = 2.998$ and $df = 7$. Using Excel, for **tails** = 1, we enter the formula **=TDIST(2.998,7,1)** and Excel returns **0.01**. For **tails** = 2, we enter the formula **=TDIST(2.998,7,2)** and Excel returns **0.02**. These are the same answers we found earlier.

Example 10.19

For a t -distribution with $df = 7$,

- find the t_p value for which the area under the curve to the right of t_p is **0.05**.
- find the t_p value for which the area under the curve to the left of t_p is **0.05**.
- using the results of (a) and (b), find the absolute value $|t_p|$ for which the area between $-t_p$ and $+t_p$ is **0.11**.

Solution:

a. Since the t -distribution is symmetrical,

$$t_{1-p} = -t_p$$

and thus

$$t_{0.05 \text{ (right)}} = -t_{0.95 \text{ (left)}}$$

From a t -distribution table,

$$t_p \Big|_{df=7}^{\text{Area} = 0.05 \text{ to the right of } t_p} = t_p \Big|_{df=7}^{\text{Area} = 0.95 \text{ to the left of } t_p} = 1.895$$

b.

$$t_p \Big|_{df=7}^{\text{Area} = 0.05 \text{ to the left of } t_p} = t_p \Big|_{df=7}^{0.95 \text{ to the right of } t_p} = -1.895$$

c. From (a) and (b),

Area left of $t_p = -1.895$ is **0.05** and Area right of $t_p = +1.895$ is also **0.05** yielding a total area of **0.10**. Then, the area between these values of t_p is $1 - 0.10 = 0.90$. Therefore, the absolute value $|t_p|$ for which the area between $-t_p$ and $+t_p$ is **0.90** is $|t_p| = 1.895$.

Excel has the build-in function **TINV** which returns the inverse of the Student's t -distribution. Its syntax is

=TINV(probability,degrees_freedom)

where

probability = probability associated with the two-tail t -distribution.

degrees_freedom = an integer indicating the number of degrees of freedom.

We can verify the answer for part (c) of Example 10.19 using Excel's formula **=TINV(0.10,7)** and Excel returns **1.895**. This is the same answer as that we found earlier.

10.8 The Chi-Square Distribution

The χ^2 (chi-square) distribution, like the t -distribution, is a family of probability distributions, and each separate distribution is determined by the *degrees of freedom*, which we will denote as **df**.

The pdf of the χ^2 distribution is given in terms of the Γ (gamma) function which is discussed in Appendix C. For the χ^2 distribution,

$$f_X(x) = \begin{cases} \frac{1}{2^{df/2}\Gamma(df/2)} x^{(df/2-1)} e^{-x/2} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (10.99)$$

Figure 10.19 shows the χ^2 distribution for $df = 5$, $df = 10$, and $df = 15$.

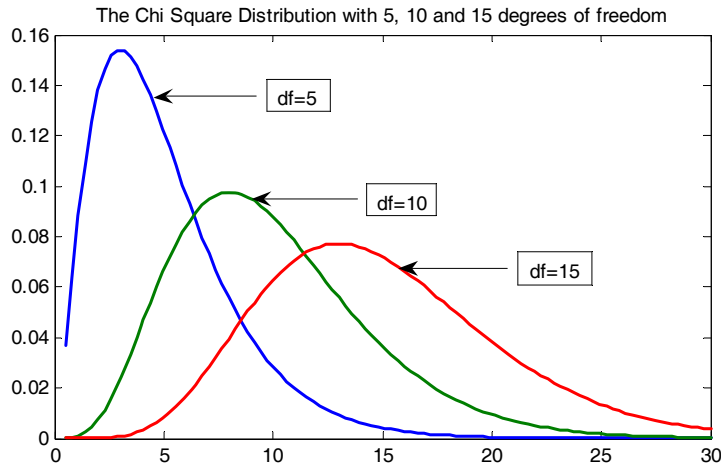


Figure 10.19. The χ^2 distribution for different degrees of freedom.

The MATLAB script to plot the pdf of the χ^2 distribution is listed below.

```
x=linspace(0.5,30); chi_sq5=1./(2.^(5./2).*gamma(5./2)).*x.^((5./2)-1).*exp(-x./2);
chi_sq10=1./(2.^(10./2).*gamma(10./2)).*x.^((10./2)-1).*exp(-x./2);
chi_sq15=1./(2.^(15./2).*gamma(15./2)).*x.^((15./2)-1).*exp(-x./2);
plot(x,chi_sq5,x,chi_sq10,x,chi_sq15);
title('The Chi Square Distribution with 5, 10 and 15 degrees of freedom')
```

The mean of the χ^2 distribution is

$$\mu = df \quad (10.100)$$

and the variance is

$$\sigma^2 = 2df \quad (10.101)$$

Percentile values of the χ^2 distribution for different degrees of freedom are denoted as χ_p^2 and can be found in math tables such as the *CRC Standard Mathematical Tables*. The tables give values of x for the cdf

$$P_X(x) = \int_{-\infty}^x \frac{1}{2^{df/2}\Gamma(df/2)} x^{(df/2-1)} e^{-x/2} dx \quad (10.102)$$

Chapter 10 Common Probability Distributions and Tests of Significance

These tables give values for the number of degrees of freedom equal to 1, 2, ..., 30 and for $P_X(x)$ from 0.005 up to 0.995.

Example 10.20

Figure 10.20 shows an χ^2 distribution with $df = 5$.

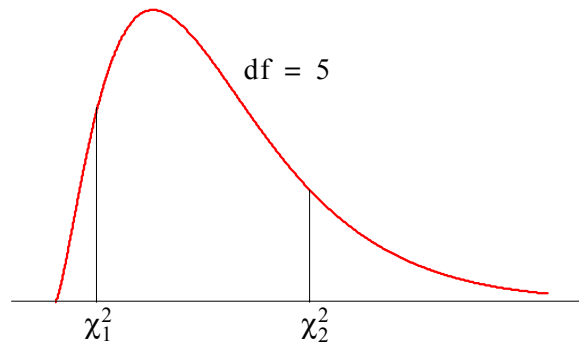


Figure 10.20. χ^2 distribution for Example 10.20

Find the values of χ_1^2 and χ_2^2 for which

- the shaded area on the right of χ_2^2 is equal to 0.05
- the shaded area on the left of χ_1^2 is equal to 0.10
- the shaded area on the right of χ_2^2 is equal to 0.01
- the total of the shaded areas to the left of χ_1^2 and to the right of χ_2^2 is 0.05

Solution:

- Referring to a table of the χ^2 distribution, we first locate the row corresponding to $df = 5$. Now, if the shaded area on the right of χ^2 is to be equal to 0.05, the area to the left must be $1 - 0.05 = 0.95$. Then, on the table, we move to the right until we reach the column headed 0.950. There, we read 11.11 and this is the value of χ_2^2 that will result in the area of 0.05 to the right of it.
- If the shaded area to the left of χ_1^2 is to be equal to 0.10, the value of χ_1^2 will represent the 10th percentile. Then, in the table, again with $df = 5$, we move to the right until we reach the

column headed 0.100. There, we read 1.61 and this is the value of χ_1^2 that will result in the area of 0.10 to the left of it.

- c. If the shaded area on the right of χ_2^2 is to be equal to 0.01, the area to the left must be $1 - 0.01 = 0.99$. Then, on the table, we move to the right until we reach the column headed 0.990. We read 15.1 and this is the value of χ_2^2 that will result in the area of 0.01 to the right.
- d. Since this distribution is not symmetric, we can arbitrarily decide on the area of one side and subtract this area from 0.05 to find the area of the other side. Let us decide on equal areas, that is 0.025 on the left and 0.025 on the right. Therefore, if the area to the right of χ_2^2 is to be 0.025, the area to the left of it will be $1 - 0.025 = 0.975$ and thus χ_2^2 represents the 97.5th percentile. From tables, $\chi_{0.975}^2 = 12.8$. Likewise, if the shaded area on the left is 0.025 and this represents the 2.5th percentile and from tables $\chi_{0.025}^2 = 0.831$. Therefore, this condition will be satisfied if $\chi_1^2 = \chi_{0.975}^2 = 12.8$ and $\chi_2^2 = \chi_{0.025}^2 = 0.831$.

We can use Excel's CHIDIST and CHIINV functions to find values associated with the χ^2 distribution. The first has the syntax

`=CHIDIST(x,degrees_freedom)`

and returns the one-tail probability of this distribution.

For example,

`=CHIDIST(12.8,5)` returns 0.025.

The second returns the inverse of the distribution and has the syntax

`=CHIINV(probability,degrees_freedom)`

For example,

`=CHIINV(0.025,5)` returns 12.8. This is the same value as in part (d) of Example 10.20.

10.9 The F Distribution

The *F distribution* is another family of curves whose shape differs according to the degrees of freedom. The pdf of the F distribution is

$$f_X(x) = \frac{\Gamma\left(\frac{df_1 + df_2}{2}\right)}{\Gamma\left(\frac{df_1}{2}\right)\Gamma\left(\frac{df_2}{2}\right)} df_1^{\left(\frac{df_1}{2}\right)} df_2^{\left(\frac{df_2}{2}\right)} x^{\left(\frac{df_1}{2} - 1\right)} (df_2 + df_1 x)^{-\frac{(df_1 + df_2)}{2}} \quad x > 0$$

$$= 0 \quad x \leq 0$$
(10.103)

These curves differ not only according to the size of the samples, but also on how many samples are being compared. The F distribution has two sets of degrees of freedom, df_1 (degrees of freedom for the numerator) and df_2 (degrees of freedom for the denominator). The designation of an F distribution with 10 degrees of freedom for the numerator, and 15 degrees of freedom for the denominator is $df = (df_1, df_2) = (10, 15)$.

Figure 10.21 shows two F distributions, one with $df(8, 40)$, and the other with $df(8, 4)$.

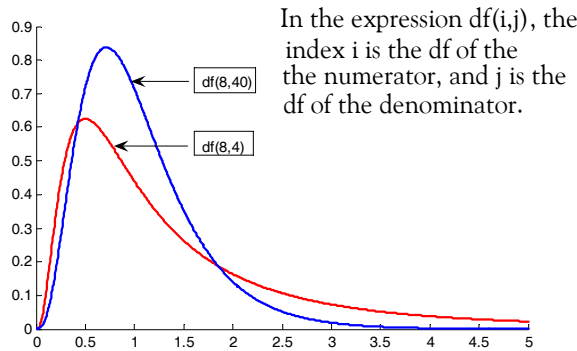


Figure 10.21. The F distribution for different degrees of freedom.

The mean μ of the F distribution is given by

$$\mu = \frac{df_2}{df_2 - 2} \quad df_2 \geq 3$$
(10.104)

Thus, for the F distribution with $df(8, 40)$,

$$\mu_{df(8, 40)} = \frac{df_2}{df_2 - 2} = \frac{40}{40 - 2} = 1.053$$

and for the other with $df(8, 4)$,

$$\mu_{df(8, 4)} = \frac{df_2}{df_2 - 2} = \frac{4}{4 - 2} = 2$$

The variance of the F distribution is given by

$$\sigma^2 = \frac{2df_2(df_1 + df_2 - 2)}{df_1(df_2 - 2)^2(df_2 - 4)} \quad df_2 \geq 5 \quad (10.105)$$

Percentile values of the F distribution for different degrees of freedom are denoted as F_p and can be found in math tables such as the *CRC Standard Mathematical Tables*. The tables give values of x for the cdf

$$P_X(x) = \int_0^x \frac{\Gamma\left(\frac{df_1 + df_2}{2}\right)}{\Gamma\left(\frac{df_1}{2}\right)\Gamma\left(\frac{df_2}{2}\right)} df_1^{\left(\frac{df_1}{2}\right)} df_2^{\left(\frac{df_2}{2}\right)} x^{\left(\frac{df_1}{2} - 1\right)} (df_2 + df_1 x)^{-\frac{(df_1 + df_2)}{2}} dx \quad (10.106)$$

The F distribution tables are based on the df_1 and df_2 degrees of freedom and provide values corresponding to

$$P_X(x) = 0.90, 0.95, 0.975, 0.99, 0.995, \text{ and } 0.999$$

Lower percentile values for the F distribution can be found from the relation

$$F_{1-p, df_1, df_2} = \frac{1}{F_{p, df_2, df_1}} \quad (10.107)$$

For example,

$$F_{0.05, 4, 7} = \frac{1}{F_{0.95, 7, 4}} = \frac{1}{6.09} = 0.164$$

Excel provides the functions **FDIST** and **FINV** to find values associated with the F distribution.

The first has the syntax

=FDIST(x,degrees_freedom1,degrees_freedom2)

and returns the probability of this distribution.

For example,

=FDIST(0.164,4,7) returns **0.95**.

The second returns the inverse of the distribution and has the syntax

=FINV(probability,degrees_freedom1,degrees_freedom2)

Thus, for the above example, **=FINV(0.95,4,7)** returns **0.164**.

The F distribution is used extensively in the *Analysis of Variance (ANOVA)* which compares the means of several populations. ANOVA is discussed in Chapter 12.

10.10 Chebyshev's Inequality

Let X be an rv (continuous or discrete) with mean μ and variance σ^2 . *Chebyshev's inequality* states that

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \quad (10.108)$$

That is, the probability of X differing from its mean by more than some positive number ε , is less than or equal to the ratio of the variance σ^2 to the square of the number ε . This inequality is often expressed as

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \quad (10.109)$$

where $\varepsilon = k\sigma$ and $k > 1$.

If, in (10.109), we let $k = 2$, we obtain

$$P(|X - \mu| \geq 2\sigma) \leq \frac{1}{2^2} = \frac{1}{4} = 0.25$$

and this implies that the probability of X differing from its mean by more than two standard deviations, is less than or equal to 0.25. That is, the probability that X will lie within two standard deviations of its mean is greater or equal to 0.75 since

$$P(|X - \mu| < 2\sigma) \geq 1 - 0.25 = 0.75$$

Likewise, if $k = 3$,

$$P(|X - \mu| \geq 3\sigma) \leq \frac{1}{3^2} = \frac{1}{9}$$

or

$$P(|X - \mu| < 3\sigma) \geq 1 - \frac{1}{9} = \frac{8}{9}$$

that is, at least $8/9$ or 89% of the data will fall between $\mu - 3\sigma$ and $\mu + 3\sigma$.

10.11 Law of Large Numbers

Let

$$X_1, X_2, \dots, X_n$$

be mutually independent random variables each having mean μ and variance σ^2 . If

$$S_n = X_1 + X_2 + \dots + X_n$$

then,

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) = 0 \quad (10.110)$$

and since S_n/n is the arithmetic mean of X_1, X_2, \dots, X_n , the *law of large numbers* states that the probability of the arithmetic mean S_n/n differing from its expected value μ by more than the value of ε approaches zero as $n \rightarrow \infty$.

10.12 The Poisson Distribution

The *Poisson distribution* is a pdf of the discrete type, and it is defined as

$$f(x) = P(X = x) = p(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \quad (10.111)$$

for $x = 0, 1, 2, \dots$ where λ is a given positive number.

This distribution appears in many natural phenomena such as the number of telephone calls per minute received by a receptionist, the number of customers walking into a bank at some particular time, the number of errors per page in a book with many pages, the number of α particles emitted by a radioactive substance, etc.

Besides the applications mentioned above, the Poisson distribution provides us with a close approximation of the binomial distribution for small values of x provided that the probability of success p is small so that the probability of failure q , that is, $q = 1 - p \approx 1$ and

$$\lambda = np \quad (10.112)$$

where $n =$ number of trials.

Note: An event in which $q = 1 - p \approx 1$ is called a *rare event*.

Figure 10.22 shows the graph of a Poisson distribution with $\lambda = 2$.

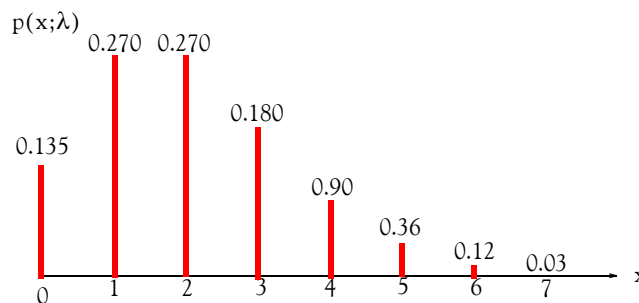


Figure 10.22. Poisson pdf with $\lambda = 2$

Chapter 10 Common Probability Distributions and Tests of Significance

To understand the meaning of the distribution of Figure 10.22, consider the case where customers arrive at a bank at the rate of two customers per minute on the average. The graph shows that the probability that three customers will arrive in one minute is approximately 0.180 or 18%, the probability that four customers will arrive in one minute is approximately 9% and so on. In other words, *the probability of having x events occurring at a given time follows a Poisson distribution.*

Example 10.21

Prove that

$$\sum_{x=0}^{\infty} p(x;\lambda) = 1 \quad (10.113)$$

that is, prove that the Poisson distribution is indeed a probability distribution.

Proof:

We recall that

$$e^y = 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \dots \quad (10.114)$$

Then,

$$e^\lambda = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \quad (10.115)$$

From (10.111), Page 10–47 and (10.115) above,

$$\sum_{x=0}^{\infty} p(x;\lambda) = \sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^\lambda = 1 \quad (10.116)$$

Example 10.22

Let X be an *rv* with Poisson distribution $p(x;\lambda)$. Prove that:

- $E[x] = \mu = \lambda$
- $\text{Var}[x] = \sigma^2 = \lambda$

Proof:

- By definition

$$p(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

and

$$E[x] = \mu = \sum_{x=0}^{\infty} xp(x;\lambda) = \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} \quad (10.117)$$

The presence of x on the right side of (10.117) above makes the entire term vanish for $x = 0$. Therefore, we make the substitution

$$\frac{x}{x!} = \frac{1}{(x-1)!} \quad (10.118)$$

The summation now is from $x = 1$ to $x \rightarrow \infty$. We factor out λ from (10.117) so that we will have the term $x - 1$ present in both the numerator and denominator. With these changes, (10.117) becomes

$$\mu = \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!} \quad (10.119)$$

Next, we make the substitution

$$u = x - 1 \quad (10.120)$$

and we observe that as x goes from 1 to ∞ , u goes from 0 to ∞ . Then,

$$\mu = \lambda \sum_{u=0}^{\infty} \frac{\lambda^u e^{-\lambda}}{u!} \quad (10.121)$$

This summation is unity as we found in (10.116). Therefore, for the Poisson distribution,

$$\mu = \lambda \quad (10.122)$$

b. From (9.46), Chapter 9, Page 9–17,

$$\text{Var}[x] = E[x^2] - \mu^2 \quad (10.123)$$

The mean-square value $E[x^2]$ is

$$E[x^2] = \sum_{x=0}^{\infty} x^2 p(x;\lambda) = \sum_{x=0}^{\infty} x^2 \frac{\lambda^x e^{-\lambda}}{x!} = \lambda \sum_{x=1}^{\infty} x \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!} \quad (10.124)$$

Again, we let $u = x - 1$. Since as x goes from 1 to ∞ , u goes from 0 to ∞ . Then,

$$\begin{aligned}
 E[x^2] &= \lambda \sum_{u=0}^{\infty} (u+1) \frac{\lambda^u e^{-\lambda}}{u!} = \lambda \sum_{u=0}^{\infty} (u+1) p(u; \lambda) \\
 &= \lambda \left[\sum_{u=0}^{\infty} u \frac{\lambda^u e^{-\lambda}}{u!} + \sum_{u=0}^{\infty} \frac{\lambda^u e^{-\lambda}}{u!} \right] = \lambda(\lambda + 1)
 \end{aligned}
 \tag{10.125}$$

where in $\lambda + 1$, λ was obtained from (10.121), and unity from (10.116). Therefore,

$$\text{Var}[x] = E[x^2] - \mu^2 = \lambda(\lambda + 1) - \lambda^2 = \lambda \tag{10.126}$$

Example 10.23

A proofreader found that there are 300 errors distributed randomly in a document consisting of 500 pages. Find the probability that any page contains

- a. exactly two errors
- b. two or more errors

Solution:

Let us represent the number of errors in any page as the number of successes p in a sequence of Bernoulli trials (binomial distribution). Since there are 300 errors, we let $n = 300$. There are 500 pages and thus the probability that exactly one error appears on any page is $p = 1/500$. Since p is small, we will use the Poisson distribution approximation to the binomial distribution, that is, (10.112). Thus,

$$\lambda = np = 300 \times \frac{1}{500} = 0.6 \tag{10.127}$$

Then,

- a. For exactly two errors in any page, $x = 2$ and from (10.111), Page 10–47,

$$p = p(x; \lambda) = (2; 0.6) = \frac{(0.6)^2 e^{-0.6}}{2!} \approx 0.1$$

that is, the probability for exactly two errors in any page is approximately 10%.

- b. The probability that any page will contain two or more errors can be found from the relation

$$p(\text{two or more errors}) = 1 - p(\text{zero or one error})$$

where

$$p(\text{zero errors}) = \frac{(0.6)^0 e^{-0.6}}{0!} = 0.549$$

and

$$p(\text{one error}) = \frac{(0.6)^1 e^{-0.6}}{1!} = 0.329$$

Then,

$$p(\text{two or more errors}) = 1 - (0.549 + 0.329) = 0.122$$

that is, the probability that any page will contain two or more errors is approximately 12.2%.

We can also use Excel's **POISSON** function whose syntax is

=POISSON(x,mean,cumulative)

where

x = the number of events

mean = expected numeric value

cumulative = exponential function. If **TRUE**, the cdf is returned, and if **FALSE**, the pdf is returned.

For example,

=POISSON(0.6,2,TRUE) returns 0.135 and this is approximately equal to the value we found in (a) of Example 10.23.

Example 10.24

An electronics company manufactures memory chips and from past data, it is estimated that 1% are defective. If a sample of 150 chips is selected, what is the probability that there will be two or less defective chips?

Solution:

Here $x = 2$, $n = 150$, and $p = 0.01$. Since p is small, we can use (10.112). Then, the mean number of defectives in a sample of 150 is

$$\mu = \lambda = np = 150 \times 0.01 = 1.5$$

Using Excel,

=POISSON(2,1.5,TRUE) we obtain 0.8088.

This means that the probability that we will find two or less defective chips in a sample of 150 is about 80.9%.

10.13 The Multinomial Distribution

The *multinomial distribution* is a generalization of the binomial distribution and states that in n repeated trials, the probability that event A_1 occurs k_1 times, event A_2 occurs k_2 times and so on until A_m occurs k_m times, where

$$k_1 + k_2 + \dots + k_m = n \quad (10.128)$$

is obtained by

$$p = \frac{n!}{k_1! k_2! \dots k_m!} p_1^{k_1} p_2^{k_2} \dots p_m^{k_m} \quad (10.129)$$

We observe that for $m = 2$, (10.129) reduces to the binomial distribution of (10.12).

The multinomial rv is the random vector

$$\mathbf{X} = [X_1, X_2, \dots, X_n]^T * \quad (10.130)$$

where X_i is the number of times that outcome i occurs.

For the multinomial distribution

$$\text{Mean of } X_i = np_i \quad (10.131)$$

$$\text{Variance of } X_i = \sigma_i^2 = np_i(1 - p_i) \quad (10.132)$$

Example 10.25

A fair die is tossed eight times. Compute the probability that the faces 5 and 6 will appear twice, and all other faces only once.

Solution:

Here, $n =$ number of trials $= 8$, $k_5 = k_6 = 2$ represent the events that 5 and 6 will appear twice, and $k_1 = k_2 = k_3 = k_4 = 1$ represent the events that all other faces will appear only once, and $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = 1/6$ states that each face has equal probability of occurrence. Then, using (10.129) we obtain

$$p = \frac{8!}{2!2!1!1!1!1!} \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right) \left(\frac{1}{6}\right) \left(\frac{1}{6}\right) \left(\frac{1}{6}\right) = \frac{35}{5832} = 0.006 \text{ or } 0.6\%$$

that is, the probability that the faces 5 and 6 will appear twice, and all other faces only once is only 0.6%.

* The exponent T means that the vector X is transposed, that is, is a column vector, not a row vector as in matrices.

10.14 The Hypergeometric Distribution

Let n represent the items selected from a sample of N items in which k items are successes and $(N - k)$ are failures. If the *rv* X is equal to the number of successes in the n selected items, X is said to be a *hypergeometric rv*.

The *hypergeometric distribution* states that

$$P(X = x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} = \frac{k!}{x!(k-x)!} \cdot \frac{(N-k)!}{(n-x)!(N-k-n+x)!} \cdot \frac{N!}{n!(N-n)!} \quad (10.133)$$

For the hypergeometric distribution, the mean μ and variance σ^2 are

$$\mu = n \frac{k}{N} \quad (10.134)$$

$$\sigma^2 = \left(\frac{N-n}{N-1} \right) n \frac{k}{N} \left(1 - \frac{k}{N} \right) \quad (10.135)$$

To illustrate this distribution with an example, consider a box that contains B blue balls and R red balls. Assume that we perform n trials in an experiment in which a ball is drawn at random, its color is observed and noted, and the ball is placed back in the box. This type of experiment is referred to as *sampling with replacement*. If the *rv* X represents the number of B balls chosen (considered as successes) in n trials, then, from the binomial distribution we see that the probability of exactly x successes is

$$P(X = x) = \binom{n}{x} \frac{B^x R^{n-x}}{(B+R)^n} \quad (10.136)$$

where we have used the facts that

$$p = \frac{B}{(B+R)}$$

and

$$q = 1 - p = \frac{R}{(B+R)}$$

Now, if we modify (10.136) so that the sampling is without replacement, we obtain (10.133).

Example 10.26

A statistics and probability class consists of 25 students of whom 5 are foreign students. Suppose

Chapter 10 Common Probability Distributions and Tests of Significance

that 3 students of this class are randomly selected to meet with the dean of mathematics. Compute the probability that at most one of these 3 students will be foreign.

Solution:

The condition that at *most one* will be satisfied when no foreign student or only one student meets with the dean. Here, X is a hypergeometric rv with $N = 25$, $k = 5$, $N - k = 25 - 5 = 20$, and $n = 3$. For this example, the probability distribution of X is

$$P(X = 0) = \frac{\binom{5}{0}\binom{20}{3}}{\binom{25}{3}} = \frac{\frac{5!}{(5-0)!0!} \cdot \frac{20!}{(20-3)!3!}}{\frac{25!}{(25-3)!3!}} = \frac{1 \times 6840}{13800} = 0.496$$

$$P(X = 1) = \frac{\binom{5}{1}\binom{20}{2}}{\binom{25}{3}} = \frac{\frac{5!}{(5-1)!1!} \cdot \frac{20!}{(20-2)!2!}}{\frac{13800}{3!}} = \frac{5 \times 190}{2300} = 0.413$$

$$P(X = 2) = \frac{\binom{5}{2}\binom{20}{1}}{\binom{25}{3}} = \frac{\frac{5!}{(5-2)!2!} \cdot \frac{20!}{(20-1)!1!}}{\frac{13800}{3!}} = \frac{10 \times 20}{2300} = 0.087$$

$$P(X = 3) = \frac{\binom{5}{3}\binom{20}{0}}{\binom{25}{3}} = \frac{\frac{5!}{(5-3)!3!} \cdot \frac{20!}{(20-0)!0!}}{\frac{13800}{3!}} = \frac{60}{13800} = 0.004$$

Figure 10.23 shows the plot of the probability distribution of the rv X for this example.

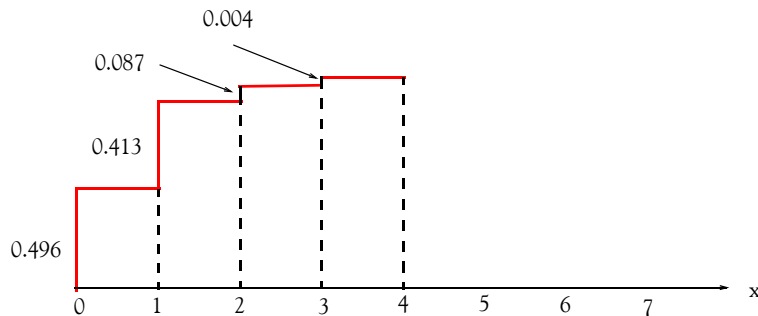


Figure 10.23. Probability distribution of the rv X for Example 10.28

Therefore, the probability that at most one foreign student will meet with the dean is

$$p = 0.496 + 0.413 = 0.909 = 90.9\%$$

We can use Excel's HYPGEOMDIST function whose syntax is

=HYPGEOMDIST(sample_s,number_sample,population_s,number_population)

where:

sample_s = number of successes in the sample

number_sample = the size of the sample

population_s = number of successes in the population

number_population = population size

For Example 10.26, success is the condition that at *most one* will be satisfied, when no foreign student or only one student meets with the dean. Then,

=HYPGEOMDIST(0,3,5,25) returns 0.496 and

=HYPGEOMDIST(1,3,5,25) returns 0.413.

Therefore, the probability that at most one foreign student will meet with the dean is $0.496 + 0.413 = 0.909$ and this is the same result as before.

Note: The binomial distribution becomes an approximation to the hypergeometric distribution when $n \leq 0.05N$.

Example 10.27

A box contains 200 computer chips and it is known that 7 of these are defective. Compute the probability of finding one defective in a sample of 5 randomly selected chips by using:

- the hypergeometric distribution
- the binomial distribution, if applicable for this example

Solution:

- Using (10.133), Page 10–53, with $n = 5$, $N = 200$ and $k = 7$, we obtain

$$P(X = 1) = \frac{\binom{7}{1} \binom{193}{4}}{\binom{200}{5}} = \frac{7!}{1!(7-6)!} \cdot \frac{(193)!}{(5-1)!(193-4)!} = \frac{7 \times 56031760}{\frac{200!}{5!(200-5)!}} = \frac{7 \times 56031760}{2.5357 \times 10^9} = 0.155 \quad (10.137)$$

- b. Let the probability of selecting a defective chip be denoted as p (success); then, the probability of failure is $q = 1 - p$. To find out if the binomial distribution is applicable, we examine the condition

$$n \leq 0.05N = 0.05 \times 200 = 10$$

Since this number is greater than $n = 5$, we can apply the binomial distribution. Then,

$$P(X = 1) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{(n-x)!x!} = \frac{5!}{4!1!} \times \left(\frac{7}{200}\right)^1 \times \left(1 - \frac{7}{200}\right)^4 = 0.152 \quad (10.138)$$

We observe that (10.137) and (10.138) are in close agreement.

10.15 The Bivariate Normal Distribution

The *bivariate normal distribution* is a generalization of the normal distribution. It applies to two continuous *rv* X and Y given by the joint density function

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\left\{ \frac{\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]}{2(1-\rho^2)} \right\}} \quad (10.139)$$

where $\mu_1, \mu_2, \sigma_1, \sigma_2$ are the means and standard deviations of the X and Y respectively, and ρ is the *correlation coefficient** between X and Y . This distribution is also referred to as the *two-dimensional normal distribution*.

If the correlation coefficient $\rho = 0$, (10.139) reduces to a product of two factors that is a function of x and y alone for all values of x and y and therefore, X and Y are independent.

10.16 The Rayleigh Distribution

This distribution is a special case of the bivariate normal distribution where the means are zero, the standard deviations are equal, and X and Y are uncorrelated, that is, $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = \sigma$, and $\rho = 0$. In this case, (10.139) reduces to

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-\{(x^2+y^2)/2\sigma^2\}} \quad (10.140)$$

* We will discuss the correlation coefficient in Chapter 11.

This expression is known as the *circular case of the bivariate normal distribution* and the distance to the origin of the x and y is denoted as r where $r = \sqrt{x^2 + y^2}$.

The *Rayleigh distribution* is defined as

$$f_r(r) = \frac{r}{\sigma^2} e^{-(r^2/2\sigma^2)} \quad \text{for } r \geq 0 \quad (10.141)$$

where $r = \sqrt{x^2 + y^2}$, and σ^2 is the variance of the variables x and y in (10.140), Page 10–56.

The Rayleigh pdf is shown in Figure 10.24 for $\sigma = 1$, $\sigma = 2$, and $\sigma = 3$.

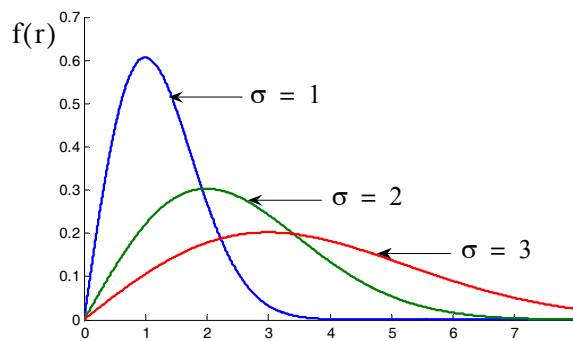


Figure 10.24. Pdf of the Rayleigh distribution

This distribution is used with the scattering of electromagnetic radiation by particles that have dimensions much smaller than the wavelength of the radiation, resulting in angular separation of colors, and responsible for the reddish color of sunset and the blue of the sky.

The Rayleigh distribution is characterized by the following constants, all normalized to the parameter σ .

1. Most probable value is σ . That is, for $\sigma = 1$, the distribution has its maximum value at $x = 1$, for $\sigma = 2$ at $x = 2$ and so on, as shown in Figure 10.24.
2. The mean μ , also known as *mean radial error*, or expected value $E(r)$ is

$$\mu = E(r) = \sigma \sqrt{\frac{\pi}{2}} \quad (10.142)$$

3. The 50% cumulative distribution point is $r = 1.18\sigma$. This is the point where, on the average, 50% of the time the function is above this value, and 50% of the time below it. It is found from

$$\int_r^{\infty} f_r(r) dr = e^{-\frac{r^2}{2\sigma^2}} = 0.5 \quad (10.143)$$

4. The second moment is

$$E(r^2) = 2\sigma^2 \quad (10.144)$$

5. The variance $\text{Var}(r)$ is

$$\text{Var}(r) = \sigma^2(2 - \pi/2) \quad (10.145)$$

Example 10.28

Use the Rayleigh distribution to compute the probability that:

- a. a transmitted signal will fade below the most probable value of σ
- b. a received signal will exceed some arbitrary value $k\sigma$

Solution:

a. Integrating (10.141), Page 10–57, we obtain

$$\int_0^{\sigma} f_r(r) dr = \int_0^{\sigma} \frac{re^{-\frac{r^2}{2\sigma^2}}}{\sigma^2} dr = \frac{1}{\sigma^2} \int_0^{\sigma} re^{-\frac{r^2}{2\sigma^2}} dr$$

and from tables of integrals (or integration by parts)

$$\int xe^{-x^2} dx = -\frac{1}{2}e^{-x^2} + C$$

Then,

$$\frac{1}{\sigma^2} \int_0^{\sigma} re^{-\frac{r^2}{2\sigma^2}} dr = \frac{1}{\sigma^2} \left(-\frac{1}{2} 2\sigma^2 e^{-\frac{r^2}{2\sigma^2}} \right) \Bigg|_0^{\sigma} = -e^{-\frac{r^2}{2\sigma^2}} \Bigg|_0^{\sigma} = 1 - e^{-1/2} = 0.394$$

This indicates that there is a probability of 39.4% that the transmitted signal will be below the most probable value of σ .

b. The probability that a received signal will exceed some arbitrary value $k\sigma$ is

$$\int_{k\sigma}^{\infty} f_r(r) dr = \int_{k\sigma}^{\infty} \frac{re^{-\frac{r^2}{2\sigma^2}}}{\sigma^2} dr = \frac{1}{\sigma^2} \int_{k\sigma}^{\infty} re^{-\frac{r^2}{2\sigma^2}} dr = -e^{-\frac{r^2}{2\sigma^2}} \Bigg|_{k\sigma}^{\infty} = e^{-k^2/2} \quad (10.146)$$

10.17 Other Probability Distributions

There are many other probability distributions used with different applications. The gamma and beta distributions are discussed in Appendix C; for a few of the others, we briefly present their basic properties in Subsections 10.17.1 through 10.17.7 below

10.17.1 The Cauchy Distribution

The *Cauchy distribution* has the pdf

$$f_X(x) = \frac{a}{\pi(x^2 + a^2)} \quad a > 0, \quad -\infty < x < \infty \quad (10.147)$$

This distribution is symmetrical about $x = 0$ and thus its mean is zero. However, the variance and higher moments do not exist.

10.17.2 The Geometric Distribution

The *geometric distribution* has the pdf

$$f_X(x) = pq^{x-1} \quad x = 1, 2, \dots \quad (10.148)$$

with mean

$$\mu = 1/p \quad (10.149)$$

and variance

$$\sigma^2 = q/p^2 \quad (10.150)$$

10.17.3 The Pascal Distribution

The *Pascal distribution* or *negative binomial distribution* has the pdf

$$f_X(x) = \binom{x-1}{r-1} p q^{x-r} \quad x = r, r+1, \dots \quad (10.151)$$

with mean

$$\mu = r/p \quad (10.152)$$

and variance

$$\sigma^2 = rq/p^2 \quad (10.153)$$

We observe that for $r = 1$ the negative binomial distribution of (10.151) reduces to the geometric distribution given by (10.148) above.

The negative binomial distribution is similar to the binomial distribution, except that the number of successes is fixed and the number of trials is variable. For example, suppose that we need to find 10 students with GPA 3.0 or higher, and we know that the probability that a student has this qualification is 0.25. This distribution calculates the probability that we must search for x stu-

dents, say 30, before we find 10 students with GPA 3.0 or higher.

We can use Excel's **NEGBINOMDIST** function to compute the probability that there will be a number of f failures before we find a number of s successes given a constant probability of success. The syntax for this function is

=NEGBINOMDIST(number_f,number_s,probability_s)

where

number_f = number of failures

number_s = number of successes

probability_s = known probability of success.

Therefore, if we want to find the probability that we must search 30 unqualified before we find 10 qualified students, and we know that the probability is 0.25, the function

=NEGBINOMDIST(30,10,0.25) returns 0.036011.

This means that the probability that we will find 30 unqualified students before we find 10 qualified is 3.601%.

10.17.4 The Weibull Distribution

The *Weibull distribution* has the pdf

$$f_X(x) = \begin{cases} \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (10.154)$$

with mean

$$\mu = \alpha^{-1/\beta} \Gamma\left(1 + \frac{1}{\beta}\right) \quad (10.155)$$

and variance

$$\sigma^2 = \alpha^{-2/\beta} \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \right] \quad (10.156)$$

The Weibull distribution is used by reliability engineers to determine probabilities such as the mean time to failure for a particular device.

We can use Excel's **WEIBULL** function for the Weibull distribution. The syntax for this function is

=WEIBULL(x,alpha,beta,cumulative)

where

x = the value at which to evaluate the distribution

alpha = value of the parameter α as it appears in (10.154)

beta = value of the parameter β as it appears in (10.154)

cumulative = **true** if the cdf is desired, or **cumulative** = **false** if the pdf is desired.

For example, if $x = 100$, $\alpha = 10$, $\beta = 95$ and **cumulative** = **true**, then

=WEIBULL(100,10,95,TRUE) returns 0.811787,

and if **cumulative** = **false**, then

=WEIBULL(100,10,95,FALSE) returns 0.031435.

10.17.5 The Maxwell Distribution

The *Maxwell distribution* has the pdf

$$f_X(x) = \begin{cases} \sqrt{\frac{2}{\pi}} a^{3/2} x^2 e^{-ax^2/2} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (10.157)$$

with mean

$$\mu = 2\sqrt{\frac{2}{\pi a}} \quad (10.158)$$

and variance

$$\sigma^2 = \frac{1}{a} \left(3 - \frac{8}{\pi} \right) \quad (10.159)$$

10.17.6 The Lognormal Distribution

The *lognormal distribution* has the pdf

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{1}{2\sigma^2} \ln(x-\mu)^2} \quad x > 0, \quad \sigma > 0 \quad (10.160)$$

with mean

$$\mu = e^{\frac{\mu + \sigma^2}{2}} \quad (10.161)$$

and variance

$$\sigma^2 = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1) \quad (10.162)$$

This distribution is used to analyze data that have been logarithmically transformed.

Excel provides the **LOGNORMDIST** and **LOGINV** functions. The first, computes the cumulative lognormal distribution of x where $\ln x$ is normally distributed with parameters the mean and standard deviation. The second, computes the inverse of the lognormal cumulative distribution. Thus, if $p = \text{LOGNORMDIST}(x, \dots)$, then $x = \text{LOGINV}(p, \dots)$.

Excel uses the following expression for the lognormal cumulative distribution.

$$\text{LOGNORMDIST}(x, \mu, \sigma) = \text{NORMDIST}\left(\frac{\ln x - \mu}{\sigma}\right) \quad (10.163)$$

The syntax for **LOGNORMDIST** is

=LOGNORMDIST(x,mean,standard_dev)

where

x = the value at which we want to evaluate the distribution

mean = the mean of $\ln x$

standard_dev = the standard deviation of $\ln x$

For example,

=LOGNORMDIST(3,4.5,1.6) returns 0.0168

The syntax for **LOGINV** is

=LOGINV(probability,mean,standard_dev)

where

probability = the probability associated with the lognormal distribution

mean = the mean of $\ln x$

standard_dev = the standard deviation of $\ln x$

For example,

=LOGINV(0.0168,4.5,1.6) returns 3.005

10.18 Critical Values of the Binomial Distribution

Often, we are interested in determining the smallest value for which the binomial distribution is greater than, or equal to a criterion value. Such values are referred to as the critical values of the binomial distribution. Of course, critical values can also be established in other probability distributions. The Excel the **CRITBINOM** function and this returns the smallest value for which the binomial distribution is greater than, or equal to a criterion value α (alpha) where $0 < \alpha < 1$. This

distribution is used in quality assurance applications. For example, it can be used to determine the largest number of defective parts that are allowed to go through the assembly line without rejecting the entire lot. Its syntax is

`=CRITBINOM(trials,probability_s,alpha)`

where:

`trials` = number of Bernoulli trials

`probability_s` = probability of success on each trial

`alpha` = criterion value

Example 10.29

Consider a plant that manufactures automobile carburetors. These are manufactured in lots of 100 and there is 90% probability that each carburetor is free of defects. In a lot of 100, what is the number of carburetors that we can say with 95% confidence are free of defects?

Solution:

For this example,

`trials` = 100

`probability_s` = 0.9

`alpha` = $1 - 0.95 = 0.05$

Then,

`=CRITBINOM(100,0.9,0.05)` returns 85 .

Therefore, in a lot 100 , the number of carburetors that we can say with 95% confidence are free from defects is 85 .

10.19 Sampling Distribution of Means

Suppose we take a large number of random samples from the same population. Each of these many samples will have its own sample statistic called the *sampling distribution for the sampled mean*, or simply the *sampled mean* which is denoted as $\mu_{\bar{x}}$. It is shown in probability theory textbooks that:

1. If we take a large number of samples, the sampled mean $\mu_{\bar{x}}$ will approach the population mean μ . That is,

$$\mu_{\bar{x}} = \mu \quad (10.164)$$

2. If the population is very large, or the sampling is with replacement, the *standard error* of the sampling distribution of means, denoted as $\sigma_{\bar{x}}$ is given by

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad (10.165)$$

where σ is the standard deviation of the population, and n is the size of a sample.

3. If the population is of size N , the sampling is without replacement, and the sample size is $n \leq N$, then,

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right) \quad (10.166)$$

4. If the population from which samples are taken, has a probability distribution with mean μ , variance σ^2 , and is not necessarily a normal distribution, then the standardized variable z associated with the sampled-mean \bar{x} is given by

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \quad (10.167)$$

5. If n is very large, z becomes asymptotically normal, that is,

$$\lim_{n \rightarrow \infty} P(Z \leq z) \Rightarrow \text{normal distribution} \quad (10.168)$$

10.20 Z-Score*

A *z-score* represents the number of standard deviations that a value x , obtained from a given observation, is above or below the mean. For the entire population, it is defined as

$$z = \frac{x - \mu}{\sigma} \quad (10.169)$$

Example 10.30

The mean salary for engineers employed in the San Francisco Bay Area is \$85,000 per year and the standard deviation is \$14,000. The mean salary for engineers employed in the Austin, Texas

* The Z-score discussed in this section has no relation to The Altman Z-score discussed in Appendix G.

area is \$65,000 per year and the standard deviation is \$8,000. An engineer who makes \$90,000 per year in the San Francisco Bay Area is offered a job in Austin, Texas for \$78,000 per year. Assuming that the overall cost of living in both locations is relative to the mean salaries, should this engineer accept the job in Austin if salary is his only consideration?

Solution:

For the engineer who makes \$90,000 per year in the San Francisco Bay Area, the z score is

$$z_1 = \frac{x_1 - \mu_1}{\sigma_1} = \frac{90,000 - 85,000}{14,000} = 0.357$$

If he moves to Austin, his z score will be

$$z_2 = \frac{x_2 - \mu_2}{\sigma_2} = \frac{78,000 - 65,000}{8,000} = 1.625$$

Since $z_2 > z_1$, we conclude the \$78,000 salary in Austin, Texas is higher relative to the \$90,000 salary in the San Francisco Bay Area.

10.21 Tests of Hypotheses and Levels of Significance

This is a lengthy topic in probability theory and thus, it will not be discussed here in detail. We will define some of the important terminology with examples.

Suppose that the mean sales of a mail order company selling computer parts for the past year was \$125, and the standard deviation was \$30. This company wishes to determine whether the mean sales per order for the current year, will be the same or different from the past year. Thus, if μ represents the mean sales per order for the current year, the company is interested in determining whether for the current year $\mu = \$125$ or $\mu \neq \$125$. If the company decides that the current year's mean sales order will be the same as in last year, that is, if $\mu = \$125$, we call this the *null hypothesis*. On the other hand, if the company wishes to conduct a research to determine if $\mu \neq \$125$, this is called the *research hypothesis* or *alternative hypothesis*.

In general, if $\mu =$ some known value represents the null hypothesis, the test to determine the alternate hypothesis for $\mu <$ some known value is called a *one-tail left test* or a *lower-tail test*. For example, if an automobile tire manufacturer claims that the mean mileage for the top-of-the-line tire is $\mu = 75,000$ miles, and a consumer group doubts the claim and decides to test for $\mu < 75,000$ miles, this is a *one left tail test*.

The test to determine the alternate hypothesis for $\mu >$ some known value is called a *one-tail right test* or an *upper-tail test*. For example, a sports league may advertise that the average cost for a family of four to attend a soccer game is \$100 including parking tickets, snacks and soft drinks, but a

Chapter 10 Common Probability Distributions and Tests of Significance

sports fan thinks that this average is low, and decides to conduct a research hypothesis for $\mu > \$100$. This is a *one right-tail test*.

If the research hypothesis is for the condition $\mu \neq$ some known value, regardless whether $\mu <$ some known value or $\mu >$ some known value, this test is referred to as a *two-tail test*. For example, if the computer company conducts a research test for both $\mu < \$125$ and $\mu > \$125$, we call this a *two-tail test*.

A *test statistic*, denoted as \bar{x} , is a statistical measure that is used to decide whether to accept or reject the null hypothesis. If the null hypothesis is true, then by the central limit theorem* $\mu = \bar{x} = \$125$ on the average. If we take a sample of 100 orders for the current year, and we assume that the standard deviation remains constant, that is, $\sigma = \$30$, the standard error, as defined in (10.165), Page 10–64, is

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\$30}{\sqrt{100}} = \$3$$

If the value of the test statistic \bar{x} is close† to $\mu = \$125$, we would not reject the null hypothesis, but if \bar{x} is considerably different than μ , we would reject the null hypothesis.

The decision we must now make, is by how much should \bar{x} differ from $\mu = \$125$ in order to reject the null hypothesis. Suppose that we decide to reject the null hypothesis if \bar{x} differs from $\mu = \$125$ by two or more standard errors, that is, $2\sigma_{\bar{x}} = 2 \times \$3 = \$6$. Therefore, if $\bar{x} \leq \$119$ or if $\bar{x} \geq \$131$, we should reject the null hypothesis. The values $\bar{x}_1 = \$119$ and $\bar{x}_2 = \$131$ are referred to as the *critical values*.

Example 10.31

An automobile tire manufacturer claims that the mean mileage for one of the best tires is $\mu = 70,000$ miles, and the standard deviation is $\sigma = 8,000$ miles. Suppose that the sampled mean is normally distributed, and the critical value is chosen as two standard errors below 70,000 miles. If 100 tires are road tested, and it is found that the mean mileage of the tested tires is $\bar{x} = 62,000$, should we accept or reject the manufacturer's claim?

Solution:

Obviously, this is a one-tail (left) test. For this example, the null hypothesis is $\mu = 70,000$ miles and the research hypothesis is $\mu < 70,000$ miles. From (10.165), Page 10–64, the standard error

* This theorem states that in any sequence of n repeated trials ($n \geq 30$), the standardized sampled mean approaches the standard normal curve as n increases.

† Of course, we have to decide on how close this value should be.

is

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{8,000}{\sqrt{100}} = 800 \text{ miles}$$

The value of two standard errors is equal to $2\sigma_{\bar{x}} = 2 \times 800 = 1,600$ miles below $\mu = 70,000$ miles or $70,000 - 1,600 = 68,400$ miles. Since the mean mileage of the tested tires was given as $\bar{x} = 62,000$ miles, the tire manufacturer's claim should be rejected.

Example 10.32

A sports league advertises that the average cost for a family of four to attend a soccer game, is $\mu = \$100$ including admission, parking, snacks and soft drinks. The standard deviation is $\sigma = \$20$. Suppose that the critical value is chosen as three standard errors above $\mu = \$100$, and the mean cost in a sample of 100 families of four, was found to be $\bar{x} = \$112$. Should we accept or reject the advertising claim of this sports league?

Solution:

For this example, the null hypothesis is $\mu = \$100$ and the research hypothesis is $\mu > \$100$. The standard error is

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\$20}{\sqrt{100}} = \$2$$

The value of three standard errors is equal to $3\sigma_{\bar{x}} = 3 \times 2 = \6 above $\mu = \$100$ or $\$100 + \$6 = \$106$. Since the mean cost of the 100 families of four is given as $\bar{x} = \$112.00$, the claim of this sports league should be rejected.

If a hypothesis is rejected when it should be accepted, we say that a *Type I error* has occurred. If a hypothesis is accepted when it should be rejected, we say that a *Type II error* has occurred.

The probabilities of Type I and Type II errors are denoted as follows:

α = probability of making a Type I error

β = probability of making a Type II error

In other words,

α = P(rejection of the null hypothesis when it should be accepted)

β = P(acceptance of the null hypothesis when it should be rejected)

The probability α is referred to as the *level of significance*.

Chapter 10 Common Probability Distributions and Tests of Significance

Typically, we use a level of significance of $\alpha = 0.05$ or $\alpha = 0.01$. For example, if we choose a 0.05 or 5% level of significance when deciding on a test of hypothesis, this means that there are about 5 chances in 100 that we would reject the hypothesis when it should be accepted. In other words, we are about 95% confident that we have made the right decision.

For the mail-order computer company, the level of significance α can be expressed as

$$\alpha = P(\bar{x} < 119 \text{ or } \bar{x} > 131) = P(\bar{x} < 119) + P(\bar{x} > 131) \quad (10.170)$$

and $\mu = 125$.

To evaluate the probability P , we use the relation of (10.167), Page 10–64, that is,

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \quad (10.171)$$

which, as we know, for large n becomes asymptotically normal, that is, it approximates the standard normal distribution.

Assuming that the standard error is $\sigma_{\bar{x}} = \$3$, as before, and the null hypothesis is true, then from (10.171) above,

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{\bar{x} - 125}{3} \quad (10.172)$$

To evaluate α in (10.170) above, we make use of the fact that the event $\bar{x} < 119$ is equivalent to the event

$$z = \frac{\bar{x} - 125}{3} < \frac{119 - 125}{3} = -2$$

that is,

$$z < -2 \quad (10.173)$$

Also, the event $\bar{x} > 131$ is equivalent to the event

$$z = \frac{\bar{x} - 125}{3} > \frac{131 - 125}{3} = 2$$

that is,

$$z > 2 \quad (10.174)$$

Therefore, with (10.170), (10.173) and (10.174), α may be expressed as

$$\alpha = P(\bar{x} < 119) + P(\bar{x} > 131) = P(z < -2) + P(z > 2) \quad (10.175)$$

We will extract the values of z from the standard normal distribution table using Figure 10.25 below.

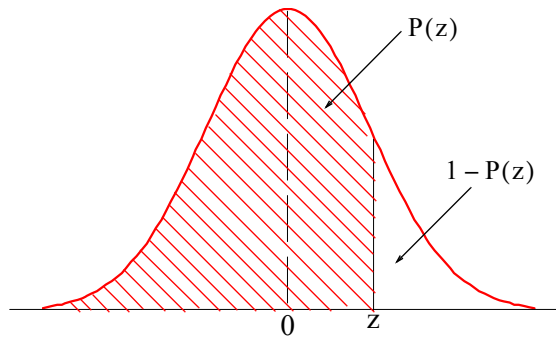


Figure 10.25. The standard normal probability function

From normal distribution tables, we find that the area to the left of $z = 2$ is 0.9772. The same value is found by using Excel's **NORMSDIST(z)** function which produces the standard normal cumulative distribution. Thus,

`=NORMSDIST(2)` returns 0.9772.

This is the area to the left of $z = 2$; we want to find the area to the right of it. Obviously, the area to the right is $1 - 0.9772 = 0.0228$ and thus

$$P(z > 2) = 0.0228 \quad (10.176)$$

Next, we need to find the value for $P(z < -2)$. For this, we use the fact that the normal distribution is symmetric, and thus we can use the relation

$$F(-z) = 1 - F(z) \quad (10.177)$$

Therefore,

$$P(z < -2) = P(z > 2) = 0.0228$$

and by substitution into (10.175), Page 10-68,

$$\alpha = P(z < -2) + P(z > 2) = 0.0228 + 0.0228 = 0.0456 \quad (10.178)$$

This means that when using the mean of a sample of 100 sales, and the rejection regions are as before, i.e., $\bar{x} \leq \$119$ or $\bar{x} \geq \$131$, we say that 4.56% of the time the mailing order computer company will conclude that the mean has changed when in fact it has not. Stated differently, we can say that 95.44% of the time, the company will conclude that the mean has not changed.

In the discussion above, the acceptance and rejection regions were specified and we found the level of significance.

Now, suppose that the level of significance is specified and we want to find the acceptance and rejection regions.

Chapter 10 Common Probability Distributions and Tests of Significance

We will again consider the example for the mailing order computer company and will assume the level of significance $\alpha = 0.01$

Because the hypothesis test is a two-tail test, we divide α into two halves, i.e., 0.005 each. Then, we need to find a value a such that $P(z > a) = 0.005$ and $P(z < -a) = 0.005$.

Using the standard normal distribution table, we find that for the probability $P(z < -a) = 0.005$, $a = -2.576$ and for $P(z > a) = 0.005$, $a = 2.576$.

The same value can be found with Excel's **NORMSINV(p)** function which produces the inverse of the standard normal cumulative distribution. Thus, for our example,

=NORMSINV(0.005) returns -2.576 .

These values define the rejection regions and are shown in Figure 10.26.

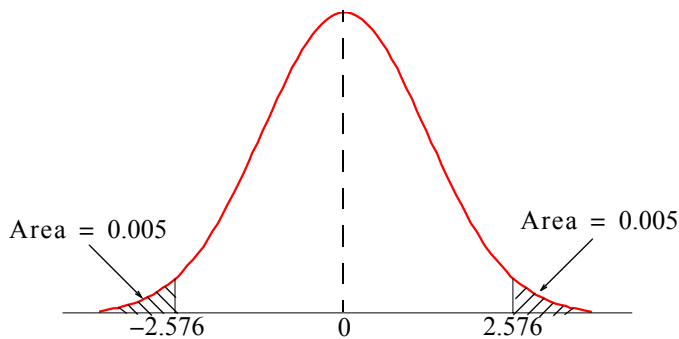


Figure 10.26. Calculated areas for rejection region

The final step is to determine the rejection region in terms of the test statistics \bar{x}_1 and \bar{x}_2 . For this, we use relation

$$z = \frac{\bar{x} - 125}{3} \quad (10.179)$$

and we replace $z < -2.576$ with

$$\frac{\bar{x}_1 - 125}{3} < -2.576 \quad (10.180)$$

Also, we replace $z > 2.576$ with

$$\frac{\bar{x}_2 - 125}{3} > 2.576 \quad (10.181)$$

To find \bar{x}_1 , we multiply both sides of (10.180) above by 3 and add 125 to both sides. Then,

$$\bar{x}_1 = \$117 \quad (10.182)$$

Likewise, to find \bar{x}_2 , we multiply both sides of (10.181) by 3 and add 125 to both sides. Then,

$$\bar{x}_2 = \$133 \quad (10.183)$$

If a sampling distribution of a test statistic \bar{x} , is approximately normally distributed with a sample size $n \geq 30$, we expect to find \bar{x} in the interval $\mu_{\bar{x}} - \sigma_x$ to $\mu_{\bar{x}} + \sigma_x$ about 68.27% of the time. Likewise, we expect to find \bar{x} in the interval $\mu_{\bar{x}} - 2\sigma_x$ to $\mu_{\bar{x}} + 2\sigma_x$ about 95.45% of the time, and in the interval $\mu_{\bar{x}} - 3\sigma_x$ to $\mu_{\bar{x}} + 3\sigma_x$ about 99.73% of the time. These intervals are referred to as *confidence intervals* for estimating \bar{x} , and their respective percentages are called *confidence limits* or *confidence levels*. Recalling that α is the level of significance (0.01 or 0.05), we find that

$$\text{confidence level} = 1 - \text{level of significance} = 1 - \alpha \quad (10.184)$$

From tables of the normal distribution, we find that for a confidence level of 95%, the value of z is 1.96, and for 99% is 2.58. These values of z are denoted as z_c , and the confidence interval can be found from

$$\text{confidence interval} = \bar{x} \pm z_c \frac{\sigma}{\sqrt{n}} \quad (10.185)$$

Example 10.33

In a sample of 100 taxpayers, the average federal tax was found to be \$23,000 with a standard deviation of \$4,000. Compute the 95% confidence interval.

Solution:

For this example, $\bar{x} = 23,000$, $\sigma = 4,000$, $n = 100$ and the value of z_c corresponding to the 95% confidence interval is $z_c = 1.96$. Then, from (10.185),

$$\text{confidence interval} = 23,000 \pm 1.96 \frac{4,000}{\sqrt{100}} = 23,000 \pm 784$$

Therefore, we can say with 95% confidence that the lowest tax was \$22,216 and the highest was \$23,784.

We can verify our result with Excel's **CONFIDENCE** function which returns the confidence interval for a population mean. Its syntax is

`=CONFIDENCE(alpha,standard_dev,size)`

where

alpha = level of significance, and therefore, confidence level = $100(1 - \alpha)\%$

standard_dev = standard deviation for the sample

size = sample size

For instance, =CONFIDENCE(0.05,4000,100) returns 784 and this is the same value we found with relation (10.185), Page 10–71.

10.22 Tests of Significance

In statistics, significant means probably true and a *test of significance* is the process of comparing two hypotheses. In Subsections 10.11.1 through 10.11.4 below we will briefly discuss the z–test, the t–test, the F–test, and the χ^2 –test.

10.22.1 The z–Test

The z–test is used to find the probability that a particular observation is drawn from a particular population. This test generates a standard score for a value x with respect to the data set or array, and returns the two–tail probability for the normal distribution. It is called z–test because we use the standard deviation as a unit (z–unit), for measuring the point where the critical region begins (e.g. 2 standard deviations = 5%, 2.5 standard deviations = 1% for two tail tests), and relate to the proportions of the normal curve. This test is performed for populations with large samples.

Example 10.34

The mean of a particular population of data is $\mu = 16$ and the standard deviation is $\sigma = 3.04$. Compute the probability that the observed data

{12, 19, 21, 22, 18, 16, 15, 17}

were drawn from that population.

Solution:

Here, the observed data consists of eight values, and thus $n = 8$. The standard error $\sigma_{\bar{x}}$ is found from (10.165), Page 10–64, that is,

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{3.04}{\sqrt{8}} = 1.0748$$

We use Excel's AVERAGE(number1,number2,...) function to find the mean of the given observed data, that is,

=AVERAGE(12,19,21,22,18,16,15,17) and this returns the sampled–mean $\bar{x} = 17.5$.

Using (10.167), Page 10–64, we find the z value as

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{17.5 - 16}{1.0748} = 1.3956$$

The area (probability), corresponding to this normalized z value, can be found from tables of the standard normal distribution or with the Excel's **NORMSDIST** function. Here, we seek the area to the left of $z = -1.3956$. Then, **=NORMSDIST(-1.3956)** returns 0.081418.

The same value is obtained with Excel's **ZTEST** function whose syntax is

ZTEST(array,x,sigma)

where

array = range of data

x = value to test

sigma = standard deviation of the population

For this example,

array = {12, 19, 21, 22, 18, 16, 15, 17}, **x** = 16, and **sigma** = 3.04. Then,

=ZTEST({12,19,21,22,18,16,15,17},16,3.04) which returns 0.081417.

This indicates that approximately only 8.15% of the time we can say that the observed data came from a population with mean $\mu = 16$ and standard deviation $\sigma = 3.04$.

10.22.2 The t-Test

The t -test is similar to the z -test but it is used with populations whose samples are small (30 or less). We use this test to determine the probability that two samples are drawn from populations that have the same mean. Performing a t -test is easy with Excel's **TTEST** function whose syntax is

=TTEST(array1,array2,tails,type)

where

array1 = the first data set

array2 = the second data set

tails=1 specifies the one-tail test and **tails=2** specifies the two-tail test.

If **type=1**, a *paired** test is performed.

Chapter 10 Common Probability Distributions and Tests of Significance

If $\text{type}=2$, a t -test is performed for samples from populations with the same variance (*homoscedastic*); with this test, **array1** and **array2** need not have the same number of data points.

If $\text{type}=3$, a t -test is performed for samples from populations with unequal variance (*heteroscedastic*); with this test, **array1** and **array2** need not have the same number of data points.

Example 10.35

Use Excel's **TTEST** function to perform a two-tail, paired t -test, if we are given that array 1 = {4, 5, 6, 9, 10, 2, 3, 5, 6}, and array 2 = {5, 18, 2, 1, 13, 3, 4, 16, 7}.

Solution:

We use the Excel function

$\text{=TTEST}(\{4,5,6,9,10,2,3,5,6\},\{5,18,2,1,13,3,4,16,7\},2,1)$ and this returns the value 0.361682.

Therefore, we can say that the probability that the data of **array1** and **array2** have come from populations that have the same mean is 36.17%.

Example 10.36

Use Excel to perform a one-tail, heteroscedastic t -test given that

array 1 = {4, 5, 6, 9, 10, 2, 3, 5, 6, 8} and array 2 = {5, 18, 2, 1, 13, 3, 4, 16, 7}.

Solution:

Here, we observe that array 1 consists of 10 data points, whereas array 2 has only 9 data points. This is permissible for heteroscedastic type t -tests. Then,

$\text{=TTEST}(\{4,5,6,9,10,2,3,5,6,8\},\{5,18,2,1,13,3,4,16,7\},1,3)$ returns 0.214937.

This indicates that the probability that the data of array 1 and array 2 have come from the same populations that have the same mean is 21.49%.

10.22.3 The F-Test

The F -test produces the one-tail probability that the variances in two arrays of data points are not significantly different. This test also determines whether two samples have different variances. For example, given scores of tests from two different universities, we can use the F -test to determine

* A paired test is one where array1 and array2 have the same number of data points.

whether these universities have different levels of diversity, i.e., different statistics, such as mean and variance.

We can perform an F-test with the Excel's **FTEST(array1,array2)** function.

With this function, **array1** and **array2** need not have the same number of data points.

Example 10.37

Let

$$\text{array 1} = \{56, 67, 79, 85, 91, 75, 69, 71, 82, 59\}$$

represent the test scores for an electrical engineering course at University A and

$$\text{array 2} = \{49, 88, 73, 81, 58, 95, 83, 86, 90\}$$

the test scores for the same course at University B. Use Excel's **FTEST** function to determine the probability that the scores are not significantly different.

Solution:

Here, we observe that **array1** consists of 10 scores, whereas **array2** has only 9 scores. This is permissible for F-tests. Using Excel's **FTEST** function, we obtain

=FTEST({56,67,79,85,91,75,69,71,82,59},{49,88,73,81,58,95,83,86,90}) and this returns the value 0.361613.

This indicates that the probability that the variances in the scores of these two universities are not significantly different is 36.16%.

The F-test is used in the *Analysis of Variance* (ANOVA). We will discuss it in Chapter 12.

10.22.4 The Chi-Square Test

The χ^2 test is used extensively in social statistics. It tests a null hypothesis* that the relative frequencies of occurrence of observed events follow a specified frequency distribution. The events must be mutually exclusive.

Consider the data of Table 10.3 representing the data obtained from a random sample of job applicants, 100 men and 100 women.

* We recall that a null hypothesis is a hypothesis set up to be nullified (refuted) in order to support an alternative hypothesis. When used, the null hypothesis is presumed true until statistical evidence in the form of a hypothesis test indicates otherwise.

We call these *actual frequencies*.

TABLE 10.3 Table of actual frequencies

	Work Experience		
Sex	Yes	No	Total
Male	70	30	100
Female	50	50	100
Total	120	80	200

Does this sample represent a real and reliable difference between the population proportions of male and female applicants?

To answer this question, we start with the hypothesis that men and women are equally likely to have had work experience; we call this the null hypothesis. Based on the null hypothesis, we draw a similar table with *expected frequencies* as shown in Table 10.4 where we have used the same proportion as the actual frequencies; in other words, $120/200 = 0.6 = 60\%$. That is, in constructing the table for the expected frequencies, we assume that 60 men in a sample of 100 men have work experience, and likewise, 60 women in a sample of 100 have work experience.

TABLE 10.4 Table of expected frequencies

	Work Experience		
Sex	Yes	No	Total
Male	60	40	100
Female	60	40	100
Total	120	80	200

Next, we combine the actual (A) and expected (E) frequencies into a single table to see how they differ. This is done in Table 10.5.

Table 10.5 shows that the difference between the actual and expected frequency is +10 or -10.

To decide whether this difference is large enough to be considered significant, we must calculate the χ^2 statistic, also known as the χ^2 test for goodness of fit. It is computed from the relation

TABLE 10.5 Combined actual and expected frequencies

	Work Experience				
Gender		Yes		No	Total
Male	A	70	A	30	100
	E	60	E	40	100
Female	A	50	A	50	100
	E	60	E	40	100

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(A_{ij} - E_{ij})^2}{E_{ij}} \quad (10.186)$$

where

A_{ij} = actual frequency in the i th row, j th column

E_{ij} = expected frequency in the i th row, j th column

r = number of rows

c = number of columns

If $\chi^2 = 0$, the actual and expected frequencies agree exactly. If $\chi^2 > 0$, they are in disagreement, and the larger the value of χ^2 , the greater the disparateness between actual and expected frequencies.

Using (10.186) with the data of Table 10.5, we obtain

$$\chi^2 = \frac{(70 - 60)^2}{60} + \frac{(30 - 40)^2}{40} + \frac{(50 - 60)^2}{60} + \frac{(50 - 40)^2}{40} = 8.333$$

Next, we must decide whether the value $\chi^2 = 8.333$ is large enough to be considered statistically significant. To make this decision, we look at a χ^2 distribution table where we find that a χ^2 of 3.84 is exceeded in only 5% of occasions, and a χ^2 of 7.83 is exceeded in only 0.5%. Therefore, we can say that the differences are significant even at the 0.5% level. In other words, the value of $\chi^2 = 8.333$ should occur less than once in 500 samples if there is no difference in work experience between male and female applicants. Accordingly, we reject the null hypothesis.

We can also Excel's **CHITEST** function which returns the test for independence. This is illus-

trated with the following example.

Example 10.38

Table 10.6 shows the actual and expected registered democrats, republicans and independents by gender.

Perform the χ^2 test for goodness of fit to find the level of significance.

Solution:

Using (10.186), Page 10–77, with the data of Table 10.6 below, we obtain

TABLE 10.6 Table for Example 10.38.

	A	B	C
1	Actual		
2		Men	Women
3	Democrats	61	72
4	Republicans	24	16
5	Neither	15	12
6			
7	Expected		
8		Men	Women
9	Democrats	52	65
10	Republicans	38	21
11	Neither	10	16

$$\begin{aligned} \chi^2 &= \sum_{i=1}^3 \sum_{j=1}^2 \frac{(A_{ij} - E_{ij})^2}{E_{ij}} \\ &= \frac{(61 - 52)^2}{52} + \frac{(24 - 38)^2}{38} + \frac{(15 - 10)^2}{10} + \frac{(72 - 65)^2}{65} + \frac{(16 - 21)^2}{21} + \frac{(12 - 16)^2}{16} = 12.1599 \end{aligned}$$

Now, we use Excel's CHIDIST(x,df) function

where

x = the χ^2 value

df = degrees of freedom. It is found from

$$df = (r - 1)(c - 1) \tag{10.187}$$

For this example, we found that $\chi^2 = 12.1599$. Also, from (10.187),

$$df = (r - 1)(c - 1) = (3 - 1)(2 - 1) = 2$$


Then,

=CHIDIST(12.1599,2) returns 0.002288 .

The same answer is obtained with the CHITEST(actual_range,expected_range) function. Then, with reference to Table 10.6,

=CHITEST(B3:C5,B9:C11) returns 0.002288 also.

From a χ^2 distribution table, we find that for $df = 2$ and $\chi^2 = 10.6$ this value is exceeded in only 0.5%. Therefore, we can say that the differences are significant even at the 0.5% level. In other words, the value of $\chi^2 = 12.1599$ should occur less than once in 500 samples if there is to be no difference between the actual and expected frequencies.



10.23 Summary

- The relation

$$\begin{aligned}(a + b)^n &= a^n + na^{n-1}b + \frac{n(n-1)}{2!}a^{n-2}b^2 \\ &\quad + \frac{n(n-1)(n-2)}{3!}a^{n-3}b^3 + \dots + ab^{n-1} + b^n \\ &= \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k\end{aligned}$$

is known as the *binomial expansion*, *binomial theorem*, or *binomial series*.

- The *Binomial (Bernoulli) distribution* is defined as

$$b(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k}$$

for $k = 0, 1, 2, \dots, n$ and $0 < p < 1$

- The binomial distribution $b(k;n;p)$ is a pdf of the *discrete type*.
- The mean μ and the variance σ^2 for the binomial distribution are:

$$\mu = np$$

and

$$\sigma^2 = npq$$

where n = number of samples, p = probability of success, and q = probability of failure.

- The *rv* X is said to be *uniformly distributed* in the interval $a \leq x \leq b$ if

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

- The cdf of the uniform distribution for the entire interval $0 \leq x \leq \infty$ is

$$P_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

- The mean μ and variance σ^2 of the uniform distribution are

$$\mu = \frac{a+b}{2}$$

and

$$\sigma^2 = \frac{(b-a)^2}{12}$$

- The pdf of the *exponential distribution* is defined as

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

- The mean μ and the variance σ^2 of the exponential distribution are

$$\mu = 1/\alpha$$

and

$$\sigma^2 = 1/\alpha^2$$

- The *normal distribution* is the most widely used distribution in scientific and engineering applications, demographic studies, poll results, and business applications. The pdf of the normal distribution is defined as

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-(x-\mu)^2/2\sigma_X^2}$$

where μ = mean and σ_X = standard deviation.

- The cdf of the normal distribution is

$$P_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^x e^{-(x-\mu)^2/2\sigma_X^2} dx$$

- The normal distribution is usually expressed in *standardized form*, where the standardized rv Z is referred to as the standardized rv X and, in this case, Z is related to X as

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

Then, the mean of the rv Z is zero, that is, $\mu = 0$, and the variance is one, that is, $\sigma_X^2 = 1$. The pdf of the standardized normal distribution is then simplified to

$$f_z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

and the corresponding cdf becomes

$$P_z(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-u^2/2} du = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^z e^{-u^2/2} du$$

where u is a dummy variable of integration.

- The *error function* or *probability integral* $\text{erf}(u)$ defined as

$$\text{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-\tau^2} d\tau$$

where τ is a dummy variable. It can also be expressed in series form as

$$\text{erf}(u) = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n u^{2n+1}}{n!(2n+1)}$$

- The *complementary error function* $\text{erfc}(u)$ is defined as

$$\text{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^{\infty} e^{-\tau^2} d\tau = 1 - \text{erf}(u)$$

- The cdf $P_X(x)$ of the normal distribution is often expressed in terms of the error function as

$$P_X(x) = \frac{1}{2} \left[1 + \text{erf} \left(\frac{x - \mu}{\sqrt{2}\sigma_X} \right) \right]$$

- The $Q(\alpha)$ *function* is defined as

$$Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{-\gamma^2/2} d\gamma$$

- Percentiles split the data into 100 parts. Typical percentile values are 0.05 or 5%, 0.10 or 10%,..... 0.95 or 95%, and 0.99 or 99% and these are denoted as $p_{0.05}$, $p_{0.10}$, ..., $p_{0.95}$, or $p_{0.99}$ respectively.
- For continuous functions, the j th percentile p_j is defined as

$$p_j = P[X \leq x_p] = \int_{-\infty}^{x_p} p(x) dx = \frac{p}{100}$$

- For a pdf of the discrete type, we can obtain *approximate values* of percentiles, if we first arrange the sample in ascending order of magnitude; then, we compute the j th percentile p_j , $j = 1, 2, \dots, 99$ from the relation $i = jn/100$ where $n =$ number of samples. If i turns out to be an integer, the j th percentile is the average of the observations in positions i and $i + 1$ in the sampled data set starting at the bottom (lowest value) of the distribution. If i is not an integer, we select the next integer which is greater than i .
- The *Student's t -distribution* or simply *t -distribution* is used instead of the standard normal distribution when the sample size is less than 30. The t -distribution is a family of probability distributions, and each separate t -distribution is determined by a parameter called the *degrees of freedom*. If n is the sample size for $n < 30$, then the degrees of freedom can be computed from $df = n - 1$
- The t -distribution curves, like the standard normal distribution curve, are centered at zero. Thus, the mean is zero, that is, $\mu = 0$
- The variance σ^2 of each t -distribution depends upon the value of df . It is found from

$$\sigma^2 = \frac{df}{df - 2} \quad df \geq 3$$

- The pdf of the t -distribution is given in terms of the Γ (gamma) function. The Γ function is discussed in Appendix C.
- Percentile values of the t -distribution for different degrees of freedom are denoted as t_p and can be found in math tables. The tables give values for the number of degrees of freedom 1, 2, ..., 30, 40, 60, 120, ∞ and for $P_X(x) = 0.60, 0.75, 0.90, 0.95, 0.975, 0.99, 0.995$ and 0.9995 .
- The t -distribution is symmetrical; therefore, $t_{1-p} = -t_p$
- The χ^2 (chi-square) distribution, like the t -distribution, is a family of probability distributions, and each separate distribution is determined by the *degrees of freedom*.
- The pdf of the χ^2 distribution is given in terms of the Γ (gamma) function which is discussed in Appendix C.
- The *F distribution* is another family of curves whose shape differs according to the degrees of freedom. The pdf of the F distribution is given in terms of the Γ (gamma) function which is discussed in Appendix C.
- Percentile values of the F distribution for different degrees of freedom are denoted as F_p and can be found in math tables. These tables are based on the df_1 and df_2 degrees of freedom and

Chapter 10 Common Probability Distributions and Tests of Significance

provide values corresponding to $P_X(x) = 0.90, 0.95, 0.975, 0.99, 0.995,$ and 0.999

- The F distribution is used extensively in the *Analysis of Variance (ANOVA)* which compares the means of several populations. ANOVA is discussed in Chapter 12.
- *Chebyshev's inequality* states that

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

- That is, the probability of X differing from its mean by more than some positive number ε , is less than or equal to the ratio of the variance σ^2 to the square of the number ε .
- The *law of large numbers* states that the probability of the arithmetic mean S_n/n differing from its expected value μ by more than the value of ε approaches zero as $n \rightarrow \infty$.
- The *Poisson distribution* is a pdf of the discrete type, and it is defined as

$$f(x) = P(X = x) = p(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

for $x = 0, 1, 2, \dots$ where λ is a given positive number. The Poisson distribution provides us with a close approximation of the binomial distribution for small values of x provided that the probability of success p is small so that the probability of failure q , that is, $q = 1 - p \approx 1$ and $\lambda = np$ where $n =$ number of trials.

- The *multinomial distribution* is a generalization of the binomial distribution and states that in n repeated trials, the probability that event A_1 occurs k_1 times, event A_2 occurs k_2 times and so on until A_m occurs k_m times, where $k_1 + k_2 + \dots + k_m = n$, is obtained by

$$p = \frac{n!}{k_1! k_2! \dots k_m!} p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}$$

We observe that for $m = 2$, this distribution reduces to the binomial distribution.

- The *hypergeometric distribution* states that

$$P(X = x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} = \frac{\frac{k!}{x!(k-x)!} \cdot \frac{(N-k)!}{(n-x)!(N-k-n+x)!}}{\frac{N!}{n!(N-n)!}}$$

For the hypergeometric distribution, the mean μ and variance σ^2 are

$$\mu = n \frac{k}{N}$$

and

$$\sigma^2 = \left(\frac{N-n}{N-1} \right) n \frac{k}{N} \left(1 - \frac{k}{N} \right)$$

- The *bivariate normal distribution* is a generalization of the normal distribution. It applies to two continuous rv X and Y given by the joint density function

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\left\{ \frac{\left[\left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right]}{2(1-\rho^2)} \right\}}$$

where $\mu_1, \mu_2, \sigma_1, \sigma_2$ are the means and standard deviations of the X and Y respectively, and ρ is the *correlation coefficient* between X and Y. This distribution is also referred to as the *two-dimensional normal distribution*.

- The Rayleigh distribution is a special case of the bivariate normal distribution where the means are zero, the standard deviations are equal, and X and Y are uncorrelated, that is, $\mu_1 = \mu_2 = 0, \sigma_1 = \sigma_2 = \sigma$, and $\rho = 0$. It is defined as

$$f_r(r) = \frac{r}{\sigma^2} e^{-(r^2/2\sigma^2)} \quad \text{for } r \geq 0$$

where $r = \sqrt{x^2 + y^2}$, and σ^2 is the variance of the variables x and y.

- The *Cauchy distribution* has the pdf

$$f_X(x) = \frac{a}{\pi(x^2 + a^2)} \quad a > 0, \quad -\infty < x < \infty$$

This distribution is symmetrical about $x = 0$ and thus its mean is zero. However, the variance and higher moments do not exist.

- The *geometric distribution* has the pdf

$$f_X(x) = pq^{x-1} \quad x = 1, 2, \dots$$

with mean $\mu = 1/p$ and variance $\sigma^2 = q/p^2$

- The *negative binomial distribution* or *Pascal distribution* has the pdf

$$f_X(x) = \binom{x-1}{r-1} p q^{x-r} \quad x = r, r+1, \dots$$

with mean $\mu = r/p$ and variance $\sigma^2 = rq/p^2$

- The *Weibull distribution* has the pdf

$$f_X(x) = \begin{cases} \alpha \beta x^{\beta-1} e^{-\alpha x^\beta} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

with mean $\mu = \alpha^{-1/\beta} \Gamma\left(1 + \frac{1}{\beta}\right)$ and variance $\sigma^2 = \alpha^{-2/\beta} \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \right]$

- The *Maxwell distribution* has the pdf

$$f_X(x) = \begin{cases} \sqrt{\frac{2}{\pi}} a^{3/2} x^2 e^{-ax^2/2} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

with mean $\mu = 2 \sqrt{\frac{2}{\pi a}}$ and variance $\sigma^2 = \frac{1}{a} \left(3 - \frac{8}{\pi} \right)$

- The *lognormal distribution* has the pdf

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{1}{2\sigma^2} \ln(x-\mu)^2} \quad x > 0, \quad \sigma > 0$$

with mean $\mu = e^{(\mu + \sigma^2)/2}$ and variance $\sigma^2 = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$

- The *critbinom* function returns the smallest value for which the binomial distribution is greater than, or equal to a criterion value α (alpha) where $0 < \alpha < 1$.
- The *sampling distribution for the sampled mean*, or simply the *sampled mean*, denoted as $\mu_{\bar{x}}$ is a sample statistic taken from a large number of random samples from the same population.
- A *z score* represents the number of standard deviations that a value x , obtained from a given observation, is above or below the mean. For the entire population, it is defined as

$$z = \frac{x - \mu}{\sigma}$$

- A *null hypothesis* is a statistical decision where we conclude that there is no difference between procedures in sampling from the same population.
- An alternate hypothesis is a hypothesis which differs from a given hypothesis.
- *Tests of significance* are procedures that we follow to decide whether to accept or reject hypotheses.
- If $\mu =$ some known value represents the null hypothesis, the test to determine the alternate hypothesis for $\mu <$ some known value is called a *one-tail left test* or a *lower-tail test*.
- If $\mu =$ some known value represents the null hypothesis, the test to determine the alternate hypothesis for $\mu >$ some known value is called a *one-tail right test* or an *upper-tail test*.
- If the alternate hypothesis is for the condition $\mu \neq$ some known value, regardless whether $\mu <$ some known value or $\mu >$ some known value, this test is referred to as a *two-tail test*.
- A *test statistic*, denoted as \bar{x} , is a statistical measure that is used to decide whether to accept or reject the null hypothesis.
- A *Type I error* occurs if a hypothesis is rejected when it should be accepted.
- A *Type II error* occurs if a hypothesis is accepted when it should be rejected.
- The probabilities of Type I and Type II errors are denoted as follows:

$\alpha =$ probability of making a Type I error

$\beta =$ probability of making a Type II error

- The probability α is referred to as the *level of significance*. Typical levels of significance of $\alpha = 0.05$ or $\alpha = 0.01$.
- If a sampling distribution of a test statistic \bar{x} , is approximately normally distributed with a sample size $n \geq 30$, we expect to find \bar{x} in the interval $\mu_{\bar{x}} - \sigma_x$ to $\mu_{\bar{x}} + \sigma_x$ about 68.27% of the time. Likewise, we expect to find \bar{x} in the interval $\mu_{\bar{x}} - 2\sigma_x$ to $\mu_{\bar{x}} + 2\sigma_x$ about 95.45% of the time, and in the interval $\mu_{\bar{x}} - 3\sigma_x$ to $\mu_{\bar{x}} + 3\sigma_x$ about 99.73% of the time. These intervals are referred to as *confidence intervals* for estimating \bar{x} , and their respective percentages are called *confidence limits* or *confidence levels*. Recalling that α is the level of significance (0.01 or 0.05), we find that

$$\text{confidence level} = 1 - \text{level of significance} = 1 - \alpha$$

- The *z-test* is used to find the probability that a particular observation is drawn from a particular population. This test generates a standard score for a value x with respect to the data set or array, and returns the two-tail probability for the normal distribution.

Chapter 10 Common Probability Distributions and Tests of Significance

- The t -test is similar to the z -test but it is used with populations whose samples are small (30 or less). We use this test to determine the probability that two samples are drawn from populations that have the same mean.
- The F -test produces the one-tail probability that the variances in two arrays of data points are not significantly different. This test also determines whether two samples have different variances. The F -test is used in the *Analysis of Variance* (ANOVA). It is discussed in Chapter 12.
- The χ^2 statistic, also known as the χ^2 test for goodness of fit, is used to decide whether the difference between actual frequencies and expected frequencies is large enough to be considered significant. It is computed from

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(A_{ij} - E_{ij})^2}{E_{ij}}$$

where

A_{ij} = actual frequency in the i th row, j th column

E_{ij} = expected frequency in the i th row, j th column

r = number of rows

c = number of columns

If $\chi^2 = 0$, the actual and expected frequencies agree exactly. If $\chi^2 > 0$, they are in disagreement, and the larger the value of χ^2 , the greater the disparateness between actual and expected frequencies.

10.24 Exercises

1. A fair die is tossed 8 times and it is considered to be a success if 2 or 3 appears. Using the binomial distribution, compute:
 - a. the probability that a 2 or 3 occurs exactly three times.
 - b. the probability that a 2 or 3 never occurs.
 - c. the probability that a 2 or 3 occurs at least once.
2. A fair die is tossed 240 times. Using the binomial distribution, compute:
 - a. the expected number of sixes.
 - b. the standard deviation.
3. Using the Poisson distribution, compute:
 - a. $p(2;1)$.
 - b. $p(3;0.5)$
 - c. $p(2;0.7)$
4. For college dropouts, the time between dropping from college and returning back to continue their education for students under the age of 25 years, is normally distributed with mean $\mu = 2.5$ semesters and standard deviation $\sigma = 1.0$ semester. Compute the percentage of those who will be returning to college within two years after dropping out.
5. The average number of revolutions per minute (RPM) for a sample of 500 small motors produced by a manufacturing plant, is normally distributed with mean $\mu = 1,200$ RPM and standard deviation $\sigma = 50$ RPM. The RPM of each motor must have a value in the $1130 \leq \text{RPM} \leq 1230$ range, and all motors outside this range are considered defective, and thus will be rejected. Compute the percentage of the defective motors produced by this plant.
6. A company produces colored light bulbs in percentages of 60% red, 25% blue, and 15% green. Using the multinomial distribution, compute the probability that in a sample of 5 bulbs, 3 are red, 1 is blue, and 1 is green.
7. The table below shows the actual and expected test results for mathematics, physics, and chemistry by male and female students.

Perform the χ^2 test for goodness of fit to find the level of significance.

Actual		
	Male Students	Female Students
Mathematics	72	76
Physics	84	81
Chemistry	79	87
Expected		
	Male Students	Female Students
Mathematics	83	82
Physics	78	79
Chemistry	87	89

8. Compute the Z–score corresponding to an actual score of 145 from a normal distribution whose mean is 100 and standard deviation is 15.
9. The Z–score of a normal distribution with mean 12 and standard deviation 3 is 1.8. Compute the actual score.
10. Use the partial Z–score table below to compute $P(z < 2.45)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936

11. Use the Excel function =NORMSDIST(z) to compute $P(z > 1.45)$.
12. Use the Excel function =NORMSDIST(z) to compute $P(-1.29 < z < 2.18)$.

10.25 Solutions to End-of-Chapter Exercises

1.

a. $n = 8$, $k = 3$, and probability of success each time the die is tossed is $p = 2/6 = 1/3$

$$\begin{aligned} p_{3 \text{ times}} &= \binom{n}{k} p^k (1-p)^{n-k} = \binom{8}{3} \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^5 = \frac{8!}{3!(8-3)!} \cdot \frac{1}{27} \cdot \frac{32}{243} \\ &= \frac{8 \times 7 \times 6}{3 \times 2} \cdot \frac{1}{27} \cdot \frac{32}{243} = \frac{10752}{39366} = \frac{927}{3394} \end{aligned}$$

b. The probability of getting 2 or 3 each time the die is tossed is $p = 2/6 = 1/3$, and with $n = 8$ and $k = 8$ we have

$$p_{\text{never 2 or 3}} = \binom{n}{k} p^k (1-p)^{n-k} = \binom{8}{8} \left(\frac{1}{3}\right)^8 (1-p)^0 = \left(\frac{1}{3}\right)^8 = \frac{1}{6561}$$

c. The probability that a 2 or 3 occurs at least once is 1 – probability of never occurring and thus with the result of part (b)

$$p_{2 \text{ or } 3 \text{ at least once}} = 1 - \frac{1}{6561} = \frac{6560}{6561}$$

2.

a. Here, $n = 240$, $p = 1/6$ and from (10.16), Page 10–5, $\mu = np = 240/6 = 40$ b. From (10.17), Page 10–5, $\sigma^2 = npq$ or

$$\begin{aligned} \sigma &= \sqrt{np(1-p)} = \sqrt{240 \cdot 1/6 \cdot (1-1/6)} = \sqrt{40 \times 5/6} \\ &= \sqrt{200/6} = \sqrt{100/3} = 5.77 \end{aligned}$$

3.

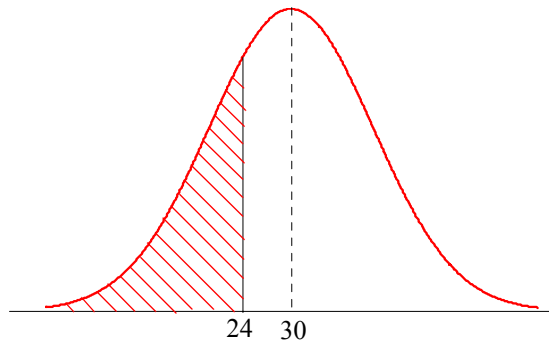
$$p(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$\text{a. } p(2;1) = \frac{1^2 e^{-1}}{2!} = \frac{1}{2e} = 0.184$$

$$\text{b. } p(3;0.5) = \frac{(1/2)^3 e^{-0.5}}{3!} = \frac{e^{-0.5}}{48} = 0.013$$

$$\text{c. } p(2;0.7) = \frac{0.7^2 e^{-0.7}}{2!} = 0.12$$

4. The percentage of those who will be returning within 2 years (24 months) is $P(X \leq 24)$ with mean $\mu = 30$ months and standard deviation $\sigma = 6$ months. This is indicated by the shaded area below.



In terms of normalized units,

$$Z = \frac{X - \mu}{\sigma} = \frac{24 - 30}{6} = -1$$

Now, we refer to the spreadsheet of Figure 10.12, Page 10–17, where under Column E for $Z = -1$ we read 15.94%. That is, there is a probability of approximately 16% that college students will be returning to college within two years after dropping out.

- 5.
- $$Z_1 = \frac{X_1 - \mu}{\sigma} = \frac{1130 - 1200}{50} = -1.4$$
- $$Z_2 = \frac{X_2 - \mu}{\sigma} = \frac{1250 - 1200}{50} = 1.4$$

Now, we refer to the spreadsheet of Figure 10.12, Page 10–17, where under Column I for $Z = 1.4$ we read 83.69%. That is, the area between the interval $-1.0 \leq Z \leq 1.0$ is 0.8369 and the area outside this interval is $1 - 0.8369 = 0.1631$ and thus there is a probability that approximately 16.31% motors will be defective.

- 6.
- $$p = \frac{n!}{k_1! k_2! \dots k_m!} p_1^{k_1} p_2^{k_2} \dots p_m^{k_m} = \frac{5!}{3! \cdot 1! \cdot 1!} \cdot 0.6^3 \cdot 0.25^1 \cdot 0.15^1 = \frac{120}{6} \cdot 0.216 \cdot 0.25 \cdot 0.15 = 0.162$$

Thus, the probability that in a sample of 5 bulbs, 3 are red, 1 is blue, and 1 is green is 16.2%

	A	B	C
1	Actual		
2		Male Students	Female Students
3	Mathematics	72	76
4	Physics	84	81
5	Chemistry	79	87
6			
7	Expected		
8		Male Students	Female Students
9	Mathematics	83	82
10	Physics	78	79
11	Chemistry	87	89

7

$$\begin{aligned}\chi^2 &= \sum_{i=1}^r \sum_{j=1}^c \frac{(A_{ij} - E_{ij})^2}{E_{ij}} = \sum_{i=1}^3 \sum_{j=1}^2 \frac{(A_{ij} - E_{ij})^2}{E_{ij}} \\ &= \frac{(72 - 83)^2}{83} + \frac{(84 - 78)^2}{78} + \frac{(79 - 87)^2}{87} + \frac{(76 - 82)^2}{82} + \frac{(81 - 79)^2}{79} + \frac{(87 - 89)^2}{89} = 3.1896\end{aligned}$$

Now, we use the Excel =CHIDIST(x,df) function where x = the χ^2 value and df = degrees of freedom which is

$$df = (r - 1)(c - 1) = (3 - 1)(2 - 1) = 2$$

Then,

=CHIDIST(3.1896,2) returns 0.203.

The same answer is obtained with the CHITEST(actual_range,expected_range) function.

Then, with reference to the table above,

=CHITEST(B3:C5,B9:C11) returns 0.203 also.

From a χ^2 distribution table, we find that for $df = 2$ and $\chi^2 = 4.61$ this value is exceeded in only $1 - 0.75 = 0.25 = 2.50\%$. Therefore, we can say that the differences are significant at the 2.50% level. In other words, the value of $\chi^2 = 3.1896$ should occur less than 2.5 times in 100 sets of test scores if there is to be no difference between the actual and expected frequencies.

8.

Using (10.169), Page 10–64, we obtain

$$z = \frac{x - \mu}{\sigma} = \frac{145 - 100}{15} = 3$$

Chapter 10 Common Probability Distributions and Tests of Significance

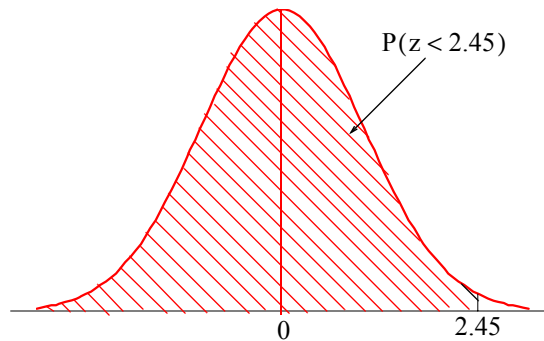
9.

Solving (10.169), Page 10–64, for x we obtain

$$x = z\sigma + \mu = 1.8 \times 3 + 12 = 17.4$$

10.

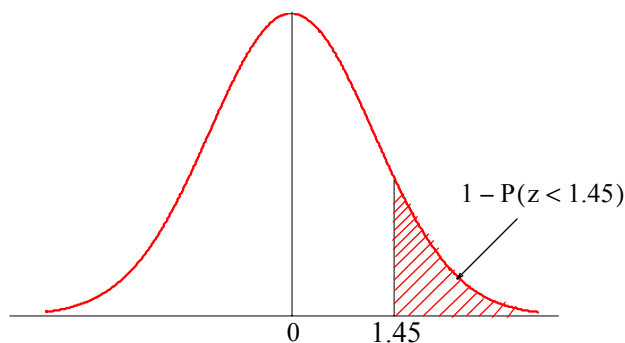
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936



In the table above, the columns correspond to the ones and tenths digits of the z -score and the rows correspond to the hundredths digits. Thus,

$$P(z < 2.45) = 0.9929$$

11.



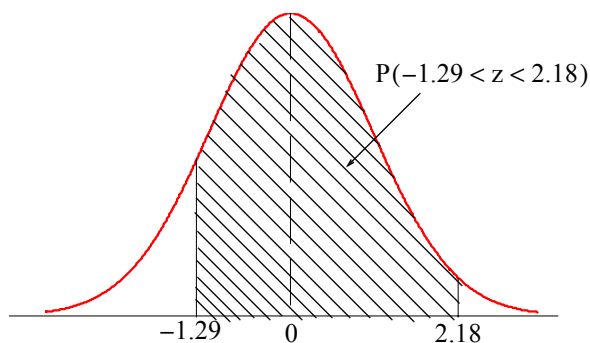
The table gives the area to the left of 1.45 but we want the area to the right of 1.45. Therefore,

$$P(z > 1.45) = 1 - P(z < 1.45)$$

To find the value of $P(z < 1.45)$, we use the Excel function `=NORMSDIST(1.45)` and this returns the value 0.9265. Thus,

$$P(z > 1.45) = 1 - P(z < 1.45) = 1 - 0.9265 = 0.0735$$

12.



From the curve above, we observe that

$$P(-1.29 < z < 2.18) = P(z < 2.18) - P(z < -1.29)$$

Using the Excel function `=NORMSDIST(z)` we obtain

$$=NORMSDIST(2.18) = 0.9854$$

$$=NORMSDIST(-1.29) = 0.0985$$

Therefore,

$$P(-1.29 < z < 2.18) = 0.9854 - 0.0985 = 0.8869$$

Chapter 11

Curve Fitting, Regression, and Correlation

This chapter introduces the concepts of curve fitting, regression, covariance, and correlation, as applied to probability and statistics. Several examples are presented to illustrate their use in practical applications.

11.1 Curve Fitting

Curve fitting is the process of finding equations to approximate straight lines and curves that best fit given sets of data. For example, for the data of Figure 11.1, we can use the equation of a straight line, that is,

$$y = mx + b \quad (11.1)$$

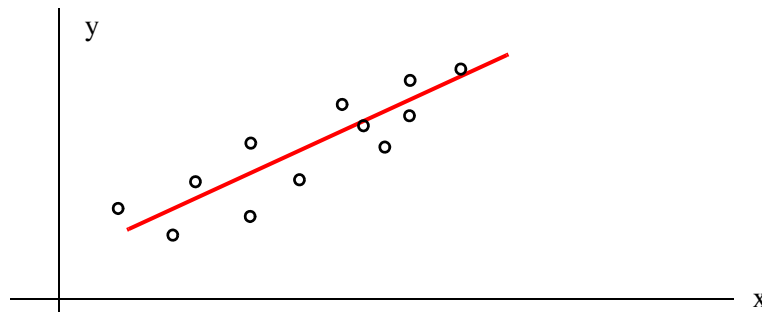


Figure 11.1. Straight line approximation.

For Figure 11.2, we can use the equation for the quadratic or parabolic curve of the form

$$y = ax^2 + bx + c \quad (11.2)$$

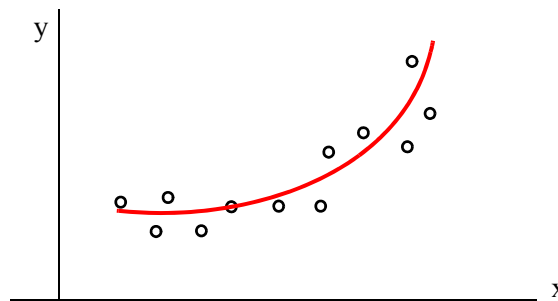


Figure 11.2. Parabolic line approximation

Chapter 11 Curve Fitting, Regression, and Correlation

In finding the best line, we normally assume that the data, shown by the small circles in Figures 11.1 and 11.2, represent the independent variable x , and our task is to find the dependent variable y . This process is called *regression*.

Regression can be linear (straight line) or curved (quadratic, cubic, etc.), and it is not restricted to engineering applications. Investment corporations use regression analysis to compare a portfolio's past performance versus index figures. Financial analysts in large corporations use regression to forecast future costs, and the Census Bureau use it for population forecasting.

Obviously, we can find more than one straight line or curve to fit a set of given data, but we are interested in finding the most suitable.

Let the distance of data point x_1 from the line be denoted as d_1 , the distance of data point x_2 from the same line as d_2 , and so on. The best fitting straight line or curve has the property that

$$d_1^2 + d_2^2 + \dots + d_n^2 = \text{minimum} \quad (11.3)$$

and it is referred to as the *least-squares curve*. Thus, a straight line that satisfies (11.3) is called a *least squares line*. If it is a parabola, we call it a *least-squares parabola*.

11.2 Linear Regression

We perform linear regression with the *method of least squares*. With this method, we compute the coefficients m (slope) and b (y-intercept) of the straight line equation

$$y = mx + b \quad (11.4)$$

such that the sum of the squares of the errors will be minimum. We derive the values of m and b , that will make the equation of the straight line to best fit the observed data, as follows:

Let x and y be two related variables, and let us assume that corresponding to the values $x_1, x_2, x_3, \dots, x_n$, we have observed the values $y_1, y_2, y_3, \dots, y_n$. Now, let us suppose that we have plotted the values of y versus the values of x , and we have observed that the points $(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)$ approximate a straight line. We denote the straight line equations passing through these points as

$$\begin{aligned} y_1 &= mx_1 + b \\ y_2 &= mx_2 + b \\ y_3 &= mx_3 + b \\ &\dots \\ y_n &= mx_n + b \end{aligned} \quad (11.5)$$

In (11.5), the slope m and y -intercept b are the same in all equations since we have assumed that all points lie close to one straight line. However, we need to determine the values of the unknowns m and b from all n equations; we will not obtain valid values for all points if we solve just two equations with two unknowns.*

The *error* (difference) between the observed value y_1 , and the value that lies on the straight line, is $y_1 - (mx_1 + b)$. This difference could be positive or negative, depending on the position of the observed value, and the value at the point on the straight line. Likewise, the error between the observed value y_2 and the value that lies on the straight line is $y_2 - (mx_2 + b)$ and so on. The straight line that we choose must be a straight line such that the distances between the observed values, and the corresponding values on the straight line, will be minimum. This will be achieved if we use the magnitudes (absolute values) of the distances; if we were to combine positive and negative values, some may cancel each other and give us an erroneous sum of the distances. Accordingly, we find the sum of the squared distances between observed points and the points on the straight line. For this reason, this method is referred to as the *method of least squares*.

The values of m and b can be found from

$$\begin{aligned} (\Sigma x^2)m + (\Sigma x)b &= \Sigma xy \\ (\Sigma x)m + nb &= \Sigma y \end{aligned} \tag{11.6}$$

where

Σx = sum of the numbers x

Σy = sum of the numbers y

Σxy = sum of the numbers of the product xy

Σx^2 = sum of the numbers x squared

n = number of data x

We can solve the equations of (11.6) simultaneously by Cramers' rule, or with Excel using matrices.

With Cramer's rule, m and b are computed from

$$m = \frac{D_1}{\Delta} \quad b = \frac{D_2}{\Delta} \tag{11.7}$$

where

* A linear system of independent equations that has more equations than unknowns is said to be **overdetermined** and no exact solution exists. On the contrary, a system that has more unknowns than equations is said to be **underdetermined** and these systems have infinite solutions.

$$\Delta = \begin{vmatrix} \Sigma x^2 & \Sigma x \\ \Sigma x & n \end{vmatrix} \quad D_1 = \begin{vmatrix} \Sigma xy & \Sigma x \\ \Sigma y & n \end{vmatrix} \quad D_2 = \begin{vmatrix} \Sigma x^2 & \Sigma xy \\ \Sigma x & \Sigma y \end{vmatrix} \quad (11.8)$$

Example 11.1

In a typical resistor, the resistance R in Ω (ohms), denoted as y in the equations above, increases with an increase in temperature T in $^{\circ}\text{C}$, denoted as x . The temperature increments and the observed resistance values are shown in Table 11.1. Compute the straight line equation that best fits the observed data.

TABLE 11.1 Data for Example 11.1 - Resistance versus Temperature

T ($^{\circ}\text{C}$)	x	0	10	20	30	40	50	60	70	80	90	100
R (Ω)	y	27.6	31.0	34.0	37	40	42.6	45.5	48.3	51.1	54	56.7

Solution:

Here, we are given 11 sets of data and thus $n = 11$. To save considerable time, we use the spreadsheet of Figure 11.3 where we enter the given values, and we perform the computations using spreadsheet formulas. Accordingly, we enter the x (temperature) values in Column A, and y (the measured resistance corresponding to each temperature value) in Column B. Columns C and D show the x^2 and xy products. Then, we compute the sums so they can be used with (11.7) and (11.8). All work is shown on the spreadsheet of Figure 11.3. The values of m and b are shown in cells I20 and I24 respectively. Thus, the straight line equation that best fits the given data is

$$y = mx + b = 0.288x + 28.123 \quad (11.9)$$

We can use Excel to produce quick answers to regression problems. We will illustrate the procedure with the following example.

Example 11.2

Repeat Example 11.1 using Excel's *Add Trendline* feature.

Solution:

We first enter the given data in columns A and B as shown on the spreadsheet of Figure 11.4. To produce the plot of Figure 11.4, we perform the following steps:

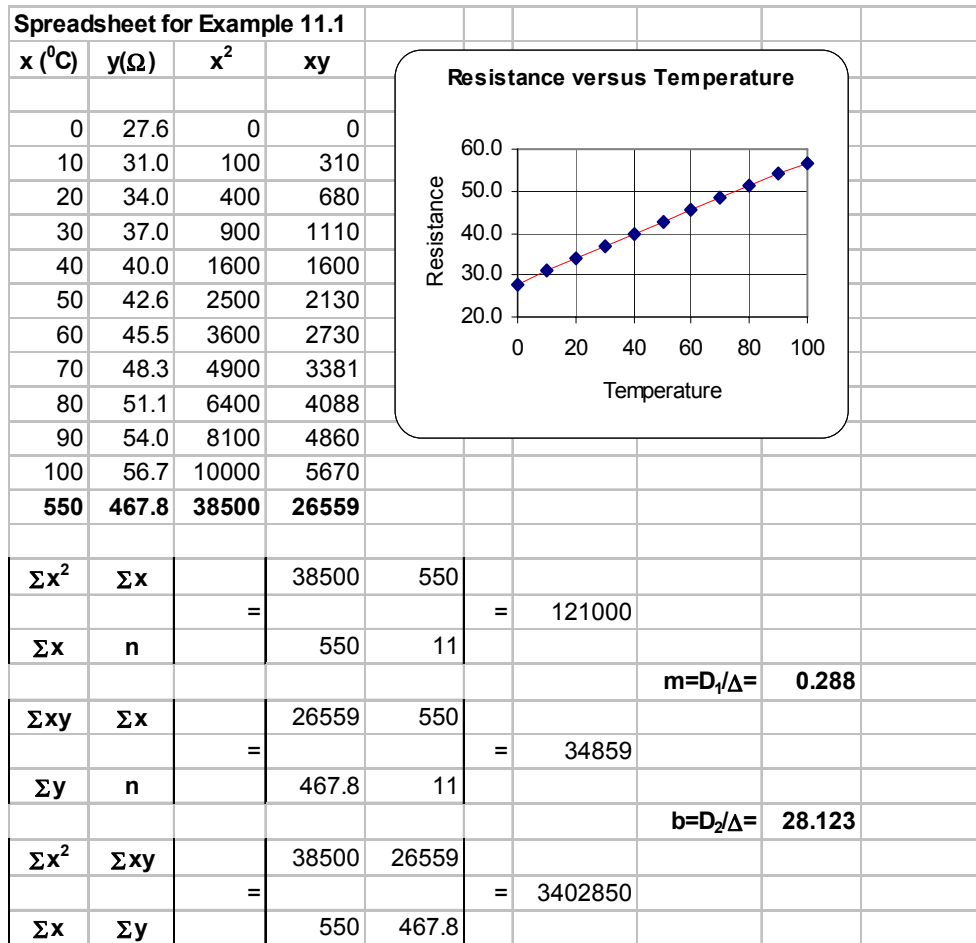


Figure 11.3. Spreadsheet for Example 11.1

1. We click on the *Chart Wizard* icon. The displayed chart types appear on the *Standard Types* tab. We click on *XY (Scatter) Type*. On the *Chart sub-types* options, we click on the top (*scatter*) *sub-type*. Then, we click on *Next>Next>Next>Finish*, and we observe that the plot appears next to the data. We click on the *Series 1 block* inside the *Chart* box, and we press the *Delete* key to delete it.
2. To change the plot area from gray to white, we choose *Plot Area* from the taskbar below the main taskbar, we click on the small (with the hand) box, on the *Patterns* tab we click on the white box (below the selected gray box), and we click on *OK*. We observe now that the plot area is white. Next, we click anywhere on the perimeter of the *Chart* area and observe six square handles (small black squares) around it. We click on *Chart* on the main taskbar, and on the *Gridlines* tab. Under the *Value (Y) axis*, we click on the *Major gridlines* box to deselect it.

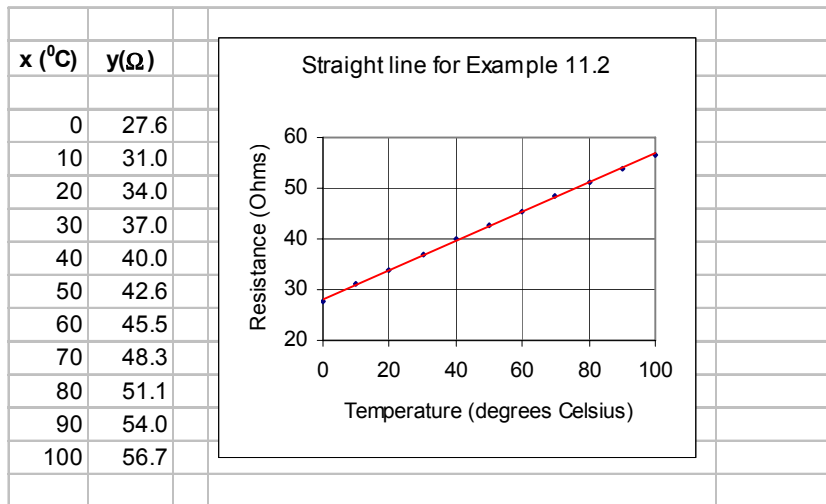


Figure 11.4. Plot of the straight line for Example 11.2

3. We click on the *Titles* tab, and on the *Chart* title box, we type *Straight line for Example 11.2*, on the *Value X-axis*, we type *Temperature, degrees Celsius*, and on the *Value Y-axis*, we type *Resistance in Ohms*. We click anywhere on the x-axis to select it, and we click on the small (with the hand) box. We click on the *Scale* tab, we change the maximum value from 150 to 100, and we click *OK*. We click anywhere on the y-axis to select it, and we click on the small (with the hand) box. We click on the *Scale* tab, we change the minimum value from 0 to 20, we change the *Major Unit* to 10, and we click on *OK*.
4. To make the plot more presentable, we click anywhere on the perimeter of the *Chart* area, and we observe the six handles around it. We place the cursor near the center handle of the upper side of the graph, and when the two-directional arrow appears, we move it upwards by moving the mouse in that direction. We can also stretch (or shrink) the height of the *Chart* area by placing the cursor near the center handle of the lower side of the graph and move it downwards with the mouse. Similarly, we can stretch or shrink the width of the plot to the left or to the right, by placing the cursor near the center handle of the left or right side of the *Chart* area.
5. We click anywhere on the perimeter of the *Chart* area to select it, and we click on *Chart* above the main taskbar. On the pull-down menu, we click on *Add Trendline*. On the *Type* tab, we click on the first (*Linear*) and we click on *OK*. We now observe that the points on the plot have been connected by a straight line.

We can also use Excel to compute and display the equation of the straight line. This feature will be illustrated in Example 11.4

The *Data Analysis Toolpack* in Excel includes the *Regression Analysis* tool which performs linear regression using the least squares method. It provides a wealth of information for statisticians, and contains several terms used in probability and statistics.

11.3 Parabolic Regression

We find the *least-squares parabola* that fits a set of sample points with

$$y = ax^2 + b + c \tag{11.10}$$

where the coefficients a , b , and c are found from

$$\begin{aligned} (\Sigma x^2)a + (\Sigma x)b + nc &= \Sigma y \\ (\Sigma x^3)a + (\Sigma x^2)b + (\Sigma x)c &= \Sigma xy \\ (\Sigma x^4)a + (\Sigma x^3)b + (\Sigma x^2)c &= \Sigma x^2y \end{aligned} \tag{11.11}$$

where n = number of data points.

Example 11.3

Find the least-squares parabola for the data of Table 11.2.

TABLE 11.2 Data for Example 11.3

x	1.2	1.5	1.8	2.6	3.1	4.3	4.9	5.3
y	4.5	5.1	5.8	6.7	7.0	7.3	7.6	7.4
x	5.7	6.4	7.1	7.6	8.6	9.2	9.8	
y	7.2	6.9	6.6	5.1	4.5	3.4	2.7	

Solution:

We construct the spreadsheet of Figure 11.5, and from the data of Columns A and B, we compute the values shown in Columns C through G. The sum values are shown in Row 18, and from these we form the coefficients of the unknowns a , b , and c .

By substitution into (11.11) above,

$$\begin{aligned} 530.15a + 79.1b + 15c &= 87.8 \\ 4004.50a + 530.15b + 79.1c &= 437.72 \\ 32331.49a + 4004.50b + 530.15c &= 2698.37 \end{aligned} \tag{11.12}$$

We solve the equations of (11.12) with matrix inversion and multiplication, as shown in Figure 11.6. The procedure was presented in Chapter 2, Page 2–23. Thus, the least-squares parabola is

	A	B	C	D	E	F	G
1	x	y	x ²	x ³	x ⁴	xy	x ² y
2	1.2	4.5	1.44	1.73	2.07	5.40	6.48
3	1.5	5.1	2.25	3.38	5.06	7.65	11.48
4	1.8	5.8	3.24	5.83	10.50	10.44	18.79
5	2.6	6.7	6.76	17.58	45.70	17.42	45.29
6	3.1	7.0	9.61	29.79	92.35	21.70	67.27
7	4.3	7.3	18.49	79.51	341.88	31.39	134.98
8	4.9	7.6	24.01	117.65	576.48	37.24	182.48
9	5.3	7.4	28.09	148.88	789.05	39.22	207.87
10	5.7	7.2	32.49	185.19	1055.60	41.04	233.93
11	6.4	6.9	40.96	262.14	1677.72	44.16	282.62
12	7.1	6.6	50.41	357.91	2541.17	46.86	332.71
13	7.6	5.1	57.76	438.98	3336.22	38.76	294.58
14	8.6	4.5	73.96	636.06	5470.08	38.70	332.82
15	9.2	3.4	84.64	778.69	7163.93	31.28	287.78
16	9.8	2.7	96.04	941.19	9223.68	26.46	259.31
17	Σx=	Σy=	Σx ² =	Σx ³ =	Σx ⁴ =	Σxy=	Σx ² y=
18	79.1	87.8	530.15	4004.50	32331.49	437.72	2698.37

Figure 11.5. Spreadsheet for Example 11.3

Matrix Inversion and Matrix Multiplication for Example 11.3						
		530.15	79.10	15.00	Σy=	87.80
A=		4004.50	530.15	79.10	Σxy=	437.72
		32331.49	4004.50	530.15	Σx ² y=	2698.37
		0.032	-0.016	0.002	a=	-0.20
A ⁻¹ =		-0.385	0.181	-0.016	b=	1.94
		0.979	-0.385	0.032	c=	2.78

Figure 11.6. Spreadsheet for the solution of the equations of (11.12)

$$y = -0.20x^2 + 1.94x + 2.78$$

The plot for this parabola is shown in Figure 11.7.

Example 11.4

The voltages (volts) shown in Table 11.3 were applied across the terminal of a non-linear device and the current ma (milliamps) values were observed and recorded. Use Excel's *Add Trendline* feature to derive a polynomial that best approximates the given data.

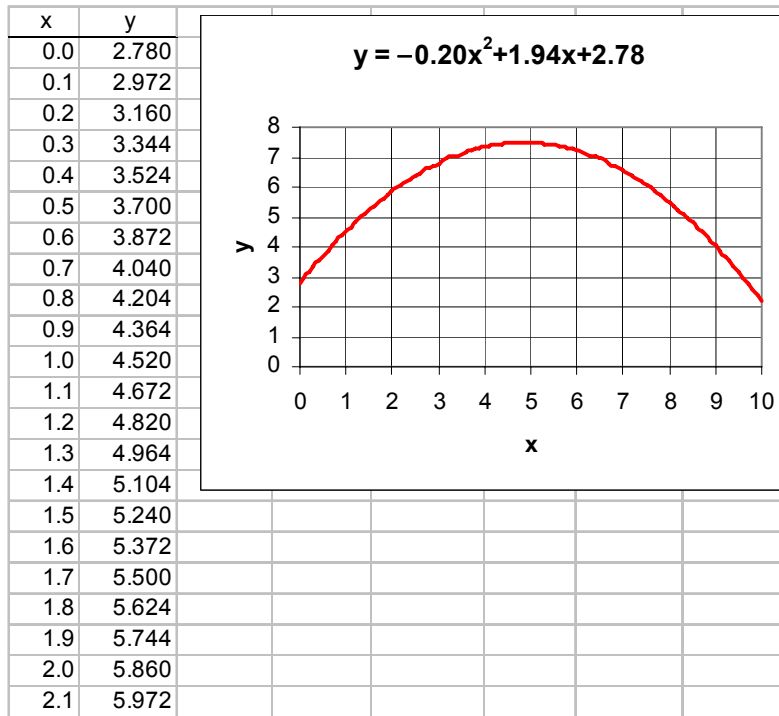


Figure 11.7. Parabola for Example 11.3

TABLE 11.3 Data for Example 11.4

Experimental Data											
Volts	0.00	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50
ma	0.00	0.01	0.03	0.05	0.08	0.11	0.14	0.18	0.23	0.28	0.34
Volts	2.75	3.00	3.25	3.50	3.75	4.00	4.25	4.50	4.75	5.00	
ma	0.42	0.50	0.60	0.72	0.85	1.00	1.18	1.39	1.63	1.91	

Solution:

We enter the given data on the spreadsheet of Figure 11.8 where, for brevity, only a partial list of the given data is shown. However, to obtain the plot, we need to enter all data in Columns A and B.

Following the steps of Example 11.2, we create the plot shown next to the data. Here, the smooth curve is chosen from the *Add trendline* feature, but we click on the *polynomial order 3* on the *Add trendline Type* tab. On the *Options* tab, we click on *Display equation on chart*, we click on *Display R squared value on chart*, and on *OK*. The quantity R^2 is a measure of the goodness of fit for a straight line or, as in this example, for parabolic regression. This is the *Pearson correlation*

Chapter 11 Curve Fitting, Regression, and Correlation

coefficient r and it is discussed in Section 11.5. The correlation coefficient can vary from 0 to 1. When $R^2 \approx 0$, there is no relationship between the dependent y and independent x variables.

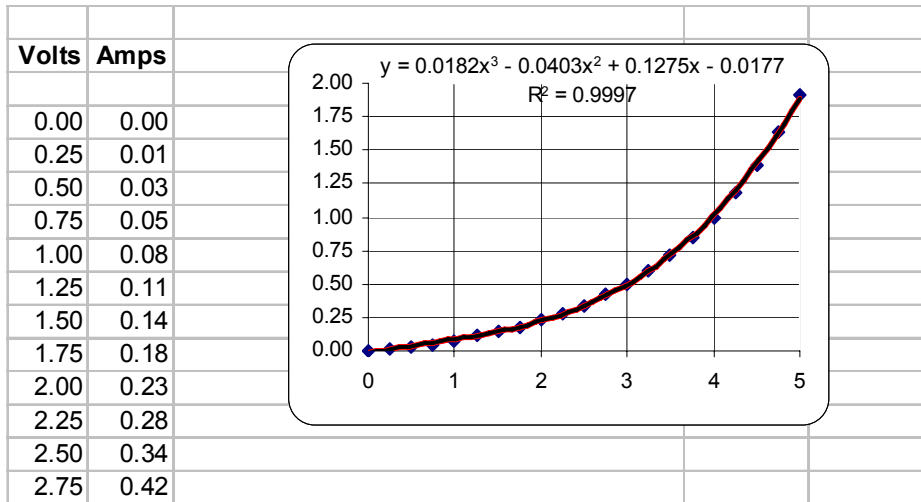


Figure 11.8. Plot for the data of Example 11.4

When $R^2 \approx 1$, there is a nearly perfect linear relationship between these variables. Thus, the result of Example 11.4 indicates that there is a strong linear relationship between the variables x and y , that is, a nearly perfect fit between the cubic polynomial and the experimental data.

11.4 Covariance

In probability and statistics, the *covariance* is a statistical measure of the variance of two random variables that are observed or measured in the same mean time period. This measure is equal to the product of the deviations of corresponding values of the two variables for their respective means.

Let us review the basic statistical averages from Chapter 9.

1. The expected value $E(x)$ of a discrete rv X that has possible values x_1, x_2, \dots, x_n with probabilities P_1, P_2, \dots, P_n respectively, is defined as

$$E(x) = x_1P_1 + x_2P_2 + \dots + x_nP_n = \sum_{i=1}^n x_iP_i \quad (11.13)$$

and if x_1, x_2, \dots, x_n have equal probability of occurrence, then

$$E(x) = \mu = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (11.14)$$

The expected value is referred to as the mean or average of x_1, x_2, \dots, x_n and it is denoted as μ .

2. For a continuous rv X the *expected value* $E(x)$ is defined as

$$E(x) = \int_{-\infty}^{\infty} xf(x)dx \quad (11.15)$$

where $f(x)$ is the probability density function (pdf) of some distribution.

3. The *variance* (Var) of X is defined as

$$\text{Var}(X) = \sigma_X^2 = E[(x - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx \quad (11.16)$$

and the *standard deviation* σ_X is defined as the positive square root of the variance, that is,

$$\sigma_X = \sqrt{\text{Var}(X)} \quad (11.17)$$

In this chapter we introduce the *covariance* $\text{Cov}(X, Y)$ of two random variables X and Y which is defined as

$$\text{Cov}(X, Y) = \sigma_{XY} = E[(x - \mu_X)(y - \mu_Y)] \quad (11.18)$$

If X and Y are independent ($X \neq Y$) random variables, then

$$\text{Cov}(X, Y) = \sigma_{XY} = 0 \quad (11.19)$$

and if they are completely dependent ($X = Y$),

$$\text{Cov}(X, Y) = \sigma_{XY} = \sigma_X \sigma_Y \quad (11.20)$$

In general,

$$|\sigma_{XY}| \leq \sigma_X \sigma_Y \quad (11.21)$$

Excel provides the **COVAR(array1,array2)** function which returns the average of the products of paired deviations. This function uses the relation

$$\text{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y) \quad (11.22)$$

Chapter 11 Curve Fitting, Regression, and Correlation

Example 11.5

The test scores shown in A2:A11 of the spreadsheet of Figure 11.9 are those achieved by college students enrolled in Section A of a particular course, and B2:B11 are those achieved by different college students enrolled in Section B of the same course, at the same college. Compute the covariance for these scores using (11.22), and then, verify your answer with the Excel COVAR(array1,array2) function.

	A	B	C	D	E	F
1	x	y	$x-\mu_x$	$y-\mu_y$	$(x-\mu_x)(y-\mu_y)$	
2	59	70	-15.70	-8.70	136.59	
3	63	69	-11.70	-9.70	113.49	
4	64	76	-10.70	-2.70	28.89	
5	70	79	-4.70	0.30	-1.41	
6	74	76	-0.70	-2.70	1.89	
7	78	80	3.30	1.30	4.29	
8	79	86	4.30	7.30	31.39	
9	82	77	7.30	-1.70	-12.41	
10	86	84	11.30	5.30	59.89	
11	92	90	17.30	11.30	195.49	
12						
13	μ_x	μ_y			$\Sigma(x-\mu_x)(y-\mu_y)=$	558.10
14	74.70	78.70			Cov(X,Y)=	55.81

Figure 11.9. Spreadsheet for Example 11.5

Solution:

The computations are shown on the spreadsheet of Figure 11.9, where the covariance was found to be 55.81. This value represents the average of the products of deviations of corresponding values in the scores of the two sections.

We obtain the same value with the Excel the formula =COVAR(A2:A11,B2:B11)

11.5 Correlation Coefficient

The *population correlation coefficient* ρ is a measure of the dependence of the random variables X and Y . It is defined as

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \quad (11.23)$$

and it exists in the interval

$$-1 \leq \rho \leq 1 \quad (11.24)$$

In other words, the correlation coefficient is a measure of the interdependence of two random variables whose values vary from -1 to $+1$. These values indicate perfect negative correlation at

-1, no correlation at zero, and perfect positive correlation at +1. For instance, the age of an used car and its value would indicate a highly negative correlation, whereas the level of education and salary of an individual would indicate a highly positive correlation. On the other hand, the sales of golf balls and cinema attendances would indicate a zero correlation.

Excel provides the `=CORREL(array1,array2)` function to compute the population correlation coefficient and uses relation (11.23). For example,

`=CORREL({59,63,64,70,74,78,79,82,86,92},{70,69,76,79,79,80,86,77,84,90})` returns 0.877.

Another correlation coefficient is the *linear correlation coefficient* or *Pearson correlation coefficient*, and it is denoted as r . It indicates the linear relationship between two data sets. In other words, it measures the closeness with which the pairs of the data fit a straight line.

Excel provides the `=PEARSON(array1,array2)` function to compute the Pearson correlation coefficient r . It uses the relation

$$r = \frac{n\sum xy - (\sum x)\sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (11.25)$$

For example,

`=PEARSON({59,63,64,70,74,78,79,82,86,92},{70,69,76,79,79,80,86,77,84,90})` returns 0.877.

We observe that this is the same value obtained with the `CORREL` function.

A third type is the *rank correlation coefficient*. Here, the data may be ranked in order of importance, size, etc., and thus it said to use *non-parametric** statistical methods. The rank correlation coefficient is computed by *Spearman's formula*

$$r_{\text{rank}} = 1 - \frac{6\sum d^2}{n(n-1)} \quad (11.26)$$

where:

d = differences between ranks of corresponding data in arrays x and y

n = number of pairs in arrays x and y

Excel provides the function `=RANK(number,ref,order)`

where

number = the number whose rank we want to find

ref = an array numbers or a reference to a list of numbers.

* A quantity, such as a mean, that is calculated from data and describes a population, is said to have been derived by parametric statistical methods.

Chapter 11 Curve Fitting, Regression, and Correlation

order = a number which specifies how to rank a number; if 0 or omitted, it ranks a number as if **ref** were a list sorted in descending order. If **order** = a non-zero value, it ranks a number as if **ref** were a list sorted in ascending order.

Also, the **=RANK(number,ref,order)** function assigns the same rank to duplicate numbers. However, the presence of duplicate numbers affects the ranks of subsequent numbers. For instance, if in a list of integers, the number 15 appears twice and has a rank of 8, then 16 would be assigned a rank of 10, that is, no number would have a rank of 9.

Example 11.6

The test scores shown in A2:A11 of the spreadsheet of Figure 11.10 are those achieved by college students enrolled in Section A of a particular course, and in B2:B11, are those achieved by different college students enrolled in Section B of the same course, at the same college. Compute the rank correlation coefficient using (11.26).

	A	B	C	D	E	F
1	x	y	Rank x	Rank y	d	d ²
2	59	70	1	2	-1	1
3	63	69	2	1	1	1
4	64	76	3	3	0	0
5	70	79	4	6	-2	4
6	74	76	5	3	2	4
7	78	80	6	7	-1	1
8	79	86	7	9	-2	4
9	82	77	8	5	3	9
10	86	84	9	8	1	1
11	92	90	10	10	0	0
12					Σd ² =	25
13		r _{rank} =	0.8485			

Figure 11.10. Spreadsheet for Example 11.6

Solution:

The scores are the same as those of Example 11.5. These scores are listed in A2:A11 and B2:B11. Then, the following entries were made:

C2: **=RANK(A2,\$A\$2:\$A\$11,1)**

D2: **=RANK(B2,\$B\$2:\$B\$11,1)**

These were copied to C3:D11. The values of d, d², and Σd² were then computed, and r_{rank} = 0.8485 was found from (11.26).

In practice, we can construct a correlation matrix to relate two or more items which may complement each other. For instance, the sales manager of a computer store may want to find out which products complement each other most. The procedure is illustrated with the following example.

Example 11.7

The sales manager of a computer store wants to do a sales promotion of a computer system that does not include the monitor, printer, scanner, and an external drive. He has gathered sales data for the past year, and wants to lower the price of the computer system while at the same time, raising the prices for the monitor, printer, scanner, and the external drive. How can he determine the correlation among these five items, so he can price them appropriately?

Solution:

We first enter the data in Rows 1 through 4, and in 11 through 14, as shown on the spreadsheet of Figure 11.11. We also enter the data shown in A5:B9 and in B15:G26.

We use Excel's **CORREL(array1,array2)** function which returns the correlation coefficient between two data sets. We could also use the **PEARSON(array1,array2)** function and get the same answers. Thus, in C5:C9 we type the following entries:

=CORREL(\$C\$15:\$C\$26, C15:C26)

=CORREL(\$D\$15:\$D\$26, C15:C26)

=CORREL(\$E\$15:\$E\$26, C15:C26)

=CORREL(\$F\$15:\$F\$26, C15:C26)

=CORREL(\$G\$15:\$G\$26, C15:C26)

respectively. Then, we highlight C5:C9, we click on *Copy*, we highlight D5:G5, and we click on *Paste*. We observe that the correlation coefficients among these 5 sets appear in C5:G9. The negative values indicate that the sales of some of the items vary inversely, for example as the sales of printers go up, the sales of scanners go down and vice versa.

Chapter 11 Curve Fitting, Regression, and Correlation

	A	B	C	D	E	F	G
1	Correlation Matrix of Sales Data						
2							
3			System	Monitor	Printer	Scanner	Ext. Drive
4			1	2	3	4	5
5	System	1	1.0000	0.9749	0.2739	-0.4990	0.7339
6	Monitor	2	0.9749	1.0000	0.1585	-0.3738	0.7789
7	Printer	3	0.2739	0.1585	1.0000	-0.6657	0.3830
8	Scanner	4	-0.4990	-0.3738	-0.6657	1.0000	-0.3517
9	Ext. Drive	5	0.7339	0.7789	0.3830	-0.3517	1.0000
10							
11	Historical sales data						
12							
13		Mo/Yr	System	Monitor	Printer	Scanner	Ext. Drive
14			1	2	3	4	5
15		Jan-00	256	242	16	14	12
16		Feb-00	249	253	14	17	11
17		Mar-00	254	243	16	19	15
18		Apr-00	250	256	12	22	14
19		May-00	252	244	13	17	10
20		Jun-00	261	227	14	15	9
21		Jul-00	268	228	16	14	11
22		Aug-00	257	242	16	14	16
23		Sep-00	249	246	15	18	13
24		Oct-00	259	253	15	13	17
25		Nov-00	362	335	17	12	18
26		Dec-00	457	445	15	13	21

Figure 11.11. Spreadsheet for Example 11.7

11.6 Summary

- *Curve fitting* is the process of finding equations to approximate straight lines and curves that best fit given sets of data.
- *Regression* is the process of finding the dependent variable y from some data of the independent variable x . Regression can be linear (straight line) or curved (quadratic, cubic, etc.)
- The best fitting straight line or curve has the property that $d_1^2 + d_2^2 + \dots + d_n^2 = \text{minimum}$ and it is referred to as the *least-squares curve*. A straight line that satisfies this property is called a *least squares line*. If it is a parabola, we call it a *least-squares parabola*.
- We perform linear regression with the *method of least squares*. With this method, we compute the coefficients m (slope) and b (y-intercept) of the straight line equation $y = mx + b$ such that the sum of the squares of the errors will be minimum. The values of m and b can be found from the relations

$$\begin{aligned}(\Sigma x^2)m + (\Sigma x)b &= \Sigma xy \\ (\Sigma x)m + nb &= \Sigma y\end{aligned}$$

where

Σx = sum of the numbers x

Σy = sum of the numbers y

Σxy = sum of the numbers of the product xy

Σx^2 = sum of the numbers x squared

n = number of data x

The values of m and b are computed from

$$m = \frac{D_1}{\Delta} \quad b = \frac{D_2}{\Delta}$$

where

$$\Delta = \begin{vmatrix} \Sigma x^2 & \Sigma x \\ \Sigma x & n \end{vmatrix} \quad D_1 = \begin{vmatrix} \Sigma xy & \Sigma x \\ \Sigma y & n \end{vmatrix} \quad D_2 = \begin{vmatrix} \Sigma x^2 & \Sigma xy \\ \Sigma x & \Sigma y \end{vmatrix}$$

- We find the *least-squares parabola* that fits a set of sample points with $y = ax^2 + b + c$ where the coefficients a , b , and c are found from

$$\begin{aligned}(\Sigma x^2)a + (\Sigma x)b + nc &= \Sigma y \\ (\Sigma x^3)a + (\Sigma x^2)b + (\Sigma x)c &= \Sigma xy \\ (\Sigma x^4)a + (\Sigma x^3)b + (\Sigma x^2)c &= \Sigma x^2y\end{aligned}$$

where n = number of data points.

- The *covariance* $\text{Cov}(X, Y)$ is a statistical measure of the variance of two random variables X and Y that are observed or measured in the same mean time period. This measure is equal to the product of the deviations of corresponding values of the two variables for their respective means. In other words,

$$\text{Cov}(X, Y) = \sigma_{XY} = E[(x - \mu_X)(y - \mu_Y)]$$

If X and Y are independent ($X \neq Y$) random variables, then

$$\text{Cov}(X, Y) = \sigma_{XY} = 0$$

and if they are completely dependent ($X = Y$),

$$\text{Cov}(X, Y) = \sigma_{XY} = \sigma_X \sigma_Y$$

In general,

$$|\sigma_{XY}| \leq \sigma_X \sigma_Y$$

- The *population correlation coefficient* ρ is a measure of the dependence of the random variables X and Y . It is defined as

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

and it exists in the interval

$$-1 \leq \rho \leq 1$$

In other words, the correlation coefficient is a measure of the interdependence of two random variables whose values vary from -1 to $+1$. These values indicate perfect negative correlation at -1 , no correlation at zero, and perfect positive correlation at $+1$.

- The *linear correlation coefficient* or *Pearson correlation coefficient*, denoted as r , indicates the linear relationship between two data sets. In other words, it measures the closeness with which the pairs of the data fit a straight line. It uses the relation

$$r = \frac{n\Sigma xy - (\Sigma x)\Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$$

- The *rank correlation coefficient* is used when it is desired to rank the data in order of importance. The rank correlation coefficient is computed by *Spearman's formula*

$$r_{\text{rank}} = 1 - \frac{6\sum d^2}{n(n-1)}$$

where d = differences between ranks of corresponding data in arrays x and y , and n = number of pairs in arrays x and y .

11.7 Exercises

1. Consider a system of equations derived from experimental data whose general form is

$$a_1x + b_1y = c_1$$

$$a_2x + b_2y = c_2$$

$$a_3x + b_3y = c_3$$

...

$$a_nx + b_ny = c_n$$

The values of unknowns x and y can be found from

$$x = \frac{D_1}{\Delta} \quad y = \frac{D_2}{\Delta}$$

where

$$\Delta = \begin{vmatrix} \Sigma a^2 & \Sigma ab \\ \Sigma ab & \Sigma b^2 \end{vmatrix} \quad D_1 = \begin{vmatrix} \Sigma ac & \Sigma ab \\ \Sigma bc & \Sigma b^2 \end{vmatrix} \quad D_2 = \begin{vmatrix} \Sigma a^2 & \Sigma ac \\ \Sigma ab & \Sigma bc \end{vmatrix}$$

Using these relations, find the values of x and y that best fit the following equations:

$$2x + y = -1$$

$$x - 3y = -4$$

$$x + 4y = 3$$

$$3x - 2y = -6$$

$$-x + 2y = 3$$

$$x + 3y = 2$$

2. Measurements made on an electronics device, yielded the following sets of values:

millivolts	100	120	140	160	180	200
milliamps	0.45	0.55	0.60	0.70	0.80	0.85

Use the procedure of Example 11.1 to compute the straight line equation that best fits the given data.

3. Repeat Exercise 2 using Excel's *Trendline* feature.

4. A sales manager wishes to forecast sales for the next three years, for a company that has been in business for the past 15 years. The sales during these years are shown below.

Year	Sales
1	\$9,149,548
2	13,048,745
3	19,147,687
4	28,873,127
5	39,163,784
6	54,545,369
7	72,456,782
8	89,547,216
9	112,642,574
10	130,456,321
11	148,678,983
12	176,453,837
13	207,547,632
14	206,147,352
15	204,456,987

Using Excel's *Trendline* feature, choose an appropriate polynomial to smooth the given data. Then, using the polynomial found, compute the sales for the next three years. You may round the sales to the nearest thousand.

11.8 Solutions to End-of-Chapter Exercises

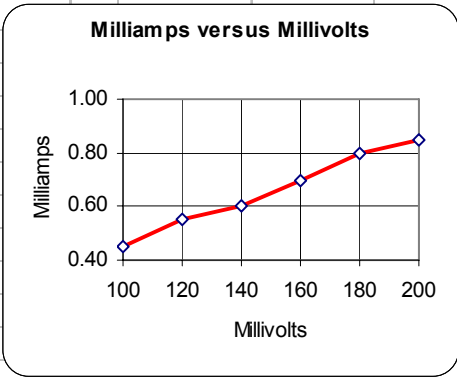
1. We construct the Excel spreadsheet below by entering the given values and computing the values from the formulas given.

Spreadsheet for Exercise 11.1									
	a	b	c	a ²	ab	b ²	ac	bc	
	2	1	-1	4	2	1	-2	-1	
	1	-3	-4	1	-3	9	-4	12	
	1	4	3	1	4	16	3	12	
	3	-2	-6	9	-6	4	-18	12	
	-1	2	3	1	-2	4	-3	6	
	1	3	2	1	3	9	2	6	
Σ	7	5	-3	17	-2	43	-22	47	
	Σa²	Σab	=	17	-2		=	727	
Δ	Σab	Σb²	=	-2	43				
								x=D₁/Δ=	-1.172
	Σac	Σab	=	-22	-2		=	-852	
D₁	Σbc	Σb²	=	47	43				
								y=D₂/Δ=	1.039
	Σa²	Σac	=	17	-22		=	755	
D₂	Σab	Σbc	=	-2	47				

Thus, $x = -1.172$ and $y = 1.039$

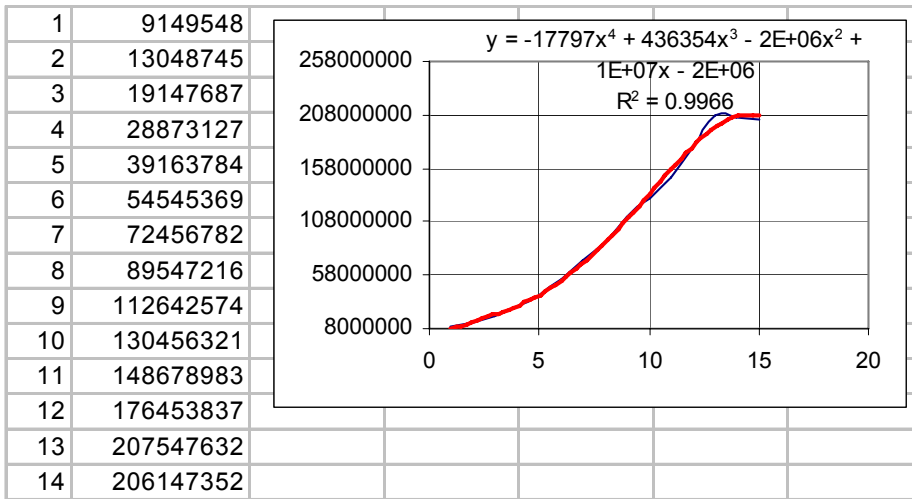
2. We construct the Excel spreadsheet below by entering the given values and computing the values from the formulas given.

Spreadsheet for Exercise 11.2				
x (mV)	y(mA)	x ²	xy	
100	0.45	10000	45	
120	0.55	14400	66	
140	0.60	19600	84	
160	0.70	25600	112	
180	0.80	32400	144	
200	0.85	40000	170	
900	3.95	142000	621	
Σx^2	Σx		142000	900
		=		= 42000
Σx	n		900	6
				$m=D_1/\Delta=$ 0.004
Σxy	Σx		621	900
		=		= 171
Σy	n		4.0	6
				$b=D_2/\Delta=$ 0.0476
Σx^2	Σxy		142000	621
		=		= 2000
Σx	Σy		900	4.0



Thus, $y = mx + b = 0.004x + 0.0476$

4. Following the procedure of Example 11.4, using Excel we choose Polynomial 4 and we get the trendline shown below.



The sales for the next 3 years are:

$$y_{16} = -17797x^4 + 436354x^3 - 2 \times 10^6 x^2 + 10^7 x - 2 \times 10^6 \Big|_{x=16} = 266961792$$

$$y_{17} = -17797x^4 + 436354x^3 - 2 \times 10^6 x^2 + 10^7 x - 2 \times 10^6 \Big|_{x=17} = 247383965$$

$$y_{18} = -17797x^4 + 436354x^3 - 2 \times 10^6 x^2 + 10^7 x - 2 \times 10^6 \Big|_{x=18} = 206558656$$

These results indicate that the trendline sets more weight on the last four sales, i.e., years 12, 13, 14, and 15, and hence the jump on the year 16 and declines the years 17 and 18.

Chapter 12

Analysis of Variance (ANOVA)

This chapter introduces the Analysis of Variance, briefly known as ANOVA. Both one-way and two-way ANOVA are discussed, and are illustrated with practical examples.

12.1 Introduction

The analysis of variance, henceforth referred to as ANOVA, is the analysis of the variation in the outcomes of an experiment to assess the contribution of each variable to the total variation. ANOVA provides answers to questions such as “In a manufacturing plant, is the day shift really producing units with fewer defects than the swing and graveyard shifts?” Or, “Is one age group more productive than others in a department store?” ANOVA makes use of the F distribution discussed in Chapter 10. It allows us to compare several samples with a single test. It was developed by the British mathematician R. A. Fisher.

We use ANOVA in situations where we have a number of observed values, and these are divided among three or more groups. We want to know whether all these values belong to the same population, regardless of group, or we want to find out if the observations in at least one of the groups come from a different population. This determination is made by comparing the variation of values *within* groups and the variation of values *between* groups. The *variation within groups* is computed from the variation within groups of samples, whereas the *variation between groups* is computed from the variation among different groups of samples.

The *F ratio* is the variance between groups divided by the variance within groups, that is,

$$F \text{ ratio} = \frac{\text{Variance between groups}}{\text{Variance within groups}} \quad (12.1)$$

Once the F ratio is found, we compare it with a value obtained from an F distribution table, to determine whether the samples are, or are not statistically different. The procedure will be illustrated with examples.

We will discuss only one- and two-way ANOVA although three-, four-way ANOVA and so on, are also possible.

12.2 One-way ANOVA

We will illustrate one-way ANOVA with the following example.

Example 12.1

Suppose that the same test was given to five different groups of students at five different universities denoted as A, B, C, D, and E. Their scores are shown in Table 12.1.

TABLE 12.1 Scores achieved by students at five universities

A	B	C	D	E
92	91	87	93	92
94	93	88	94	91
91	92	87	90	94
93	89	90	89	93
94	90	89	91	95
92	89	91		94
	91			

Determine whether or not the statistical averages are significant at the 1% level of confidence.

Solution:

We observe that the groups consist of unequal number of students. For instance, we see that in University B, there is a group of seven students, but the group at University D has only five students. This is not a problem with one-way ANOVA, that is, the sample sizes need not be equal for all groups of samples.

From Table 12.1 we might get tempted to conclude that the group of students at University E achieved the highest scores. However, it is possible that observed differences between the mean scores of these groups are not statistically different. To find out, we will perform a one-way ANOVA. The results will reveal whether the differences among the average scores of the groups of students are significant. If the differences of the averages are not statistically significant, we will conclude that the performance of all groups of students is about the same. However, if we find that the differences of the averages are statistically significant, we will conclude that the performance of all five groups of students is not the same. Then, we need to find which group of students performed best.

The given data and the computations are shown on the spreadsheet of Figure 12.1. For this example, we want to find out with 99% certainty that there are, or are not significant differences among the groups of students.

In the spreadsheet of Figure 12.1, Rows 1 through 11 contain the given data, and range H5:H11 shows the sum of the squares.

Formulas for the cells that contain results of computations are given below.

	A	B	C	D	E	F	G	H	
1	One-way ANOVA for the scores of five different groups of students								
2							GRAND	SUM OF	
3		A	B	C	D	E	TOTALS	SQUARES	
4									
5	Scores	92	91	87	93	92		41427	
6		94	93	88	94	91		42346	
7		91	92	87	90	94		41250	
8		93	89	90	89	93		41240	
9		94	90	89	91	95		42163	
10		92	89	91		94		33502	
11			91					8281	
12									
13									
14	Totals	556	635	532	457	559	2739	250209	
15									
16	No. of samples	6	7	6	5	6	30		
17									
18	Averages	92.67	90.71	88.67	91.40	93.17	91.30		
19									
20	No. of Groups		5						
21									
22	ANOVA Table								
23									
24		Sum of	Deg. Of						
25		Squares	Freedom	Variance	F-ratio				
26									
27	Between								
28	Groups	76.17	4	19.04	7.66				
29									
30	Within								
31	Groups	62.13	25	2.49					
32									
33	Totals	138.30	29						
34									
35	F-table Lookup num. 4 denom. 25				4.18				
36	at 1% (0.01)								
37									
38	CONCLUSION: Averages are significantly different at 1% level of confidence								
39									
40	Probability that averages are not significantly different is						0.04%		

Figure 12.1. Spreadsheet for one-way ANOVA of Example 12.1

H5: =B5^2+C5^2+D5^2+E5^2+F5^2 and it is copied to H6:H11

B14: =SUM(B4:B13) and it is copied to C14:H14

The COUNT function counts the number of cells that contain numbers in a range. It does not count blank cells, or cells that contain labels. It is very useful when a spreadsheet contains long ranges of numbers, labels, and blank cells. For this example,

B16: =COUNT(B5:B12) and it is copied to C16:F16

G16: =SUM(B16:F16)

Chapter 12 Analysis of Variance (ANOVA)

The averages are shown in B18:G18.

B18: =B14/B16 and it is copied to C18:G18

C20: =COUNT(B5:F5). This counts the number of groups. For this example, it is 5.

The *sum of squares between groups* ss_{bg} is computed from

$$ss_{bg} = \sum \frac{x^2}{n_i} - \frac{(\sum x)^2}{n} \quad (12.2)$$

where:

x = sum of each group. For this example, it is range B14:F14.

n_i = number of samples in i th group. For this example, it is range B16:F16.

n = the total number of samples in all groups. For this example, it is cell G16.

We enter the formula of (12.2) in B28.

B28: =B14^2/B16+C14^2/C16+D14^2/D16+E14^2/E16+F14^2/F16-G14^2/G16

The *degrees of freedom between groups* df_{bg} is the number of groups minus one, that is,

$$df_{bg} = \text{number of groups} - 1 = 5 - 1 = 4 \quad (12.3)$$

For this example, the number of groups is shown in C20, and df_{bg} in C28.

C28: =C20-1

The *variance between groups* σ_{bg}^2 is obtained from

$$\sigma_{bg}^2 = \frac{\text{Sum of Squares Between Groups}}{\text{Degrees of Freedom Between Groups}} = \frac{ss_{bg}}{df_{bg}} \quad (12.4)$$

σ_{bg}^2 is shown in D28.

D28: =B28/C28

The *sum of squares within groups* ss_{wg} is computed from

$$ss_{wg} = \sum x^2 - \frac{(\sum x)^2}{n} - ss_{bg} \quad (12.5)$$

where:

x = the value of each sample in all groups. For this example, x represents the values in range A5:F11.

n = number of samples in all groups. For this example, it is the number of non-zero entries in range A5:F11.

ss_{bg} = sum of squares between groups as defined in (12.2). For this example, it is cell B28.

ss_{wg} is shown in B31.

$$B31: =H14-G14^2/G16-B28$$

The *degrees of freedom within groups* df_{wg} is the total number of samples n in all groups minus the number of groups, that is,

$$df_{wg} = n - \text{number of groups} = 30 - 5 = 25 \quad (12.6)$$

df_{wg} appears in C31.

$$C31: =G16-C20$$

The *variance within groups* σ_{wg}^2 is obtained from

$$\sigma_{wg}^2 = \frac{\text{Sum of Squares Within Groups}}{\text{Degrees of Freedom Within Groups}} = \frac{ss_{wg}}{df_{wg}} \quad (12.7)$$

σ_{wg}^2 is shown in D31.

$$D31: =B31/C31$$

The *F ratio* is obtained from

$$F \text{ ratio} = \frac{\text{Variance Between Groups}}{\text{Variance Within Groups}} = \frac{\sigma_{bg}^2}{\sigma_{wg}^2} \quad (12.8)$$

The *F ratio* is shown in E28.

$$E28: =D28/D31.$$

B33 shows the sum of the squares of the between groups and within groups, and C33 shows the sum of the degrees of freedom.

B33: =B28+B31 and it is copied to C33. These values are not required in subsequent calculations.

Cell E35 shows the *F distribution value*. We can obtain this value from an *F-distribution table*. This value must correspond to the degree of freedom values of $df_1 = df_{bg} = 4$ and

Chapter 12 Analysis of Variance (ANOVA)

$df_2 = df_{wg} = 25$ for 1% level of confidence. The table shows the value of 4.18. For convenience, we can use Excel's FINV function which is discussed in Chapter 10, Page 10–45. Thus,

E35: =FINV(0.01,C28,C31)

and this returns 4.18. This is the same value that we find in an F–distribution table.

Therefore, with an F ratio ≥ 4.18 , $df_1 = df_{bg} = 4$ and $df_2 = df_{wg} = 25$, we can say with 1% probability that there is no significant difference between the scores of the groups of students. Stated in other words, we conclude with 99% probability that there are significant differences among the scores of the groups of students.

The entry in A38 makes a comparison of the values of E28 and E35 and displays the appropriate message.

A38:="CONCLUSION: Averages are "&IF(E28>E35,"","not ")&"significantly different at 1% level of confidence."

To find out what the probability is that the scores are not significantly different, we use Excel's FDIST function. We use the formula

G40: =FDIST(7.66,4,25) where 7.66 is the F ratio, 4 is the degrees of freedom for between groups, and 25 is the degrees of freedom for within groups. This returns 0.0004 or 0.04%

Since we have determined that the results of the ANOVA indicate variation in test scores, we may want to find out whether there is a significant difference between the test scores in groups A and E, since these indicate a very close resemblance. Accordingly, we erase the test scores for groups B, C and D. Doing this, we find that the degrees of freedom between groups is 1, and the degrees of freedom within groups is 10. The F ratio for the two groups is 0.41 and from an F–distribution table, or with the FINV function, we find the value 10.04. Since 0.41 is less than 10.04, we can say with 99% certainty that, statistically, there is no difference between the test scores of student groups A and E.

Figure 12.2 shows the revised spreadsheet that compares groups A and E. This was obtained from the spreadsheet of Figure 12.1, where we erased the contents of C5:E12, we entered zeros in C14:E14, we changed the contents of B28 to =B14^2/B16+F14^2/F16–G14^2/G16, and we changed the value in E35 to 10.04.

The Microsoft Excel package contains the *Data Analysis Toolpack* that provides three ANOVA analysis tools, the *Single Factor Analysis Tool*, the *Two–Factor with Replication Analysis Tool*, and the *Two–Factor without Replication Analysis Tool*. The first applies to one–way ANOVA and the others to two–way ANOVA.

	A	B	C	D	E	F	G	H	
1	One-way ANOVA for the scores five different groups of students								
2							GRAND	SUM OF	
3		A	B	C	D	E	TOTALS	SQUARES	
4									
5	Scores	92				92		16928	
6		94				91		17117	
7		91				94		17117	
8		93				93		17298	
9		94				95		17861	
10		92				94		17300	
11								0	
12									
13									
14	Totals	556	0	0	0	559	1115	103621	
15									
16	No. of samples	6				6	12		
17									
18	Averages	92.67				93.17	92.92		
19									
20	No. of Groups		2						
21									
22	ANOVA Table								
23									
24		Sum of	Deg. Of						
25		Squares	Freedom	Variance	F-ratio				
26									
27	Between								
28	Groups	0.75	1	0.75	0.41				
29									
30	Within								
31	Groups	18.17	10	1.82					
32									
33	Totals	18.92	11						
34									
35	F-table Lookup num. 1 denom. 10				10.04				
36	at 1% (0.01)								
37									
38	CONCLUSION: Averages are not significantly different								

Figure 12.2. Revised spreadsheet for one-way ANOVA.

The Single Factor Analysis Tool (one-way ANOVA) computes and displays the averages, variances, degrees of freedom, the F ratio, and the F value shown on tables. Excel denotes the F value as *F critical*.

To use the Single Factor Analysis Toolpack* we perform the following steps:

1. We invoke the file that has the spreadsheet of Figure 12.1. We give this file another name, we click on *Tools>Data Analysis>Anova: Single Factor>OK*.

* We should make sure that the Data Analysis Toolpack has been installed; if not, we must click on Help to get instructions on how to install it.

Chapter 12 Analysis of Variance (ANOVA)

2. For the Input Range, we type B5:F11, we change the value of Alpha to 0.01, and for the Output Range, we type I1. Then, Excel displays the data shown in Figure 12.3.

	I	J	K	L	M	N	O
1	Anova: Single Factor						
2							
3	SUMMARY						
4	<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>		
5	Column 1	6	556	92.67	1.47		
6	Column 2	7	635	90.71	2.24		
7	Column 3	6	532	88.67	2.67		
8	Column 4	5	457	91.40	4.30		
9	Column 5	6	559	93.17	2.17		
10							
11							
12	ANOVA						
13	<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
14	Between Groups	76.17	4	19.04	7.66	0.0004	4.18
15	Within Groups	62.13	25	2.49			
16							
17	Total	138.3	29				
18							

Figure 12.3. One-way ANOVA using Excel's Single Factor analysis.

For simplicity, the values have been rounded to two decimal places. We observe that these values are the same as the corresponding values on the spreadsheet of Figure 12.1. The *P-value* indicates that with 0.04% certainty, we can say that the averages are not significantly different, or we can state with 99.96% certainty, that the averages are significantly different.

12.3 Two-way ANOVA

In one-way ANOVA, there is only one variable to consider. Thus, in Example 1 only the average scores of different groups of students were considered. In two-way ANOVA we consider two variables. A two-way ANOVA can be *two-factor without replication*, or a *two-factor with replication*. They are discussed in Subsections 12.3.1 and 12.3.2. An example for each is presented.

12.3.1 Two-factor without Replication ANOVA

This method performs a two-way ANOVA that does not include more than one sampling per group. With this method, we must have the same number of samples within each group. We will illustrate this method with the following example.

Example 12.2

The test scores for five groups of students in three universities are shown in Table 12.2.

TABLE 12.2 Test scores for five groups of students in three universities.

	Student Scores in Each Group				
	Group 1	Group 2	Group 3	Group 4	Group 5
University A	92	91	87	93	95
University B	84	89	81	90	87
University C	88	92	85	89	86

Perform a test at the 0.01 level of significance to determine whether

- a. there is a significant difference in the averages of scores at each university
- b. there is a significant difference in the averages of scores in each group

Solution:

Table 12.2 shows only one sample (test score) per group in each of the three universities. Here, we consider the groups of students as one factor, and the universities as the other factor.

As before, we will use a spreadsheet for the computations. This is shown in Figure 12.4. The formulas can be found in probability and statistics textbooks.

We enter the given scores in range B4:F6. We compute the row sums and row means with the **SUM** and **AVERAGE** functions; we enter these in G4:H6. Using the same functions, we enter the column sums in B8:F8, and the column means in B11:F11.

Using the **SUM** function, in G9 we enter the *grand sum* of the column sums B8:F8. Using the **AVERAGE** function, in G12 we enter the *grand mean*; this is the mean of the column means. We denote the number of rows as r , and the number of columns as c . We use the **COUNT** function to compute the number of rows (universities), and the number of columns (groups). Thus, in B14 we type **=COUNT(B4:B6)**, and in B18 we type **=COUNT(B4:F4)**.

The *variation between rows*, denoted as v_r , is defined as

$$v_r = c \sum_{i=1}^r (\bar{x}_i - \bar{\bar{x}})^2 \tag{12.9}$$

where:

c = number of columns. For this example, it is the number of columns in B through F, that is, 5. It appears in B18.

r = number of rows. For this example, it is the number of Rows 4 through 6, that is, 3. It appears in B14.

Chapter 12 Analysis of Variance (ANOVA)

	A	B	C	D	E	F	G	H	
1	ANOVA: Two-factor without replication								
2			Groups of Students				Row	Row	
3		1	2	3	4	5	Sum	Means	
4	University A	92	91	87	93	95	458	91.6	
5	University B	84	89	81	90	87	431	86.2	
6	University C	88	92	85	89	86	440	88.0	
7									
8	Column Sums	264	272	253	272	268			
9	Grand Sum (Sum of Column Sums)						1329		
10									
11	Column Means	88.00	90.67	84.33	90.67	89.33			
12	Grand Mean (Mean of Column means)						88.60		
13									
14	Number of rows $r =$	3							
15	Variation of Row Means from Mean of all Columns								
16	(H4, H5, and H6 minus G12 squared) $vr =$	75.60							
17									
18	Number of columns $c =$	5							
19	Variation of Column Means from Mean of all Columns								
20	(B11, C11, D11, E11, and F11 minus G12 squared) $vc =$	82.93							
21									
22	Total variation								
23	Each value of B4 through F6 minus G12 squared) $vt =$	195.6							
24									
25	Error variation, ve								
26	$ve = vt - vr - vc$ $ve =$	37.07							
27									
28	Degrees of freedom for rows $dfr = r - 1$ $dfr =$	2							
29	Degrees of freedom for columns $dfc = c - 1$ $dfc =$	4							
30	Degrees of freedom for error $dfev = dfr * dfc$ $dfev =$	8							
31	Deg. of freedom for total variation $dfr + dfc + dfev = dftv =$	14							
32									
33	Mean square for vr $msvr = vr / (r - 1)$ $msvr =$	37.80							
34	Mean square for vc $msvc = vc / (c - 1)$ $msvc =$	20.73							
35	Mean square for ve $msve = ve / [(r - 1) * (c - 1)]$ $msve =$	4.63							
36	F ratio for $vr = msvr / msve$ $msvr / msve$ F ratio =	8.16							
37	F ratio for $vc = msvc / msve$ $msvc / msve$ F ratio =	4.47							
38				Mean	F	From F tables			
39	Variations Summary		df	Square	ratio	or with FINV			
40									
41		$vr =$	75.60	2	37.80	8.16	8.65		
42		$vc =$	82.93	4	20.73	4.47	7.01		
43		$ve =$	37.07	8	4.63				
44		$vt = vr + vc + ve =$	195.60						
45									
46	CONCLUSION 1: At 0.01 level of significance, the averages are not significantly different among universities								
47	CONCLUSION 2: At 0.01 level of significance, the averages are not significantly different among groups								

Figure 12.4. Two-way ANOVA without replication.

\bar{x}_i = the mean of the i th row. For this example i varies from 1 to 3, that is, H4 through H6.

\bar{x} = grand mean. For this example, it is the value in G12.

The variation v_r of (12.9) appears in B16, that is,

$$B16: =B18*((H4-\$G\$12)^2+(H5-\$G\$12)^2+(H6-\$G\$12)^2)$$

The *variation between columns*, denoted as v_c , is defined as

$$v_c = r \sum_{j=1}^c (\bar{x}_j - \bar{x})^2 \quad (12.10)$$

where:

r = number of rows

c = number of columns

\bar{x}_j = the mean of the j th column. For this example j varies from 1 to 5, that is, B11 through F11.

\bar{x} = grand mean. For this example, it is the value in G12.

The variation v_c of (12.10) appears in B20, that is,

$$B20: =B14*((B11-\$G\$12)^2+(C11-\$G\$12)^2+(D11-\$G\$12)^2+(E11-\$G\$12)^2+(F11-\$G\$12)^2)$$

The *total variation*, denoted as v_t , is defined as

$$v_t = \sum_{i,j} (x_{ij} - \bar{x})^2 \quad (12.11)$$

where

x_{ij} = the value in i th row and j th column for the entire range of given samples. For this example, it represents the values in B4:F6.

\bar{x} = grand mean. For this example, it is the value in G12.

The variation v_t of (12.11) appears in B23.

$$B23: =(B4-\$G\$12)^2+(C4-\$G\$12)^2+(D4-\$G\$12)^2+(E4-\$G\$12)^2+(F4-\$G\$12)^2+(B5-\$G\$12)^2+(C5-\$G\$12)^2+(D5-\$G\$12)^2+(E5-\$G\$12)^2+(F5-\$G\$12)^2+(B6-\$G\$12)^2+(C6-\$G\$12)^2+(D6-\$G\$12)^2+(E6-\$G\$12)^2+(F6-\$G\$12)^2)$$

The *variation due to error*, or *random variation* v_e is defined as

$$v_e = v_t - v_r - v_c \quad (12.12)$$

The variation v_e of (12.12) appears in B26.

Chapter 12 Analysis of Variance (ANOVA)

B26: =B23–B16–B20

The formulas for the degrees of freedom for rows df_r and for columns df_c are also written on the spreadsheet. These are:

$$df_r = r - 1 \quad (12.13)$$

B28: =B14–1

$$df_c = c - 1 \quad (12.14)$$

B29: =B18–1

The formulas of the *degrees of freedom for error variation* df_{ev} and *total variation* df_{tv} appear on the spreadsheet also. These are:

$$df_{ev} = df_r \cdot df_c \quad (12.15)$$

B30: =B28*B29

$$df_{tv} = df_r + df_r + df_{ev} \quad (12.16)$$

B31: =SUM(B28:B30)

The formulas for the *mean squares* of v_r , v_c , and v_e , denoted as msv_r , msv_c , and msv_e respectively, appear in B33:B35. These are:

$$msv_r = \frac{v_r}{(r - 1)} \quad (12.17)$$

B33: =B16/(B14–1)

$$msv_c = \frac{v_c}{(c - 1)} \quad (12.18)$$

B34: =B20/(B18–1)

$$msv_e = \frac{v_e}{(r - 1)(c - 1)} \quad (12.19)$$

B35: =B26/((B14–1)*(B18–1))

For a two-way ANOVA without replication, we must consider two F ratios; the ratio due to variation between rows, denoted as F ratio_{vr}, and the ratio due to variation between columns, denoted as F ratio_{vc}. These are:

$$F \text{ ratio}_{vr} = \frac{msv_r}{msv_e} \quad (12.20)$$

B36: =B33/B35

$$F \text{ ratio}_{vc} = \frac{msv_c}{msv_e} \quad (12.21)$$

B37: =B34/B35

For convenience, we summarize the degrees of freedom, mean square, and F ratio values in B41:E43.

The final step is to look-up the appropriate F values in an F-distribution table, or compute these with the FINV function. We choose the latter. Thus,

F41: =FINV(0.01,2,8) = 8.65

F42: =FINV(0.01,4,8) = 7.01

These values are greater than those obtained by computations; therefore, with 1% level of significance, we can state that the averages are not significantly different, or stated differently, we can say with 99% certainty, that the averages are significantly different.

Let us check our results with Excel's Two-factor without replication analysis Toolpack.

We refer to the spreadsheet of Figure 12.5 and we type all labels (words) and values shown in Rows 1 through 4.

	A	B	C	D	E	F	G
1		Group1	Group 2	Group 3	Group 4	Group 5	
2	University A	92	91	87	93	95	
3	University B	84	89	81	90	87	
4	University C	88	92	85	89	86	
5							
6	Anova: Two-Factor Without Replication						
7							
8	SUMMARY	Count	Sum	Average	Variance		
9	University A	5	458	91.6	8.8		
10	University B	5	431	86.2	13.7		
11	University C	5	440	88	7.5		
12							
13	Group1	3	264	88	16		
14	Group 2	3	272	90.67	2.33		
15	Group 3	3	253	84.33	9.33		
16	Group 4	3	272	90.67	4.33		
17	Group 5	3	268	89.33	24.33		
18							
19							
20	ANOVA						
21	Source of Variation	SS	df	MS	F	P-value	F crit
22	Rows	75.6	2	37.8	8.16	0.01	8.65
23	Columns	82.93	4	20.73	4.47	0.03	7.01
24	Error	37.07	8	4.63			
25							
26	Total	195.6	14				

Figure 12.5. Two-way ANOVA without replication with Excel's Data Analysis.

Then, we click on *Tools>Data Analysis>Anova: Two-factor Without Replication>OK*. On the Anova: Two-Factor Without Replication box we enter the following:

Input Range: A1:F4. We click on the Labels small box to place a check mark, we change *Alpha* to

Chapter 12 Analysis of Variance (ANOVA)

0.01, we click on *Output Range*, we type A6 and click on OK. Excel then displays the answers shown in Rows 6 through 26. We observe that these values are the same as before.

12.3.2 Two-factor with Replication ANOVA

This method is an extension of the one-way ANOVA to include two or more samples for each group of data. In general, let A and B be two factors. Also, let a denote the *number of levels of Factor A* and b denote the *number of levels of Factor B*. In a two-factor with replication ANOVA, each level of Factor A is combined with each level of Factor B to form a total of ab *treatments*. We will illustrate this method with the following example.

Example 12.3

Suppose that the same test was given to male and female students at five different universities denoted as A, B, C, D, and E, and their scores are shown in Table 12.3.

TABLE 12.3 Scores achieved by male and female students at five universities

		A	B	C	D	E
Male Students	Test 1	92	91	87	93	95
	Test 2	94	93	88	94	94
	Test 3	91	92	87	90	94
	Test 4	93	89	90	89	93
Female Students	Test 1	94	90	86	91	93
	Test 2	92	89	87	91	91
	Test 3	90	91	86	89	92
	Test 4	89	89	86	90	92

Determine whether

- averages among university students are significantly different at the 0.01 level of confidence
- averages between male and female students are significantly different at the 0.01 significance level
- there is an interaction between male and female student scores at the 0.01 significance level

Solution:

As stated earlier, unlike the one-way ANOVA where we can use groups with unequal number of

samples, with two-way ANOVA we must have the same number of samples within each group. Table 12.4 shows an alternate method of displaying the information of Table 12.3.

TABLE 12.4 Alternate method of displaying the scores shown in Table 12.3

Factor A University	Factor B Gender	Replicates			
		Test 1	Test 2	Test 3	Test 4
A	Male	92	94	91	93
	Female	94	92	90	89
B	Male	91	93	91	89
	Female	90	89	91	89
C	Male	87	88	87	90
	Female	86	87	86	86
D	Male	93	94	90	89
	Female	91	91	89	90
E	Male	95	94	94	93
	Female	93	91	92	92

Either the universities or the students' gender can be chosen as Factor A or Factor B. Suppose we choose the universities to be Factor A, and the students' gender as Factor B. Then, A has five levels and B has 2 levels. Therefore, $a = 5$, $b = 2$, and there are $ab = 5 \times 2 = 10$ treatments.

In two-way, three-way ANOVA etc. with replication, we must also consider the *interaction* between conditions. For instance, in Table 12.4, we see that students performance on tests depends not only on a particular group, but also on the gender* of the students.

We enter the given scores on the spreadsheet of Figure 12.6.

Rows 1 through 22 contain the given scores and the means. The means were computed with the **AVERAGE** function. We enter the values that define the dimension of the data in D26:D28. For this example, Factor A has five levels (universities) and thus $a = 5$. Factor B has two levels (male and female students), and therefore, $b = 2$. We have used the letter r to denote the number of scores per group. Then, $r = 4$.

The *variation due to interaction* v_i is defined as

* Of course, this is an example. To suggest that male students make higher grades than female students, or vice versa, is absurd.

Chapter 12 Analysis of Variance (ANOVA)

	A	B	C	D	E	F	G	H
1	Two-way ANOVA for scores of five different groups of male and female students							
2								
3			A	B	C	D	E	Overall
4								Means
5	Male Students		92	91	87	93	95	
6			94	93	88	94	94	
7			91	92	87	90	94	
8			93	89	90	89	93	
9								
10		Means	92.50	91.25	88.00	91.50	94.00	91.45
11								
12	Female Students		94	90	86	91	93	
13			92	89	87	91	91	
14			90	91	86	89	92	
15			89	89	86	90	92	
16								
17		Means	91.25	89.75	86.25	90.25	92.00	89.90
18								
19	Combined Male &							
20	Female Means		91.88	90.50	87.13	90.88	93.00	
21								
22	Grand Mean	90.68						
23								
24	Dimensions of Data Table							
25								
26	b = number of row groups (gender)			2				
27	a = number of groups (columns)			5				
28	r = number of scores in each group			4				
29								
30	Variation due to interaction							
31								
32	Male Students		0.0225	0.0006	0.0100	0.0225	0.0506	
33	Female Students		0.0225	0.0006	0.0100	0.0225	0.0506	
34								
35	Variation between columns							
36								
37	Male Students		5.0000	8.7500	6.0000	17.0000	2.0000	
38	Female Students		14.7500	2.7500	0.7500	2.7500	2.0000	

Figure 12.6. Spreadsheet for two-way ANOVA with replication.

$$v_i = (\bar{x}_{A_j B_i} - \bar{x}_{A_j B_\Sigma} - \bar{x}_{A_\Sigma B_i} + \bar{x})^2 \quad (12.22)$$

Recalling that Factor A denotes the universities, and Factor B the gender of the students, in (12.22),

subscript i varies from 1 to 2 for rows (levels of Factor B).

subscript j varies from 1 to 5 for columns (levels of Factor A).

$\bar{x}_{A_j B_i}$ denotes the mean of each level of Factor A (universities) and each level of Factor B. This is shown in C10:G10 (male students) and C17:G17 (female students).

$\bar{x}_{A_j B_\Sigma}$ denotes the mean of each level of Factor A (universities) and both levels of Factor B (combined means of male and female students). This is shown in C20:G20.

$\bar{x}_{A_\Sigma B_i}$ denotes the mean of all levels of Factor A and the first level of Factor B. For this example, it is the value in H10 (for male students) and H17 (for female students).

\bar{x} denotes the grand mean. For this example is the value in B22.

Accordingly, we enter (12.22) in C32 as

C32: =(C10-C20-\$H\$10+\$B\$22)^2 and we copy it to D32:G32.

Also,

C33: =(C17-C19-\$H\$17+\$B\$22)^2 and we copy it to D33:G33.

Next, we define the *variation between columns*, v_c , as

$$v_c = \sum_{k=1}^r x_{kA_j}^2 - r\bar{x}_{A_j}^2 \quad (12.23)$$

where:

subscript k varies from 1 to 4 for number of scores per group.

subscript j varies from 1 to 5 for columns (levels of Factor A).

r = number of scores per group. For this example, r = 4.

x_{kA_j} = sample value of kth score in each group of scores in jth column. For this example, they are the values in C5:C8 ... G5:G8 (male students), and C12:C15 ... G12:G15 (female students).

\bar{x}_{A_j} = mean of sampled values in jth column. For this example, it appears in C10:G10 (male students) and C17:G17 (female students)

Then, we enter (12.23) in C37 as

C37: =C5^2+C6^2+C7^2+C8^2-\$D\$28*C10^2

and we copy it to D37:G37.

Also,

C38: =C12^2+C13^2+C14^2+C15^2-\$D\$28*C17^2

and we copy it to D38:G38.

Chapter 12 Analysis of Variance (ANOVA)

The remaining part of the spreadsheet appears in Figure 12.7.

	A	B	C	D	E
41	Computations				
42		Sum of	Deg. Of		
43		squares	freedom	Variance	F ratio
44					
45	Factor A	156.15	4	39.04	18.97
46					
47	Factor B	24.02	1	24.02	11.67
48					
49	Interaction	0.85	4	0.21	0.10
50					
51	Error	61.75	30	2.06	
52					
53	Total	242.77	39		
54					
55	Significance tests				
56					
57	Alpha	0.01			
58					F values
59	Factor A				
60					
61	F-table Lookup num. 4 denom. 30 or with FINV				4.02
62	CONCLUSION: Averages are significantly				
63	different at the 0.01 level of confidence				
64					
65	Factor B				
66					
67	F-table Lookup num. 1 denom. 30 or with FINV				7.56
68	CONCLUSION: Averages are significantly				
69	different at the 0.01 level of confidence				
70					
71	Interaction				
72					
73	F-table Lookup num. 4 denom. 30 or with FINV				4.02
74	CONCLUSION: There is no significant interaction				
75	at the 0.01 level of confidence				

Figure 12.7. Remaining part of the spreadsheet of Figure 12.6

The sum of squares for Factor A ss_A is given by

$$ss_A = br \sum_{j=1}^a \bar{x}_{jB_\Sigma}^2 - abr\bar{x}^2 \quad (12.24)$$

where:

b = row groups. For this example $b = 2$, i.e., male and female students. The value of b appears in D26.

r = number of rows that contain sampled values. For this example, $r = 4$, i.e., number of scores per gender. The value r appears in D28.

The subscript j varies from 1 to 5 for columns (levels of Factor A) where $a = 5$ and it appears in D27.

$\bar{x}_{jB_{\Sigma}}$ denotes the mean of the j th level of Factor A and the *combined* levels of Factor B (both genders). The values appear in C20:G20.

\bar{x} denotes the grand mean. It appears in B22.

For this example, (12.24) appears in B45 as

$$B45: =D26*D28*(C20^2+D20^2+E20^2+F20^2+G20^2)-D26*D27*D28*B22^2.$$

The *degrees of freedom for Factor A* df_A is defined as

$$df_A = a - 1 \tag{12.25}$$

Then,

$$C45: =D27-1$$

The *variance for Factor A* σ_A^2 is defined as

$$\sigma_A^2 = \frac{ss_A}{df_A} \tag{12.26}$$

Then,

$$D45: =B45/C45$$

The *sum of squares for Factor B* ss_B is defined as

$$ss_B = ar \sum_{i=1}^b \bar{x}_{A_{\Sigma}B_i}^2 - abr\bar{x}^2 \tag{12.27}$$

where

$\bar{x}_{A_{\Sigma}B_i}^2$ = the mean of the means of each gender. They are shown in H10 and H17. The other variables are as defined in (12.24).

Then,

$$B47: =D27*D28*(H10^2+H17^2)-D26*D27*D28*B22^2$$

The *degrees of freedom for Factor B* df_B is defined as

$$df_B = b - 1 \tag{12.28}$$

Then,

Chapter 12 Analysis of Variance (ANOVA)

C47: =D26-1

The *variance for Factor B* σ_B^2 is defined as

$$\sigma_B^2 = \frac{SS_B}{df_B} \quad (12.29)$$

Then,

D47: =B47/C47

The *sum of squares for interaction* ss_I is defined as

$$ss_I = r \sum v_i = r \sum (\bar{x}_{A_j B_i} - \bar{x}_{A_j B_\Sigma} - \bar{x}_{A_\Sigma B_i} + \bar{x})^2 \quad (12.30)$$

where the right side of (12.30) is as defined in (12.22). Then,

B49: =D28*SUM(C32:G33)

The *degrees of freedom for interaction* df_I is defined as

$$df_I = (a - 1)(b - 1) \quad (12.31)$$

Then,

C49: =(D27-1)*(D26-1)

The *variance for interaction* σ_I^2 is defined as

$$\sigma_I^2 = \frac{ss_I}{df_I} \quad (12.32)$$

Then,

D49: =B49/C49

The *sum of squares for error* ss_E is obtained by summing the formula of (12.23) which represents the variation between columns, that is,

$$ss_E = \sum v_c = \sum \left(\sum_{k=1}^r x_{kA_j}^2 - r\bar{x}_{A_j}^2 \right) \quad (12.33)$$

Then,

B51: =SUM(C37:G38)

The *degrees of freedom for error* df_E is defined as

$$df_E = ab(r - 1) \quad (12.34)$$

Then,

$$C51: =D27*\$D26*(D28-1)$$

The variance for error σ_E^2 is defined as

$$\sigma_E^2 = \frac{SS_E}{df_E} \quad (12.35)$$

Then,

$$D51: =B51/C51$$

The totals for the sum of squares and degrees of freedom are

$$B53: =SUM(B45:B51)$$

$$C53: =SUM(C45:C51)$$

In a Two-Factor with replication ANOVA, we consider three F ratios. We denote these as F ratio_{FA} for Factor A, F ratio_{FB} for Factor B, and F ratio_I for Interaction. They are defined as:

$$F \text{ ratio}_{FA} = \frac{\text{Variance for Factor A}}{\text{Variance for Error}} = \frac{\sigma_A^2}{\sigma_E^2} \quad (12.36)$$

$$F \text{ ratio}_{FB} = \frac{\text{Variance for Factor B}}{\text{Variance for Error}} = \frac{\sigma_B^2}{\sigma_E^2} \quad (12.37)$$

$$F \text{ ratio}_I = \frac{\text{Variance for Interaction}}{\text{Variance for Error}} = \frac{\sigma_I^2}{\sigma_E^2} \quad (12.38)$$

Then,

$$E45: =D45/D51$$

$$E47: =D47/D51$$

$$E49: =D49/D51$$

Finally, we make the following entries:

B57: 0.01 (value of alpha; can be changed to 0.05, 0.10, and so on.)

A61: =CONCATENATE("From F table with df1 = "&C45,", df2 = "&C51," or with FINV")

A62: ="CONCLUSION: Averages are "&IF(E45>E61,"","not ")&"significantly"

A63: =CONCATENATE("different at the "&B57," level of confidence")

Chapter 12 Analysis of Variance (ANOVA)

A67: =CONCATENATE("From F table with df1 = "&C47," , df2 = "&C51," or with FINV")

A68: ="CONCLUSION: Averages are "&IF(E47>E67,"","not ")&"significantly"

A69: =CONCATENATE("different at the "&B57," level of confidence")

A73: =CONCATENATE("From F table with df1 = "&C49," , df2 = "&C51," or with FINV")

A74: ="CONCLUSION: There is "&IF(E49>E73,"a","no ")&" significant interaction"

A75: =CONCATENATE("at the "&B57," level of confidence")

With the FINV function we obtain the values shown in E61, E67, and E72.

E61: =FINV(B57,C45,C51)

E67: =FINV(B57,C47,C51)

E73: =FINV(B57,C49,C51)

Figure 12.8 shows the so-called *interaction plot*. The upper line indicates that the male students' scores are slightly higher consistently. The fact that the lines are nearly parallel indicates that there is no interaction between male and female student scores. If the lines crossed or diverged, we would say that definitely there is some interaction.

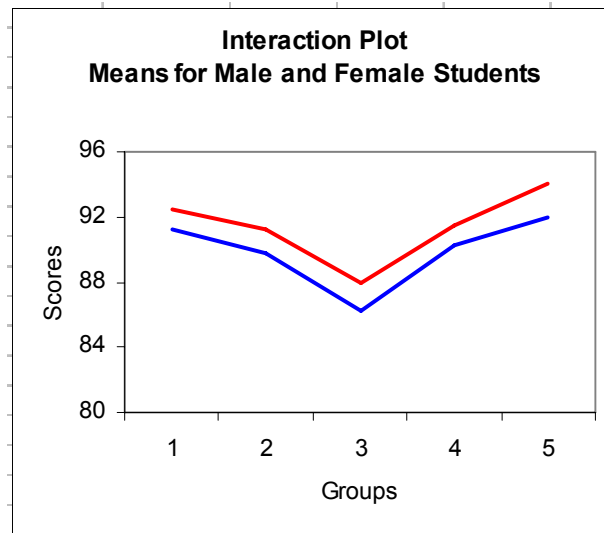


Figure 12.8. Interaction plot for two-way ANOVA with replication.

The plot of Figure 12.8 can be produced with Excel's *Chart Wizard* as follows:

We click on *Chart Wizard* icon on the main taskbar, we click on the *Line chart type*, and on the *Chart Sub-type*, we click on the left chart of the second row. We click *Next* and on the *Data Range* tab. In the *Data range* box we type C10:G10,C17:G17, and in *Series in* we click on *Rows*. We click on *Next*, and on the *Titles* tab make the entries:

Chart title: Interaction Plot

Category (X) axis: Groups

Value (Y) axis: Scores

With the Chart selected (square handles are visible), on the block below the main taskbar we select *Plot Area*, we click on small (with hand) box next to it, and we click on the white box to change the plot from gray to white. We click on OK, we select *Value Axis*, we click on small box next to it, we click on the *Scale* tab, and we enter the following:

Minimum: 80

Maximum: 96

Major Unit: 4

We click on OK, we select *Category Axis*, we click on small box next to it, and we type 1 in each of the three boxes. We click on OK to return to Chart.

We will check the results of Example 12.3 with Excel's Data Analysis toolpack.

We first arrange the rows and columns as shown in A79:F87 of the spreadsheet of Figure 12.9 and we perform the following steps:

Tools > Data Analysis > Anova: Two-Factor With Replication > OK, *Input Range:* A78:F87, *Rows per sample:* 4, *Alpha:* 0.01, we click on *Output Range* and there we type A89. Excel then displays the data shown in A89:G118 where the values have been rounded to two decimal places. We observe that they are the same as before.

Chapter 12 Analysis of Variance (ANOVA)

	A	B	C	D	E	F	G
79		Univ A	Univ B	Univ C	Univ D	Univ E	
80	Male Students	92	91	87	93	95	
81		94	93	88	94	94	
82		91	92	87	90	94	
83		93	89	90	89	93	
84	Female Students	94	90	86	91	93	
85		92	89	87	91	91	
86		90	91	86	89	92	
87		89	89	86	90	92	
88							
89	Anova: Two-Factor With Replication						
90							
91	SUMMARY	Univ A	Univ B	Univ C	Univ D	Univ E	Total
92	<i>Male Students</i>						
93	Count	4	4	4	4	4	20
94	Sum	370	365	352	366	376	1829
95	Average	92.5	91.25	88	91.5	94	91.45
96	Variance	1.67	2.92	2	5.67	0.67	6.16
97							
98	<i>Female Students</i>						
99	Count	4	4	4	4	4	20
100	Sum	365	359	345	361	368	1798
101	Average	91.25	89.75	86.25	90.25	92	89.9
102	Variance	4.92	0.92	0.25	0.92	0.67	5.36
103							
104	<i>Total</i>						
105	Count	8	8	8	8	8	
106	Sum	735	724	697	727	744	
107	Average	91.88	90.50	87.13	90.88	93.00	
108	Variance	3.27	2.29	1.84	3.27	1.71	
109							
110							
111	ANOVA						
112	<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
113	Sample	24.03	1	24.03	11.67	0.00	7.56
114	Columns	156.15	4	39.04	18.97	0.00	4.02
115	Interaction	0.85	4	0.21	0.10	0.98	4.02
116	Within	61.75	30	2.06			
117							
118	Total	242.78	39				

Figure 12.9. Two-way ANOVA with replication using Excel's Data Analysis tool.

12.4 Summary

- We use ANOVA in situations where we have a number of observed values, and these are divided among three or more groups. We want to know whether all these values belong to the same population, regardless of group, or we want to find out if the observations in at least one of the groups come from a different population. This determination is made by comparing the variation of values *within* groups and the variation of values *between* groups. The *variation within groups* is computed from the variation within groups of samples, whereas the *variation between groups* is computed from the variation among different groups of samples.
- We use one-way ANOVA when observations are obtained for A independent groups of samples where the number of observations in each group is B.
- For one-way ANOVA the *sum of squares between groups* ss_{bg} is computed from

$$ss_{bg} = \sum \frac{x^2}{n_i} - \frac{(\sum x)^2}{n}$$

where x = sum of each group, n_i = number of samples in i th group, and n = the total number of samples in all groups.

- For one-way ANOVA the *degrees of freedom between groups* df_{bg} is the number of groups minus one, that is,

$$df_{bg} = \text{number of groups} - 1$$

- For one-way ANOVA the *variance between groups* σ_{bg}^2 is obtained from

$$\sigma_{bg}^2 = \frac{\text{Sum of Squares Between Groups}}{\text{Degrees of Freedom Between Groups}} = \frac{ss_{bg}}{df_{bg}}$$

- For one-way ANOVA The *sum of squares within groups* ss_{wg} is computed from

$$ss_{wg} = \sum x^2 - \frac{(\sum x)^2}{n} - ss_{bg}$$

where x = the value of each sample in all groups, n = number of samples in all groups, and ss_{bg} = sum of squares between groups.

- For one-way ANOVA the *degrees of freedom within groups* df_{wg} is the total number of samples n in all groups minus the number of groups, that is,

$$df_{wg} = n - \text{number of groups}$$

Chapter 12 Analysis of Variance (ANOVA)

- For one-way ANOVA the *variance within groups* σ_{wg}^2 is obtained from

$$\sigma_{wg}^2 = \frac{\text{Sum of Squares Within Groups}}{\text{Degrees of Freedom Within Groups}} = \frac{SS_{wg}}{df_{wg}}$$

- For one-way ANOVA the *F ratio* is obtained from

$$F \text{ ratio} = \frac{\text{Variance Between Groups}}{\text{Variance Within Groups}} = \frac{\sigma_{bg}^2}{\sigma_{wg}^2}$$

Once the F ratio is found, we compare it with a value obtained from an F distribution table, to determine whether the samples are, or are not statistically different.

- Excel's *Data Analysis Toolpack* provides three ANOVA analysis tools, the *Single Factor Analysis Tool*, the *Two-Factor with Replication Analysis Tool*, and the *Two-Factor without Replication Analysis Tool*. The first applies to one-way ANOVA and the others to two-way ANOVA.
- In two-way ANOVA we consider two variables. A two-way ANOVA can be *two-factor without replication*, or a *two-factor with replication*.
- Two-factor without replication ANOVA performs a two-way ANOVA that does not include more than one sampling per group. With this method, we must have the same number of samples within each group.
- For two-factor without replication ANOVA the *variation between rows*, denoted as v_r , is defined as

$$v_r = c \sum_{i=1}^r (\bar{x}_i - \bar{x})^2$$

where c = number of columns, r = number of rows, \bar{x}_i = the mean of the i th row, and \bar{x} = grand mean.

- For two-factor without replication ANOVA the *variation between columns*, denoted as v_c , is defined as

$$v_c = r \sum_{j=1}^c (\bar{x}_j - \bar{x})^2$$

where r = number of rows, c = number of columns, \bar{x}_j = the mean of the j th column, and \bar{x} = grand mean.

- For two-factor without replication ANOVA the *total variation*, denoted as v_t , is defined as
-

$$v_t = \sum_{i,j} (x_{ij} - \bar{x})^2$$

where x_{ij} = the value in i th row and j th column for the entire range of given samples, and \bar{x} = grand mean.

- For two-factor without replication ANOVA the *variation due to error*, or *random variation* v_e is defined as

$$v_e = v_t - v_r - v_c$$

- For two-factor without replication ANOVA the degrees of freedom for rows df_r and for columns df_c are

$$df_r = r - 1$$

and

$$df_c = c - 1$$

- For two-factor without replication ANOVA the *degrees of freedom for error variation* df_{ev} and *total variation* df_{tv} are

$$df_{ev} = df_r \cdot df_c$$

and

$$df_{tv} = df_r + df_c + df_{ev}$$

For two-factor without replication ANOVA the formulas for the *mean squares* of v_r , v_c , and v_e , denoted as msv_r , msv_c , and msv_e respectively are

$$msv_r = \frac{v_r}{(r - 1)}$$

$$msv_c = \frac{v_c}{(c - 1)}$$

$$msv_e = \frac{v_e}{(r - 1)(c - 1)}$$

- For a two-way ANOVA without replication, we must consider two F ratios; the ratio due to variation between rows, denoted as F ratio $_{vr}$, and the ratio due to variation between columns, denoted as F ratio $_{vc}$. These are:

$$F \text{ ratio}_{vr} = \frac{msv_r}{msv_e}$$

$$F \text{ ratio}_{vc} = \frac{msv_c}{msv_e}$$

- The two-factor with replication ANOVA is an extension of the one-way ANOVA to include two or more samples for each group of data. In general, let A and B be two factors. Also, let a denote the *number of levels of Factor A* and b denote the *number of levels of Factor B*. In a two-factor with replication ANOVA, each level of Factor A is combined with each level of Factor B to form a total of ab *treatments*.
- In two-way, three-way ANOVA etc. with replication, we must also consider the *interaction* between conditions. The *variation due to interaction* v_i is defined as

$$v_i = (\bar{x}_{A_j B_i} - \bar{x}_{A_j B_\Sigma} - \bar{x}_{A_\Sigma B_i} + \bar{x})^2$$

where subscript i varies from 1 to m for rows (levels of Factor B), subscript j varies from 1 to n for columns (levels of Factor A), $\bar{x}_{A_j B_i}$ denotes the mean of each level of Factor A, $\bar{x}_{A_j B_\Sigma}$ denotes the mean of each level of Factor A and both levels of Factor B, $\bar{x}_{A_\Sigma B_i}$ denotes the mean of all levels of Factor A and the first level of Factor B, and \bar{x} denotes the grand mean.

- For two-factor with replication ANOVA the *variation between columns*, v_c , is defined as

$$v_c = \sum_{k=1}^r x_{kA_j}^2 - r\bar{x}_{A_j}^2$$

where the subscript k varies from 1 to m for rows, subscript j varies from 1 to n for columns, r = number of scores per group, x_{kA_j} = sample value of kth score in each group of scores in jth column, and \bar{x}_{A_j} = mean of sampled values in jth column.

- For two-factor with replication ANOVA the *sum of squares for Factor A* ss_A is given by

$$ss_A = br \sum_{j=1}^a \bar{x}_{jB_\Sigma}^2 - abr\bar{x}^2$$

where b = row groups, r = number of rows that contain sampled values, the subscript j varies from 1 to m (levels of Factor A), \bar{x}_{jB_Σ} denotes the mean of the jth level of Factor A and the *combined* levels of Factor B, and \bar{x} denotes the grand mean.

12.5 Exercises

1. The data shown in the table below, represent the mileage (miles per gallon) for automobiles manufactured by five different companies A through E, and these were obtained by different drivers in each company. Using one-way ANOVA, perform a test at the 0.05 significance level, to determine whether there is a difference in the mileage for each of the companies. Use Excel's Data Analysis to verify your answers.

Company A	Company B	Company C	Company D	Company E
32	30	25	33	32
34	32	26	34	31
31	31	29	30	34
33	28	31	29	33
34	29	30	31	35
32	28	28		34
	30			

2. The data shown in the table below, represent the automobiles assembled in an assembly line of an automobile manufacturer by three different shifts. Using two-way ANOVA without replication, perform a test at the 0.05 significance level to determine whether there is a difference in the productivity of each shift. Use Excel's Data Analysis to verify your answers.

	Automobiles assembled during days of the week				
	Monday	Tuesday	Wednesday	Thursday	Friday
Day Shift	32	31	27	33	35
Swing Shift	24	29	21	30	27
Graveyard Shift	28	32	25	29	26

3. The annual salaries (in thousands) of male and female employees of four different companies, all of which have department stores in five different cities, are shown in the table below. Use a two-way ANOVA with replication, to perform a test at the 0.05 level of significance to determine if there is a significant difference in salaries between

- gender
- cities

Use Excel's Data Analysis to verify your answers.

		City A	City B	City C	City D	City E
Male Employees	Company 1	52	51	47	53	55
	Company 2	50	49	45	54	52
	Company 3	48	50	47	51	52
	Company 4	53	51	50	49	51
Female Employees	Company 1	51	50	48	52	56
	Company 2	49	48	44	53	51
	Company 3	48	49	48	50	51
	Company 4	52	54	49	48	50

12.6 Solutions to End-of-Chapter Exercises

1. Following the procedure of Example 12.1, we construct the Excel spreadsheet below.

	A	B	C	D	E	F	G	H
1	ANOVA - Data							
2	Automobile						GRAND	SUM OF
3	Mileage	A	B	C	D	E	TOTALS	SQUARES
4								
5	Computations	32	30	25	33	32		4662
6	Miles per gallon	34	32	26	34	31		4973
7		31	31	29	30	34		4819
8		33	28	31	29	33		4764
9		34	29	30	31	35		5083
10		32	28	28		34		3748
11			30					900
12								
13								
14	Totals	196	208	169	157	199	929	28949
15								
16	Observations	6	7	6	5	6	30	
17								
18	Averages	32.67	29.71	28.17	31.40	33.17	30.97	
19								
20	No. of Groups		5					
21								
22	ANOVA Table							
23								
24		Sum of	Deg. Of					
25		Squares	Freedom	Variance	F-ratio			
26								
27	Between							
28	Groups	105.34	4	26.33	8.71			
29								
30	Within							
31	Groups	75.63	25	3.03				
32								
33	Totals	180.97	29					
34								
35	F-table Lookup num. 4 denom. 25				4.18			
36	at 1% (0.01)							
37								
38	CONCLUSION: Averages are significantly different							
39								
40	Probability that averages are not significantly different is					0.02%		

With Excel's Data Analysis Toolpack we get the following data:

Chapter 12 Analysis of Variance (ANOVA)

	I	J	K	L	M	N	O
1	Anova: Single Factor						
2							
3	SUMMARY						
4	<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>		
5	Column 1	6	196	32.66667	1.466667		
6	Column 2	7	208	29.71429	2.238095		
7	Column 3	6	169	28.16667	5.366667		
8	Column 4	5	157	31.4	4.3		
9	Column 5	6	199	33.16667	2.166667		
10							
11							
12	ANOVA						
13	<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
14	Between Groups	105.3381	4	26.33452	8.705217	0.000152	4.177423
15	Within Groups	75.62857	25	3.025143			
16							
17	Total	180.9667	29				

2. Following the procedure of Example 12.2, we construct the Excel spreadsheet below.

	A	B	C	D	E	F	G	H
1	ANOVA: Two-factor without replication							
2			Groups of Students				Row	Row
3		Mon	Tue	Wed	Thu	Fri	Sum	Means
4	Day Shift	32	31	27	33	35	158	31.6
5	Swing Shift	24	29	21	30	27	131	26.2
6	Graveyard Shift	28	32	25	29	26	140	28.0
7								
8	Column Sums	84	92	73	92	88		
9	Grand Sum (Sum of Column Sums)						429	
10								
11	Column Means	28.00	30.67	24.33	30.67	29.33		
12	Grand Mean (Mean of Column means)						28.60	
13								
14	Number of rows $r =$	3						
15	Variation of Row Means from Mean of all Columns							
16	(H4, H5, and H6 minus G12 squared) $vr =$	75.60						
17								
18	Number of columns $c =$	5						
19	Variation of Column Means from Mean of all Columns							
20	(B11, C11, D11, E11, and F11 minus G12 squared) $vc =$	82.93						
21								
22	Total variation							
23	Each value of B4 through F6 minus G12 squared) $vt =$	195.6						
24								
25	Error variation, ve							
26	$ve = vt - vr - vc$ $ve =$	37.07						
27								
28	Degrees of freedom for rows $dfr = r - 1$ $dfr =$	2						
29	Degrees of freedom for columns $dfc = c - 1$ $dfc =$	4						
30	Degrees of freedom for error $dfev = dfr * dfc$ $dfev =$	8						
31	Deg. of freedom for total variation $dfr + dfc + dfev = dftv =$	14						
32								
33	Mean square for vr $msvr = vr / (r - 1)$ $msvr =$	37.80						
34	Mean square for vc $msvc = vc / (c - 1)$ $msvc =$	20.73						
35	Mean square for ve $msve = ve / [(r - 1) * (c - 1)]$ $msve =$	4.63						
36	F ratio for $vr = msvr / msve$ $msvr / msve$ F ratio =	8.16						
37	F ratio for $vc = msvc / msve$ $msvc / msve$ F ratio =	4.47						
38				Mean	F	From F tables		
39	Variations Summary		df	Square	ratio	or with FINV		
40								
41		$vr =$	75.60	2	37.80	8.16	8.65	
42		$vc =$	82.93	4	20.73	4.47	7.01	
43		$ve =$	37.07	8	4.63			
44		$vt = vr + vc + ve =$	195.60					
45								
46	CONCLUSION 1: At 0.01 level of significance, the averages are not significantly different among shifts							
47	CONCLUSION 2: At 0.01 level of significance, the averages are not significantly different among days							

With Excel's Data Analysis Toolpack we get the following data:

Chapter 12 Analysis of Variance (ANOVA)

	J	K	L	M	N	O	P	
1	Anova: Two-Factor Without Replication							
2								
3	SUMMARY	Count	Sum	Average	Variance			
4	Row 1	5	158	31.6	8.8			
5	Row 2	5	131	26.2	13.7			
6	Row 3	5	140	28	7.5			
7								
8	Column 1	3	84	28	16			
9	Column 2	3	92	30.66667	2.333333			
10	Column 3	3	73	24.33333	9.333333			
11	Column 4	3	92	30.66667	4.333333			
12	Column 5	3	88	29.33333	24.33333			
13								
14								
15	ANOVA							
16	Source of Variation	SS	df	MS	F	P-value	F crit	
17	Rows	75.6	2	37.8	8.158273	0.011715	8.649067	
18	Columns	82.93333	4	20.73333	4.47482	0.03427	7.006065	
19	Error	37.06667	8	4.633333				
20								
21	Total	195.6	14					

3. Following the procedure of Example 12.3, we construct the Excel spreadsheet shown on Page 12–35 and continued on Page 12–36.

	A	B	C	D	E	F	G	H
1	Two-way ANOVA with replication - Salaries (in thousands) for male							
2	and female employees in 4 different companies and 5 different cities							
3			A	B	C	D	E	Overall
4								Means
5	Male Employees		52	51	47	53	55	
6			50	49	45	54	52	
7			48	50	47	51	52	
8			53	51	50	49	51	
9								
10		Means	50.75	50.25	47.25	51.75	52.50	50.5
11								
12	Female Employees		51	50	48	52	56	
13			49	48	44	53	51	
14			48	49	48	50	51	
15			52	54	49	48	50	
16								
17		Means	50.00	50.25	47.25	50.75	52.00	50.05
18								
19	Combined Male &							
20	Female Means		50.38	50.25	47.25	51.25	52.25	
21								
22	Grand Mean	50.28						
23								
24	Dimensions of Data Table							
25								
26	b = number of row groups (gender)			2				
27	a = number of groups (columns)			5				
28	r = number of scores in each group			4				
29								
30	Variation due to interaction							
31								
32	Male Students		0.0225	0.0506	0.0506	0.07562	0.0006	
33	Female Students		0.0225	0.0506	0.0506	0.07562	0.0006	
34								
35	Variation between columns							
36								
37	Male Students		14.7500	2.7500	12.7500	14.7500	9.0000	
38	Female Students		10.0000	20.7500	14.7500	14.7500	22.0000	
39								
40								
41	Computations							
42		Sum of	Deg. Of					
43		squares	freedom	Variance	F ratio			
44								
45	Factor A	112.1	4	28.03	6.17			
46								
47	Factor B	2.02	1	2.02	0.45			
48								
49	Interaction	1.6	4	0.40	0.09			
50								
51	Error	136.25	30	4.54				
52								
53	Total	251.98	39					

	A	B	C	D	E
55	Significance tests				
56					
57	Alpha	0.01			
58					F values
59	Factor A				
60					
61	From F table with $df_1 = 4$, $df_2 = 30$ or with FINV				4.02
62	CONCLUSION: Averages are significantly				
63	different at the 0.01 level of confidence				
64					
65	Factor B				
66					
67	From F table with $df_1 = 1$, $df_2 = 30$ or with FINV				7.56
68	CONCLUSION: Averages are not significantly				
69	different at the 0.01 level of confidence				
70					
71	Interaction				
72					
73	From F table with $df_1 = 4$, $df_2 = 30$ or with FINV				4.02
74	CONCLUSION: There is no significant interaction				
75	at the 0.01 level of confidence				
76					
77					

The interaction plot is shown below.



With Excel's Data Analysis Toolpack we get the following data:

	A	B	C	D	E	F	G
89	Anova: Two-Factor With Replication						
90							
91	SUMMARY	City A	City B	City C	City D	City E	Total
92	<i>Male Employees</i>						
93	Count	4	4	4	4	4	20
94	Sum	203	201	189	207	210	1010
95	Average	50.75	50.25	47.25	51.75	52.5	50.5
96	Variance	4.91667	0.91667	4.25	4.916667	3	6.26316
97							
98	<i>Female Employees</i>						
99	Count	4	4	4	4	4	20
100	Sum	200	201	189	203	208	1001
101	Average	50	50.25	47.25	50.75	52	50.05
102	Variance	3.33333	6.91667	4.916667	4.916667	7.33333	6.89211
103							
104	<i>Total</i>						
105	Count	8	8	8	8	8	
106	Sum	403	402	378	410	418	
107	Average	50.375	50.25	47.25	51.25	52.25	
108	Variance	3.69643	3.35714	3.928571	4.5	4.5	
109							
110							
111	ANOVA						
112	<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
113	Sample	2.025	1	2.025	0.445872	0.50941	7.56245
114	Columns	112.1	4	28.025	6.170642	0.00095	4.01786
115	Interaction	1.6	4	0.4	0.088073	0.98549	4.01786
116	Within	136.25	30	4.541667			
117							
118	Total	251.975	39				

Appendix A

Numbers and Arithmetic Operations

This appendix is a review of the basic arithmetic concepts. It is intended for readers feeling that they need a math review from the very beginning. It forms the basis for understanding and working with relations (formulas) encountered in business, science and technology. Readers with a fair mathematical background may skip this appendix. Others may find it useful as well as a convenient source for review.

A.1 Number Systems

The *decimal* (base 10) number system uses the *digits* (numbers) 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. This is the number system we use in our everyday arithmetic calculations such as the monetary transactions. Another number system is the *binary* (base 2) that uses the digits 0 and 1 only. The binary system is used in computers and it is being taught in electronics courses. We will not be concerned with the binary system in this text.

A.2 Positive and Negative Numbers

A *positive number* is a number greater than zero and it is understood to have a plus (+) sign in front of it. The (+) sign in front of a positive number is generally omitted. Thus, any number without a sign in front of it is understood to be a positive number. A *negative number* is less than zero and it is written with a minus (-) sign* in front of it. The minus (-) sign in front of a negative number is a must; otherwise it would not be possible to distinguish the negative from the positive numbers. Positive and negative numbers can be whole (integer) or fractional numbers. Several examples will be presented in this chapter to illustrate their designation, how they are added, subtracted, multiplied, and divided with other numbers. To avoid confusion between the addition operation (+) and positive numbers, which are also denoted with the (+) sign, we will enclose positive numbers with their sign inside parentheses whenever necessary. Likewise, we will enclose negative numbers in parentheses to distinguish them from the subtraction (-) symbol. This will be illustrated with the examples that follow.

Example A.1

Joe Smith's checking account shows a balance of \$534.29. Thus, we can say that his balance is

* The financial community, such as banks, usually enclose a negative number in parentheses without the minus sign. Most often, this designation appears in financial statements.

Appendix A Numbers and Arithmetic Operations

+534.39 dollars but we normally omit the plus (+) sign, and we say that his balance is 534.39 dollars.

Example A.2

Bill Jones, unaware that his checking account has a balance of only \$78.31, makes a purchase of \$128.74. He pays this amount with a check. His new account balance is now -128.74 dollars. Here, the minus (–) sign is a must.

The *absolute value* of a number is that number without a positive or negative sign, and is enclosed in small vertical lines. For example, the absolute value of X is written as $|X|$. The number 0 (*zero*) is considered neither positive nor negative; it is the number that separates the negative from the positive numbers. The positive and negative numbers that we are familiar with, are referred to as the *real numbers*^{*} and are shown below on the so-called *real axis of numbers*.[†]

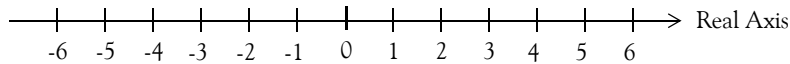


Figure A.1. Representation of Real Numbers

In our subsequent discussion, we will only be concerned with real numbers and thus the word *real* will not be used further.

A.3 Addition and Subtraction

The following rules apply for the addition of numbers.

Rule 1: *To add numbers with the same sign, we add the absolute values of these numbers and place the common sign (+ or –) in front of the result (sum). We can omit the plus sign in the result if positive. We must not omit the minus sign if the result is negative.*

Example A.3

Perform the addition

$$7 + 16 + 0.5$$

* The reader may have heard the expression “**imaginary numbers**”. The square root of minus 1, i.e. $\sqrt{-1}$, is an example of an imaginary number; it does not fit anywhere in the real axis of numbers. We will not be concerned with these numbers in this text. There is a brief discussion in Appendix B in conjunction with MATLAB.

† Only whole numbers are shown on the real axis of Figure A.1. However, it is understood that within each division, there are numbers such as 1.5, -2.75 etc.

Solution:

The plus sign between the given numbers indicates addition of three positive numbers whose sign is positive and it is omitted. However, we can enclose these numbers in parentheses just to emphasize that the numbers are positive. Addition of the absolute values of these numbers yield a sum of 23.5, and this also represents an absolute value. We should remember that the absolute value of a number is that just that number without regard to being positive or negative. Now, since all three numbers are positive, the sum is +23.5 or simply 23.5 as shown below. The final result, 23.5, does not represent an absolute number; it is a positive number whose sign has been omitted, as it is customary. Thus,

$$7 + 16 + 0.5 = (+7) + (+16) + (+0.5) \Rightarrow |7| + |16| + |0.5| = |23.5| \Rightarrow (+23.5) = 23.5$$

where the symbol \Rightarrow means conversion from signed numbers to absolute values and vice versa. Of course, these steps will be unnecessary after one becomes familiar with the rules.

Example A.4

Perform the addition

$$(-6) + (-45)$$

Solution:

Here, we are asked to add two negative numbers as indicated by the addition sign between them. Addition of the absolute values yields 51 and since both numbers are negative, we place the minus sign in front of the result. Therefore, the sum of these numbers is (-51) or -51. The minus sign cannot be omitted. Thus,

$$(-6) + (-45) \Rightarrow |6| + |45| = |51| \Rightarrow (-51) = -51$$

Example A.5

Perform the addition

$$(-1.25) + (-0.75) + (-3)$$

Solution:

In this example, we are asked to add three negative numbers. The sum of the absolute values is 5. Since all given numbers are negative, we place the minus sign in front of the sum. The result then is (-5) or simply -5. The negative sign cannot be omitted. Thus,

$$(-1.25) + (-0.75) + (-3) \Rightarrow |1.25| + |0.75| + |3| = |5| \Rightarrow (-5) = -5$$

Consider the subtraction of number B from number A, that is, $A - B$. The number A is called the

Appendix A Numbers and Arithmetic Operations

minuend and the number B is called the *subtrahend*. The result of the subtraction is called the *difference*. These definitions are illustrated with Examples A.6 and A.7 below.

Example A.6

Draw a rough sketch to indicate the minuend, subtrahend, and difference for the subtraction operation of $735 - 592$.

Solution:

The minuend, subtrahend, and difference are shown on the sketch below. The procedure for finding the difference will be explained in Rule 3 below.

$$\begin{array}{r} 735 - 592 = 143 \\ \begin{array}{l} \text{-----} \\ | \\ \text{-----} \\ | \\ \text{-----} \end{array} \end{array}$$

Difference
Subtrahend
Minuend

Example A.7

Draw a rough sketch to indicate the minuend, subtrahend, and difference for the subtraction operation of $248 - 857$.

Solution:

The minuend, subtrahend, and difference are shown on the sketch below. The procedure for finding the difference is our next topic.

$$\begin{array}{r} 248 - 857 = -609 \\ \begin{array}{l} \text{-----} \\ | \\ \text{-----} \\ | \\ \text{-----} \end{array} \end{array}$$

Difference
Subtrahend
Minuend

Rule 2: To add two numbers with different signs, we subtract the number with the smaller absolute value from the number with the larger absolute value, and we place the sign of the larger number in front of the result (sum).

Example A.8

Perform the addition

$$37 + (-15)$$

Solution:

The number with the smaller absolute value is 15; therefore, we subtract this from the absolute value of the larger number, 37, and we place the plus sign in front of the difference as shown below because the larger number is positive. Since the result is a positive number, we omit the plus sign.

$$37 + (-15) = (+37) + (-15) \Rightarrow |37| - |15| = |22| \Rightarrow (+22) = 22$$

Example A.9

Perform the addition

$$(-16) + 7$$

Solution:

The number with the smaller absolute value is 7; therefore, we subtract this from the absolute value of the larger number which is 16, and we place the minus sign in front of the difference as shown below.

$$(-16) + 7 = (-16) + (+7) \Rightarrow |16| - |7| = |9| \Rightarrow (-9) = -9$$

Rule 3: *To subtract one number (the subtrahend) from another number (the minuend) we change the sign (we replace + with - or - with +) of the subtrahend, and we perform addition instead of subtraction.*

Example A.10

Perform the subtraction

$$39 - 25$$

Solution:

For this example, both the minuend and subtrahend are positive numbers. The minus sign between them indicates subtraction. Therefore, we enclose them in parentheses with the positive (+) sign, and we change the sign of the subtrahend from plus to minus, while at the same time, we change the subtraction operation to addition. Next, we need to add these numbers, one of which is positive and the other negative. For this addition we follow Rule 2, that is, we subtract the number with the smaller absolute number from the larger. The steps are shown below.

$$39 - 25 = (+39) - (+25) = (+39) + (-25) \Rightarrow |39| - |25| = |14| \Rightarrow +14 = 14$$

Appendix A Numbers and Arithmetic Operations

Example A.11

Perform the subtraction

$$53 - (-18)$$

Solution:

Here, the minuend is positive, the subtrahend is negative, and the minus sign between them indicates subtraction. Therefore, we change the sign of the subtrahend from minus to plus and at the same time we change the subtraction operation to addition. Next, we need to add these two positive numbers, and for this addition we follow Rule 1, that is, we add the absolute values and we place the plus sign in front of the result. The steps are shown below.

$$53 - (-18) = (+53) - (-18) = (+53) + (+18) \Rightarrow |53| + |18| = |71| \Rightarrow +71=71$$

Again, these steps indicate the “train of thought” and the reader is not expected to write down all these steps when performing arithmetic operations.

Example A.12

Perform the subtraction

$$(-86) - 37$$

Solution:

Here, the minuend is negative, the subtrahend is positive, and the minus sign between them indicates subtraction. Therefore, we enclose 37 in parentheses with the positive (+) sign, then, change the sign of it from plus to minus, and at the same time we change the subtraction operation to addition. Next, we must add these two negative numbers, and for this addition we follow Rule 1, that is, we add the absolute values and we place the minus sign in front of the result. The steps are shown below.

$$(-86) - 37 = (-86) - (+37) = (-86) + (-37) \Rightarrow |86| + |37| = |123| \Rightarrow -123$$

Example A.13

Perform the subtraction

$$(-75) - (-125)$$

Solution:

Here, both the minuend and subtrahend are negative, and the minus sign between them indicates subtraction. Therefore, we change the sign of the subtrahend from minus to plus, and at the same time we change the subtraction operation to addition. Next, we must add these numbers which now have different signs. For this addition we follow Rule 2, that is, we subtract the number with

the smaller absolute number from the larger, and we place the sign of the larger number in front of the result. The steps are shown below.

$$(-75) - (-125) = (-75) + (+125) \Rightarrow |125| - |75| = |50| \Rightarrow +50 = 50$$

In general, let X be any number; then, the following relations apply for the possible combinations of addition and subtraction of positive and negative numbers.

$$\begin{aligned} +(+X) &= +X = X \\ +(-X) &= -X \\ -(+X) &= -X \\ -(-X) &= X \end{aligned} \tag{A.1}$$

A.4 Multiplication and Division

Note A.1

The multiplication (\times) and division (\div) signs do not interfere with the plus and minus signs of positive and negative numbers; therefore, for the examples which follow, we will omit the steps involving absolute values.

Note A.2

Multiplication is repeated addition, for instance, $5 \times 3 = 3 + 3 + 3 + 3 + 3 = 15$. One of the two numbers involved in multiplication is the *multiplier*, and the other is the *multiplicand*. The dictionary defines multiplier the number by which another number is multiplied. The multiplicand is defined as the number that is or is to be multiplied by another. Thus, in 5×3 , the multiplier is 5 and the multiplicand is 3. But since we can change the order of multiplication as $5 \times 3 = 3 \times 5$, either number can be the multiplicand or the multiplier. In this text, we will refer to the first number as the multiplier and the second as the multiplicand. The result of the multiplication is called the *product*.

The following rules apply for multiplication and division of numbers:

Rule 4: *When two numbers with the same sign are multiplied, the product will be positive (+). If the numbers have different signs, the product will be negative (-).*

Example A.14

Perform the multiplication

Appendix A Numbers and Arithmetic Operations

$$39 \times 25$$

Solution:

For this example, both the multiplier and multiplicand are positive. Since they have the same sign, the product will be positive. The steps are shown below.

$$39 \times 25 = (+39) \times (+25) = (+975) = 975$$

Note A.3

The reader can use a hand calculator to obtain the result. The intent here is to illustrate how the sign of the result is obtained. The same is true for the examples that follow.

Example A.15

Perform the multiplication

$$53 \times (-18)$$

Solution:

For this example, the multiplier is positive and the multiplicand is negative. Since they have different signs, in accordance with Rule 4 the product will be negative. The steps are shown below.

$$53 \times (-18) = (+53) \times (-18) = (-954) = -954$$

Example A.16

Perform the multiplication

$$(-86) \times 37$$

Solution:

Here, the multiplier is negative and the multiplicand is positive. Since they have different signs, in accordance with Rule 4, the product will be negative. The steps are shown below.

$$(-86) \times 37 = (-86) \times (+37) = (-3182) = -3182$$

Example A.17

Perform the multiplication

$$(-75) \times (-125)$$

Solution:

For this example, both the multiplier and multiplicand are negative. Since they have the same

sign, in accordance with Rule 4, the product will be positive. The steps are shown below.

$$(-75) \times (-125) = (+9375) = 9375$$

Note A.4

A *rational number* is a number that can be expressed as an integer or a quotient of integers, excluding zero as a denominator. In division, the number that is to be divided by another is called the *dividend*. For instance, the number $\frac{27}{3}$ is a rational number and the dividend is 27. The dividend is also referred to as the *numerator*. The number by which the dividend is to be divided, is called the *divisor*. For instance, in the division operation above, the divisor is 3. The divisor is also called the *denominator*. The number obtained by dividing one quantity by another is the *quotient*. Thus, in

$$\frac{27}{3} = 9$$

the quotient is 9.

Rule 5: When two numbers with the same sign are divided, the quotient will be positive (+). If the numbers have different signs, the quotient will be negative (-).

Example A.18

Perform the division

$$\frac{125}{5}$$

Solution:

Here, both the dividend and divisor are positive. Since they have the same sign, in accordance with Rule 5 the quotient will be positive. The steps are shown below.

$$\frac{125}{5} = \frac{(+125)}{(+5)} = (+25) = 25$$

Example A.19

Perform the division

$$\frac{-750}{10}$$

Solution:

Here, the dividend is negative and the divisor is positive. Since they have different signs, in accor-

Appendix A Numbers and Arithmetic Operations

dance with Rule 5 the quotient will be negative. The steps are shown below.

$$\frac{-750}{10} = \frac{(-750)}{(+10)} = (-75) = -75$$

Example A.20

Perform the division

$$\frac{1500}{-75}$$

Solution:

Here, the dividend is positive and the divisor is negative. Since they have different signs, in accordance with Rule 5 the quotient will be negative. The steps are shown below.

$$\frac{1500}{-75} = \frac{(+1500)}{(-75)} = (-20) = -20$$

Example A.21

Perform the division

$$\frac{-450}{-15}$$

Solution:

Here, both the dividend and divisor are negative. Since they have the same sign, in accordance with Rule 5 the quotient will be positive. The steps are shown below.

$$\frac{-450}{-15} = \frac{(-450)}{(-15)} = (+30) = 30$$

A.5 Integer, Fractional, and Mixed Numbers

A *decimal point* is the dot in a number that is written in decimal form. Its purpose is to indicate where values change from positive to negative powers of 10. Positive and negative powers will be discussed in Section A.8. For instance, the number 5.3 is written in decimal form, and the dot between 5 and 3 is referred to as the decimal point.

Every number has a decimal point although it may not be shown. A number is classified as an *integer*, *fractional* or *mixed* depending on the position of the decimal point in that number.

An *integer* number is a whole number and it is understood that its decimal point is positioned after the last (rightmost) digit although not shown. For instance,

$$7 = 7. \quad -305 = -305. \quad 14108 = 14108.$$

The numbers we have used in all examples thus far are integer numbers.

A *fractional* number, positive or negative, is always a number less than one. It can be expressed in a rational form such as $\frac{3}{4}$ or in *decimal point form* such as 0.75. When written in decimal point form, the decimal point appears in front of the first (leftmost) digit. For instance,

$$\frac{1}{2} = .5 = 0.5 \quad \frac{-1}{4} = -.25 = -0.25$$

where $\frac{1}{2}$ and $\frac{-1}{4}$ are fractional numbers expressed in rational form, and 0.5 and -0.25 are fractional numbers expressed in decimal point form.

Note A.5

When writing fractional numbers in decimal form, it is highly recommended that a 0 (zero) is written in front of the decimal point. This reduces the possibility of an erroneous reading if the decimal point goes unnoticed. The presence of a zero in front of the decimal point alerts the reader that a decimal point may follow the zero. We will follow this practice throughout this text.

A *mixed number* consists of an integer and a fractional number. For instance, the number 409.0875 consists of the integer part 409 and the fractional part 0.0875 as shown below.

$$409.0875 = 409. + 0.0875$$

Rule 6: Any number (integer, fractional or mixed) may be written in a rational form where the numerator (dividend) is the number itself and the denominator (divisor) is 1.

For instance, the numbers 13, 385, 2.75, 22/7, and -0.25 may be written as

$$13 = \frac{13}{1} \quad 385 = \frac{385}{1} \quad 2.75 = \frac{2.75}{1} \quad \frac{22}{7} = \frac{22/7}{1} \quad -0.25 = \frac{-0.25}{1}$$

A.6 Reciprocals of Numbers

The *reciprocal* of an integer number is a fraction with numerator 1 and denominator that number itself. Alternately, since any number can be considered as a fraction with denominator 1, the reciprocal of a fraction is another fraction with the numerator and denominator reversed. Conse-

Appendix A Numbers and Arithmetic Operations

quently, the product (multiplication) of any number by its reciprocal is 1.

Example A.22

Find the reciprocals of

a. 3 b. $\frac{1}{4}$ c. -3.875 d. $\frac{-1}{2}$

Solution:

a.

The reciprocal of 3 is $\frac{1}{3}$

b.

The reciprocal of $\frac{1}{4}$ is $\frac{4}{1}$ or 4

c.

The reciprocal of -3.875 is $\frac{1}{-3.875} = -0.2581$

d.

The reciprocal of $\frac{-1}{2}$ is $\frac{2}{-1}$

Note A.6

When either the sign of the numerator or the denominator (but not both) is negative, it is customary to place the minus sign in front of the bar that separates them. For instance,

$$\frac{-5}{8} = -\frac{5}{8} \quad \frac{4}{-11} = -\frac{4}{11}$$

We must not forget that when the sign of both the numerator and denominator is minus, the result is a positive number in accordance with *Rule 5*.

Rule 7: *A rational number in which the numerator and denominator are the same number is always equal to 1 with the proper sign.*

For instance,

$$\frac{7}{7} = 1 \quad \frac{-12}{12} = -\frac{12}{12} = -1$$

Rule 8: *As a consequence of Rule 7, the value of a rational number is not changed if both numerator and denominator are multiplied by the same number.*

A.7 Arithmetic Operations with Fractional Numbers

The following rules apply to addition, subtraction, multiplication and division with fractional numbers. As stated in Rule 8, the value of a rational number is not changed if both the numerator and denominator are multiplied by the same number.

Rule 9: *To add two or more fractional numbers in rational form, we first express the numbers in a common denominator; then, we add the numerators to obtain the numerator of the sum. The denominator of the sum is the common denominator. If the numbers are in decimal point form, we first align the given numbers with the decimal point, and we add the numbers as it is done in normal addition.*

Example A.23

Perform the addition

$$\frac{1}{2} + \frac{1}{4}$$

Solution:

Before we add these numbers, we must express them in a common (same) denominator form. Here $1/2$ is the same as $2/4$ and this and the number $1/4$ have the same (common) denominator. Then, by Rule 9,

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

Note A.7

In an expression such as the above, the individual numbers $1/2$ and $1/4$ are referred to as the *terms* of the expression. Of course, an expression may also have three or more terms.

Example A.24

Perform the addition

$$\frac{1}{25} + \frac{1}{10} + \frac{1}{20} \tag{A.2}$$

Solution:

To perform this addition, we must first express the given numbers so that all have the same denominator. This can be done by first finding the *Least Common Multiple (LCM)*. The LCM is the smallest quantity that is divisible by two or more given quantities without a remainder. For this example, the LCM is 100 which, when divided by 25, 10 and 20 yields 4, 10 and 5 respec-

Appendix A Numbers and Arithmetic Operations

tively. We can verify that the LCM is 100 with *Microsoft Excel** as follows.

Start Excel. The spreadsheet appears with a rectangle in cell A1. The cell which is surrounded by this rectangle is the *selected* cell. Thus, when Excel is first invoked (started), the selected cell is A1. You can select any other cell by moving the thick hollow white cross pointer to the desired cell and click the mouse. Select cell A3. Next, move the mouse towards the top of the screen where several *icons* (symbols and pictures) appear. This is the *Standard Toolbar*. Click on f_x . The *Function Category* menu appears. Select *Math & Trig*. Scroll down the function name list to locate LCM and click on it. In Number 1 box type 25, in Number 2 type 10, and in Number 3 type 20. Click on OK and observe that Excel displays 100. This is the LCM for this example.

Next, we need to express each of the terms of (A.2) with the common denominator 100. Of course, we must multiply the numerators of each term with the appropriate number so that numerical value of each term will not change. The new numerator values are found by multiplying each numerator by the quotient obtained by dividing 100 by each of the denominators of

(A.2). Thus, we find the new numerator of $\frac{1}{25}$ by first dividing 100 by 25 and this division yields 4. Then, we multiply this number by the numerator 1 and we get $1 \times 4 = 4$. Therefore, the number $\frac{1}{25}$ is now expressed as $\frac{4}{100}$. Likewise, we express $\frac{1}{10}$ as $\frac{10}{100}$ and $\frac{1}{20}$ as $\frac{5}{100}$. In other words, we have multiplied the numerator and denominator of each term by the same number so that their values will not change. Then, by Rule 9 we have:

$$\frac{1}{25} + \frac{1}{10} + \frac{1}{20} = \frac{1 \times 4}{25 \times 4} + \frac{1 \times 10}{10 \times 10} + \frac{1 \times 5}{20 \times 5} = \frac{4}{100} + \frac{10}{100} + \frac{5}{100} = \frac{19}{100}$$

We can verify the answer directly with Excel as follows:

Begin with a blank worksheet. Click on the letter A of Column A and observe that the entire column is now highlighted, that is, the background on this column has changed from white to black. Click on **Format** on the Menu bar (top of the screen) and on the drop menu click on **Cells**. The Format Cells menu appears. Under the Category column select **Fraction** and on the right side, under **Type**, select **Up to three digits**. Click on OK.

In cell A1 type 1/25, in A2 1/10 and in A3 1/20. Observe that these numbers appear on the spreadsheet as entered. After the last entry and pressing the <enter> key, the selected cell is A4. Now, click on the Σ (sum) symbol on the *Standard Toolbar*. Excel displays =SUM(A1:A3) in A4. Excel is asking us if we want to add the contents of A1 through A3 and display the answer in A4. Of course, this is what we want and we confirm it by pressing the <enter> key. Excel then displays the answer in A4 which is 19/100.

* No proficiency with Excel is required. However, the reader must have a Personal Computer (PC) or McIntosh, or access to one in which Excel has been installed. Otherwise, he may use a hand calculator which has the LCM feature.

Note A.8

For brevity, when using Excel in our subsequent discussion, we will omit the word cell and the key `<enter>`. Thus, B3, C11 etc., will be understood to be cell B3, cell C11 etc. After an entry in a cell has been made, it will be understood that the `<enter>` key was pressed. Also, we will denote sequential selections with the `>` (greater than) symbol. Thus, in the example above, our preference to display the numbers in rational form was accomplished by the sequential steps **Format>Cells>Fraction>Up to three digits>OK**.

In Examples A.23 and A.24 above, finding the common denominator was an easy task even without using Excel. However, in other cases, it may be a tedious process. Consider the following example.

Example A.25

Perform the following addition and express the answer in rational form.

$$\frac{3}{4} + \frac{2}{5} + \frac{4}{7} + \frac{11}{12} + \frac{7}{16} + \frac{19}{22} \tag{A.3}$$

Solution:

To find the LCM here without the use of a hand calculator or a PC, is a formidable task. Therefore, we will use Excel as we did in the previous example. Recall that LCM produces the common denominator of two or more numbers. The common multiplier for the numbers 4, 5, 7, 12, 16 and 22 is found as follows:

Begin with a blank Excel worksheet. Following the same procedure as in Example A.24, click on f_x in the Standard Toolbar. The Function Category menu appears. Select *Math & Trig*. Scroll down the function name list to locate LCM and click on it. In Number 1 box enter 4, 5, 7, 12, 16, 22 separated by commas, and observe that Excel displays 18480. This is the LCM for this example.

Now, we need to express each of the terms of (A.3) with the common denominator 18480. Of course, we must multiply the numerators of each term with the appropriate number so that numerical value of each term will not change. The new numerator value of each term is found by multiplying the present numerator by the quotient that is obtained with the division of 18480 by its respective denominator. Thus, we find the new numerator of $\frac{3}{4}$ by first dividing 18480 by 4; this division yields 4620. Then, we multiply this number by the present numerator of this term, that is, 3 and we get $3 \times 4620 = 13860$. Therefore, the number $\frac{3}{4}$ is expressed as $\frac{13860}{18480}$. In other words, we multiplied both numerator and denominator of $\frac{3}{4}$ by 4620. Likewise, to find the new

Appendix A Numbers and Arithmetic Operations

numerator of the term $\frac{2}{5}$, we first divide 18480 by 5 to get 3696. Then, multiplication of this number by 2 yields the new numerator of the second term which is $2 \times 3696 = 7392$, and therefore we express $\frac{2}{5}$ as $\frac{7392}{18480}$. Following the same procedure, we express $\frac{4}{7}$ as $\frac{10560}{18480}$, $\frac{11}{12}$ as $\frac{16940}{18480}$, $\frac{7}{16}$ as $\frac{8085}{18480}$ and $\frac{19}{22}$ as $\frac{15960}{18480}$. Finally, we find the sum of the given numbers in (A.3) as

$$\begin{aligned} & \frac{3}{4} + \frac{2}{5} + \frac{4}{7} + \frac{11}{12} + \frac{7}{16} + \frac{19}{22} \\ &= \frac{3 \times 4620}{4 \times 4620} + \frac{2 \times 3696}{5 \times 3696} + \frac{4 \times 2640}{7 \times 2640} + \frac{11 \times 1540}{12 \times 1540} + \frac{7 \times 1155}{16 \times 1155} + \frac{19 \times 840}{22 \times 840} \\ &= \frac{13860}{18480} + \frac{7392}{18480} + \frac{10560}{18480} + \frac{16940}{18480} + \frac{8085}{18480} + \frac{15960}{18480} = \frac{72797}{18480} \end{aligned}$$

This example required a considerable amount of work. This was the procedure one would follow before the advent of scientific calculators and personal computers. Let us now verify the result with Excel directly, that is, without first computing the LCM.

Start with a blank worksheet. Click on the letter A of Column A and observe that the entire column is now highlighted, that is, the background on this column has changed from white to black. Click on **Format > Cells > Fraction > Up to three digits > OK**.

In A1 through A6, type $3/4$, $2/5$, $4/7$, $11/12$, $7/16$ and $19/22$ respectively. Observe that these numbers appear on the spreadsheet as entered. After the last entry and pressing the $\langle \text{enter} \rangle$ key, the selected cell is A7. Now, click on the Σ (sum) symbol on the *Standard Toolbar*. Excel displays **=SUM(A1:A6)** in A7, and is asking us if we want to add the contents of A1 through A6 and display the answer in A7. Of course, this is what we want and we confirm it by pressing the $\langle \text{enter} \rangle$ key. Excel now displays the answer in A7 which is 3881/938. This value is an approximation since

Excel can only display rational numbers up to 3 digits. The exact value is $\frac{72797}{18480}$ as found above.

We will learn how to compute the percent error in Section A.14.

Note A.9

The reader, at this time, may ask why should one waste his time trying to find the solution to a problem when it can easily be found with a hand calculator or a PC. There is nothing wrong with this thought provided that we can predict the *approximate* answer beforehand, or we can determine whether the answer displayed is reasonably correct. We must remember that a PC is a high speed idiot and will do exactly whatever we will ask it to do. So, if we accidentally enter the wrong numbers, the answer we will get will definitely be wrong and unless we challenge it, we will accept

it as the correct answer. Likewise, with hand calculators; they display the wrong numbers when inaccurate numbers are entered, and also when the batteries are running low.

Example A.26

Perform the addition

$$\frac{1}{15} + \left(-\frac{1}{10}\right)^*$$

Solution:

The LCM for this example is 90. We can verify the answer with Excel as we did in the previous examples. This is left as an exercise for the reader. Then,

$$\frac{1}{15} + \left(-\frac{1}{10}\right) = \frac{1 \times 6}{15 \times 6} + \left(-\frac{1 \times 9}{10 \times 9}\right) = \frac{6}{90} + \left(-\frac{9}{90}\right) = -\frac{3}{90} = -\frac{1}{30}$$

Note A.10

In the last step above, we divided both the numerator and denominator of $-\frac{3}{90}$ by 3, or we can say that we multiplied each of these by $\frac{1}{3}$.

Example A.27

Perform the addition

$$-\frac{5}{12} + \frac{11}{24}$$

Solution:

The LCM for this example is 24. Then,

$$-\frac{5}{12} + \frac{11}{24} = \left(-\frac{5 \times 2}{12 \times 2}\right) + \frac{11 \times 1}{24 \times 1} = \left(-\frac{10}{24}\right) + \frac{11}{24} = \frac{1}{24}$$

* We must remember that the plus (+) and minus (-) signs, besides being used for addition and subtraction respectively, are also used to indicate positive and negative numbers respectively. Accordingly, it is recommended that negative numbers be enclosed in parentheses.

Appendix A Numbers and Arithmetic Operations


Example A.28

Perform the addition

$$0.75 + 0.005$$

Solution:

These numbers are given in decimal point (not rational) form. We can express them in rational form as $\frac{75}{100}$ and $\frac{5}{1000}$, find the LCM, which, in this case, is 1000, and add them as we did in Examples A.23 through A.27. However, it is easier to add them directly as they appear by aligning them with the decimal point, then add them as it is done in normal addition. This is shown below.

$$0.75 + 0.005 \longrightarrow \begin{array}{r} 0.750 \\ 0.005 \\ \hline 0.755 \end{array} +$$


We will do this simple example with Excel just to become more familiar with spreadsheets.

Begin with a blank worksheet or erase the contents of a previous worksheet. We erase the spreadsheet contents by surrounding all entries with a big rectangle that encloses all previous entries. When this is done properly, the selected cells within the rectangle will be highlighted. Press the *Delete* key on the keyboard and observe that all cells are now blank. Click on the letter A of Column A and observe that the entire column is now highlighted. Click on **Format > Cells > (tab) Number > (selection) Number** and the decimal places box click on the up or down arrow until the number 3 appears. Click on **OK**. On A1 type 0.750 and on A2 0.005. The selected cell is now A3. Next, click on the Σ (sum) symbol on the Standard Toolbar. Excel displays **=SUM(A1:A2)** in A3, and Excel is asking us if we want to add the contents of A1 and A2 in display the answer in A3. This is what we want and we confirm it by pressing the *<enter>* key. Excel then displays the answer in A3 which is 0.755.

Using Excel for this example, it was not necessary that the numbers are all displayed with the same number of decimal places. Excel will produce the correct answer if the numbers typed-in do not have the same number of decimal places. We chose to display them as we did just to become more familiar with Excel's format feature.

Rule 10: *To subtract a number given in rational form from another number of the same form, we first express the numbers in a common denominator; then, we subtract the numerator of the subtrahend from the numerator of the minuend to obtain the numerator of the difference. The denominator of the difference is the common denominator. If the numbers are in decimal point form, we first align the given numbers with the decimal point; then, we subtract the subtrahend from the minuend as it is done in normal subtraction.*

Example A.29

Perform the following subtractions.

a. $\frac{1}{10} - \frac{1}{20}$ b. $\frac{7}{25} - \left(-\frac{1}{10}\right)$ c. $\left(-\frac{5}{16}\right) - \frac{1}{24}$ d. $0.75 - 0.005$

Solution:

For (a), (b) and (c) we follow the same steps as we did in Examples A.23 through A.27, except that here we perform subtraction instead of addition. Thus,

a.

$$\frac{1}{10} - \frac{1}{20} = \frac{2}{20} - \frac{1}{20} = \frac{1}{20}$$

b.

$$\frac{7}{25} - \left(-\frac{1}{10}\right) = \frac{14}{50} + \frac{5}{50} = \frac{19}{50}$$

c.

$$\left(-\frac{5}{16}\right) - \frac{1}{24} = \left(-\frac{15}{48}\right) - \frac{2}{48} = \left(-\frac{17}{48}\right)$$

d. We follow the same steps as Example in A.28, except that we perform subtraction instead of addition. Then,

$$0.75 - 0.005 \longrightarrow \begin{array}{r} 0.750 \\ - 0.005 \\ \hline 0.745 \end{array}$$

Rule 11: *To multiply a number given in a rational form by another number of the same form, we multiply the numerators to obtain the numerator of the product. Then, we multiply the denominators to obtain the denominator of the product. If the numbers are in decimal point form, we multiply the given numbers as is done in normal multiplication, and we place the decimal point at the position that is equal to the sum of the decimal places of the given numbers.*

Example A.30

Perform the following multiplications.

a. $\frac{1}{12} \times \frac{1}{10}$ b. $\frac{3}{25} \times \left(-\frac{1}{10}\right)$ c. $\left(-\frac{23}{8}\right) \times \frac{1}{12}$ d. 0.75×0.005

Solution:

Appendix A Numbers and Arithmetic Operations

For each of (a), (b) and (c) we multiply the numerators to obtain the numerator of the product, then multiply the denominators to get the denominator of the product in accordance with Rule 11. Then,

a.

$$\frac{1}{12} \times \frac{1}{10} = \frac{1 \times 1}{12 \times 10} = \frac{1}{120}$$

b.

$$\frac{3}{25} \times \left(-\frac{1}{10}\right) = -\frac{3 \times 1}{25 \times 10} = -\frac{3}{250}$$

c.

$$\left(-\frac{23}{8}\right) \times \frac{1}{12} = -\frac{23 \times 1}{8 \times 12} = -\frac{23}{96}$$

d. We first perform the multiplication $75 \times 5 = 375$. Next, we place the decimal point at the position that is equal to the *sum* of the decimal places of the given numbers starting from the rightmost digit of the number we found, i.e., 375, and proceeding to the left. For this example, the multiplier has two decimal places and the multiplicand three; thus, their sum is five. The product is as shown below.

$$0.75 \times 0.005 \longrightarrow \begin{array}{r} 0.75 \times \\ 0.005 \\ \hline 0.00375 \end{array}$$

Rule 12: *To divide a number (dividend) given in a rational form, by another number (divisor) of the same form, we multiply the numerator of the dividend by the denominator of the divisor to obtain the numerator of the quotient. Then, we multiply the denominator of the dividend by the numerator of the divisor to obtain the denominator of the quotient. If the numbers are in decimal point form, we first express them in rational form and convert to integers by multiplying both numerator and denominator by the same smallest possible number that will make them integer numbers; then, we perform division as stated in Rule 5.*

Example A.31

Perform the following divisions.

$$\text{a. } \frac{1}{12} \div \frac{1}{10} \quad \text{b. } \frac{13}{12} \div \left(-\frac{1}{10}\right) \quad \text{c. } \left(-\frac{15}{8}\right) \div \frac{5}{12} \quad \text{d. } 0.75 \div 0.005$$

Solution:

For each of (a), (b) and (c) we multiply the numerator of the dividend by the denominator of the

divisor to obtain the numerator of the quotient; then we multiply the denominator of the dividend by the numerator of the divisor to get the denominator of the quotient in accordance with Rule 12. Then,

a.
$$\frac{1}{12} \div \frac{1}{10} = \frac{1 \times 10}{12 \times 1} = \frac{10}{12} = \frac{5}{6}$$

b.
$$\frac{13}{12} \div \left(-\frac{1}{10}\right) = -\frac{13 \times 10}{12 \times 1} = -\frac{130}{12} = -\frac{65}{6}$$

c.
$$\left(-\frac{15}{8}\right) \div \frac{5}{12} = -\frac{15 \times 12}{8 \times 5} = -\frac{180}{40} = -\frac{9}{2}$$

d. We first express the given numbers in rational form and we multiply both numerator and denominator by 1000 to convert them to integers. The division is carried in accordance with Rule 5. Then,

$$0.75 \div 0.005 \longrightarrow \frac{0.75 \times 1000}{0.005 \times 1000} = \frac{750}{5} = 150$$

Often, the division of two numbers in ratio form is expressed in complex fraction form. A *complex fraction* is a fraction whose numerator or denominator or both are fractions. The general form of a complex fraction is

$$\frac{\frac{X}{Y}}{\frac{W}{Z}}$$

Any complex fraction can be simplified by application of Rule 12 above twice. Alternately, any complex fraction can be replaced by a simple fraction whose numerator is the product of the outer numbers X and Z , and whose denominator is the product of the inner numbers Y and W , as shown in the sketch below.

$$\begin{array}{c} \text{Outer} \\ \text{Numbers} \end{array} \left\{ \begin{array}{l} \xrightarrow{\hspace{1.5cm}} \frac{X}{Y} \\ \xrightarrow{\hspace{1.5cm}} \frac{W}{Z} \end{array} \right. = \frac{XZ}{YW}$$

Inner
Numbers $\left\{ \begin{array}{l} \xrightarrow{\hspace{0.5cm}} \frac{X}{Y} \\ \xrightarrow{\hspace{0.5cm}} \frac{W}{Z} \end{array} \right.$

Appendix A Numbers and Arithmetic Operations

Example A.32

Simplify the following complex fractions.

$$\text{a. } \frac{\frac{3}{5}}{\frac{4}{7}} \quad \text{b. } \frac{5}{\frac{3}{4}} \quad \text{c. } \frac{1}{0.4}$$

Solution:

- a. We multiply the outer numbers, i.e., 3×7 to obtain the numerator and the inner numbers, i.e., 5×4 to obtain the denominator. Then,

$$\frac{\frac{3}{5}}{\frac{4}{7}} = \frac{3 \times 7}{5 \times 4} = \frac{21}{20}$$

- b. The numerator 5 of the given complex fraction can be expressed in rational form as $5/1$. Then, performing the same steps as in (a), we get

$$\frac{5}{\frac{3}{4}} = \frac{\frac{5}{1}}{\frac{3}{4}} = \frac{5 \times 4}{1 \times 3} = \frac{20}{3}$$

- c. The denominator 0.4 is first expressed as $\frac{4}{10}$. Then,

$$\frac{1}{0.4} = \frac{1}{\frac{4}{10}} = \frac{1 \times 10}{4 \times 1} = \frac{10}{4} = \frac{5}{2}$$

A.8 Exponents

In the mathematical expression $\mathbf{a}^{\mathbf{n}}$, \mathbf{a} is the *base*, \mathbf{n} is the *exponent* or *power*, and $\mathbf{a}^{\mathbf{n}}$ is said to be a *number in exponential form*. The exponent indicates the number of times the base appears in a string of multiplication operations as illustrated by the following example.

Example A.33

Express the following numbers in exponential form.

- a. $10 \times 10 \times 10 \times 10$ b. $5 \times 5 \times 5$ c. 3×3 d. $7 \times 7 \times 7 \times 8 \times 8$ e. 100, 000, 000

Solution:

- a. The number (base) 10 appears 4 times and thus the exponent is 4. Then,

$$10 \times 10 \times 10 \times 10 = 10^4$$

and it is read as *10 to the 4th power*. Here, 10 is the *base* and 4 is the *exponent* (or *power*).

- b. The number (base) 5 appears 3 times and thus the exponent is 3. Then,

$$5 \times 5 \times 5 = 5^3$$

and it is read as *5 to the 3rd power* or *5 cubed*. Here, 5 is the *base* and 3 is the *exponent*.

- c. The number (base) 3 appears 2 times and thus the exponent is 3. Then,

$$3 \times 3 = 3^2$$

and it is read as *3 to the 2nd power* or *3 squared*. Here, 3 is the *base* and 2 is the *exponent*.

- d.

$$7 \times 7 \times 7 \times 8 \times 8 = 7^3 \times 8^2$$

and it is read as *7 cubed times 8 squared*. Here, 7 and 8 are the *bases* and 3 and 2 are the *exponents*.

- e. As we now know, $100 = 10 \times 10 = 10^2$, $1000 = 10 \times 10 \times 10 = 10^3$ and so on. We observe that the exponent is equal to the number of zeros. Conversely, as we found in (a), the exponent 4 represents the four zeros of the number 10, 000. Here, the number 100, 000, 000 has 8 zeros and thus can be written in exponential form with base 10 and exponent 8, that is,

$$100, 000, 000 = 10^8$$

Rule 13: Any number written without an exponent is understood to have exponent 1.

Example A.34

Determine the exponents of the following numbers.

- a. 10 b. 3 c. -897 d. 0.06

Solution:

In accordance with Rule 13, the exponent of all these numbers is 1. Then,

Appendix A Numbers and Arithmetic Operations

a. 10^1 b. 3^1 c. -897^1 d. 0.06^1

Rule 14: Any number that has exponent 0 is equal to 1.

Example A.35

Express the following numbers in their simplest form.

a. 10^0 b. 8^0 c. -1093^0 d. 0.005^0

Solution:

In accordance with Rule 14, these numbers are all equal to 1, that is,

a. $10^0 = 1$ b. $8^0 = 1$ c. $-1093^0 = 1$ d. $0.005^0 = 1$

Rule 15: Any number may be written in powers of 10 multiplied by the appropriate number of digits.

Example A.36

Express the number 65,437 in powers of 10.

Solution:

$$\begin{aligned} 65437 &= 60,000 + 5,000 + 400 + 30 + 7 \\ &= 6 \times 10,000 + 5 \times 1,000 + 4 \times 100 + 3 \times 10 + 7 \\ &= 6 \times 10^4 + 5 \times 10^3 + 4 \times 10^2 + 3 \times 10^1 + 7 \times 10^0 \end{aligned}$$

We observe that the exponents appear in descending order, that is, 4, 3, 2, 1, 0. This is true for all numbers expressed as powers of 10.

Rule 16: Any number can be replaced by its reciprocal if the sign of its exponent is changed. Thus,

$$\frac{1}{X} = X^{-1} \quad \text{and} \quad Y = \frac{1}{Y^{-1}}$$

Example A.37

Express the following numbers in exponential form.

a. 0.1 b. 0.01 c. -0.001 d. 0.5 e. $\frac{1}{0.1}$ f. $\frac{1}{0.01}$ g. $-\frac{1}{0.001}$ h. $\frac{1}{0.5}$

Solution:

The given numbers can be expressed in exponential form by first expressing them in rational form and then applying Rules 13 through 16. Thus,

a.
$$0.1 = \frac{1}{10} = \frac{1}{10^1} = 10^{-1}$$

b.
$$0.01 = \frac{1}{100} = \frac{1}{10^2} = 10^{-2}$$

c.
$$-0.001 = -\frac{1}{1000} = -\frac{1}{10^3} = -10^{-3}$$

d.
$$0.5 = \frac{1}{2} = \frac{1}{2^1} = 2^{-1}$$

e.
$$\frac{1}{0.1} = \frac{1}{10^{-1}} = 10^1 = 10$$

f.
$$\frac{1}{0.01} = \frac{1}{10^{-2}} = 10^2 = 100$$

g.
$$-\frac{1}{0.001} = -\frac{1}{10^{-3}} = -10^3 = -1000$$

h.
$$\frac{1}{0.5} = \frac{1}{2^{-1}} = 2^1 = 2$$

For (e), (f), (g) and (h) we have used the results of (a), (b), (c) and (d) respectively.

A.9 Scientific Notation

In any number, the digit immediately to the left of the decimal point occupies the *units* position, the second digit occupies the *tens* position, the third the *hundreds* position, the fourth the *thousands* position and so on. The leftmost digit is the *most significant digit (msd)* and the rightmost is

Appendix A Numbers and Arithmetic Operations

the *least significant digit (lsd)*. Thus, in the number 7469.32, the msd is 7 and occupies the thousands position, 9 occupies the units position and 2 is the lsd.

Rule 17: *A number is said to be written in scientific notation when the decimal point appears immediately to the right of the msd.*

Therefore, any single digit number such as 1 or 8, or a two-digit number with a decimal point between the digits such as 6.5 or 1.9 is already in scientific notation. Other numbers can be written in scientific notation using the following rules:

Rule 18: *To write an integer or a mixed number in scientific notation, we move the decimal point to the left until we reach the msd. We place it immediately to the right of the msd. Then, we multiply the new number by 10 raised to the power that is equal to the number of positions the decimal point was moved to the left.*

Example A.38

Write the following numbers in scientific notation.

- a. 29 b. 765 c. 203982 d. -345.67 e. 98765.43

Solution:

In accordance with Rule 18, for (a) we move the decimal point to the left by one position and we multiply by 10 with exponent 1. Likewise, for (b) we move the decimal point to the left by two positions and we multiply by 10 with exponent 2. We follow the same procedure for (c), (d) and (e). Then,

a.
$$29 = 29. = 2.9 \times 10^1 = 2.9 \times 10$$

b.
$$765 = 765. = 7.65 \times 10^2$$

c.
$$203982 = 203982. = 2.03982 \times 10^5$$

d.
$$-345.67 = -3.4567 \times 10^2$$

e.
$$98765.43 = 9.876543 \times 10^4$$

Note A.11

Scientific calculators and Excel display numbers in scientific notation using the letter E followed by a number that represents the exponent. For instance, the number 98765.43 in scientific notation is displayed as 9.86543E+4. For practice, type in the number 98765.43 in any cell and high-

light that cell. Then, perform the sequential steps **Format>Cells>Scientific>6**(Decimal places)>**OK**. Observe that the number is displayed as 9.86543E+4.

Rule 19: *To write a fractional number in scientific notation, we move the decimal point to the right of the msd and we multiply the number by 10 raised to the negative power that is equal to the number of positions the decimal point was moved to the right.*

Example A.39

Write the following numbers in scientific notation.

a. 0.37 b. 0.045 c. -0.00123

Solution:

For (a), the msd is 3; therefore, in accordance with Rule 19, we move the decimal point to its right one position and we multiply by 10 with exponent -1 . For (b), the msd is 4 and thus we need to move two positions to the right; therefore, we multiply by 10 with exponent -2 . We follow the same procedure for (c). Then,

$$\text{a. } 0.37 = 3.7 \times 10^{-1} \quad \text{b. } 0.045 = 4.5 \times 10^{-2} \quad \text{c. } -0.00123 = -1.23 \times 10^{-3}$$

A.10 Operations with Numbers in Scientific Notation

Multiplication and division* of very large or very small numbers is simplified with application of the following rule:

Rule 20: *To multiply two numbers, we perform the following steps:*

1. Write the numbers in scientific notation.
2. Multiply the parts of the numbers without the 10s and their exponents.
3. Multiply the product of Step 2 above by 10 with the exponent obtained by the sum of the exponents of the numbers.

* *Addition and subtraction of numbers in scientific notation is impractical. Accordingly, no examples are given so that they will not cause confusion with multiplication and division. Nonetheless, the procedure is as follows:*

1. *Write the numbers to be added or subtracted in the same power of 10.*
2. *For addition, add the numbers without the 10s and their exponents. For subtraction, perform subtraction without the 10s and their exponents.*
3. *Multiply the result (sum or difference) by 10 raised to the same power of step 1.*

Appendix A Numbers and Arithmetic Operations

Example A.40

Multiply 4,000,000 by 23,000,000,000

Solution:

We first write the given numbers in scientific notation and in accordance with Rule 20, we first multiply 4 by 2.3 which results in the product 9.2. Then, we multiply this product by 10 raised to power 16 which is sum of the exponents of the multiplier and multiplicand as shown below.

$$\begin{array}{rcl} 4,000,000 & \longrightarrow & 4 \times 10^6 \\ 23,000,000,000 & \longrightarrow & 2.3 \times 10^{10} \times \\ \text{Product} & \longrightarrow & \underline{9.2 \times 10^{16}} \end{array}$$

Example A.41

Multiply 0.00029 by 0.000003

Solution:

We first write the given numbers in scientific notation and, in accordance with Rule 20, we first multiply 2.9 by 3 which results in the product 8.7. Then, we multiply this product by 10 raised to power -10 which is the sum of the exponents of the multiplier and multiplicand as shown below.

$$\begin{array}{rcl} 0.00029 & \longrightarrow & 2.9 \times 10^{-4} \\ 0.000003 & \longrightarrow & 3 \times 10^{-6} \times \\ \text{Product} & \longrightarrow & \underline{8.7 \times 10^{-10}} \end{array}$$

Four more examples follow.

a. Multiplication of 0.000025 by 8,000,000

$$\begin{array}{rcl} 0.000025 & \longrightarrow & 2.5 \times 10^{-5} \\ 8,000,000 & \longrightarrow & 8 \times 10^6 \times \\ \text{Product} & \longrightarrow & \underline{20 \times 10 = 2 \times 10^2} \end{array}$$

b. Multiplication of 80,000 by -0.0875

$$\begin{array}{rcl} 80,000 & \longrightarrow & 8 \times 10^4 \\ -0.0875 & \longrightarrow & -8.75 \times 10^{-2} \times \\ \text{Product} & \longrightarrow & \underline{-70 \times 10^2 = -7 \times 10^3} \end{array}$$

c. Multiplication of -250,000 by 0.025

$$\begin{array}{lcl}
 -250,000 & \longrightarrow & -2.5 \times 10^5 \\
 0.025 & \longrightarrow & \frac{2.5 \times 10^{-2}}{\phantom{2.5 \times 10^{-2}}} \\
 \text{Product} & \longrightarrow & -6.25 \times 10^2
 \end{array}$$

d. Multiplication of -0.00125 by -0.00008

$$\begin{array}{lcl}
 -0.00125 & \longrightarrow & -1.25 \times 10^{-3} \\
 -0.00008 & \longrightarrow & \frac{-8 \times 10^{-5}}{\phantom{-8 \times 10^{-5}}} \\
 \text{Product} & \longrightarrow & 10 \times 10^{-8} = 10^{-7}
 \end{array}$$

Rule 21: To divide two numbers, we perform the following steps:

1. Write the numbers in scientific notation
2. Divide the parts of the numbers without the 10s and their exponents
3. Multiply the product of Step 2 above by 10 raised to the power obtained by subtracting the exponent of the divisor from the exponent of the dividend.

Example A.42

Divide 45,000,000 by 1,500,000

Solution:

We first write the given numbers in scientific notation; then, in accordance with Rule 21, we first divide 4.5 by 1.5 which results in the quotient 3. Next, we multiply this product by 10 raised to power 1 which is obtained by subtracting the exponent 6 of the divisor from the exponent 7 of the dividend as shown below.

$$\begin{array}{lcl}
 \frac{45,000,000}{1,500,000} & \longrightarrow & \frac{4.5 \times 10^7}{1.5 \times 10^6} \\
 \text{Quotient} & \longrightarrow & 3 \times 10
 \end{array}$$

Two more examples follow.

$$\begin{array}{lcl}
 \frac{-2,000}{500,000} & \longrightarrow & \frac{-2 \times 10^3}{5 \times 10^5} \\
 \text{Quotient} & \longrightarrow & -4 \times 10^{-3}
 \end{array}$$

$$\begin{array}{lcl}
 \frac{-21,000,000}{-0.0000105} & \longrightarrow & \frac{-2.1 \times 10^7}{-1.05 \times 10^{-5}} \\
 \text{Quotient} & \longrightarrow & 2 \times 10^{12}
 \end{array}$$

Appendix A Numbers and Arithmetic Operations

Let us check the last division with Excel. We enter $-21,000,000$ in A1. We observe that Excel displays this number in scientific notation, i.e., $-2.1\text{E}+07$. We enter -0.0000105 in A2 and we observe that this number is also displayed in scientific notation as $-1.05\text{E}-05$. Generally, if the numbers are very large or very small, Excel displays them in scientific notation. Now, the selected cell is A3 where we type the formula $=\text{A1}/\text{A2}$; this formula instructs Excel to divide the contents of A1, i.e. the number $-21,000,000$, by the contents of A2, i.e., -0.0000105 and display the quotient in A3. We observe that Excel displays $2\text{E}+12$.

A.11 Square and Cubic Roots

The *square root* of a number is the number that, when it is squared, i.e., when multiplied by itself, produces the number under the square root. The square root of a non-negative number can be either positive or negative. Thus, the square root of 25 is ± 5 since both values $+5$ and -5 , when squared, will produce 25. *The square root of a negative number is an imaginary number.* We will not be concerned with the square roots of negative numbers. The square root of a number may be indicated by the square root symbol or with the fractional exponent $\frac{1}{2}$ or 0.5. Thus, the square root of 25 may be shown as $\sqrt{25}$, $25^{1/2}$, or $25^{0.5}$.

Example A.43

Evaluate the following square roots

a. $\sqrt{81}$ b. $\sqrt{225}$ c. $\sqrt{0.0625}$ d. $\sqrt{17.85}$

Solution:

a. The square root of 81 is ± 9 since $9 \times 9 = 81$ and $-9 \times (-9) = 81$ also. Thus,

$$\sqrt{81} = \pm 9$$

b. The square root of 225 is ± 15 since $15 \times 15 = 225$ and $-15 \times (-15) = 225$ also. Therefore,

$$\sqrt{225} = \pm 15$$

c. $0.25 \times 0.25 = 0.0625$ and $-0.25 \times (-0.25) = 0.0625$ also. Thus,

$$\sqrt{0.0625} = \pm 0.25$$

d. The answer is not obvious. However, since $\sqrt{25} = \pm 5$ and $\sqrt{16} = \pm 4$, we expect that the square root of 17.85 will be some value between ± 4 and ± 5 . There is a procedure for finding the square root of any number, but it is tedious and thus we will not discuss it here. It can be found in math tables books such as the *CRC Standard Mathematical Tables*, CRC Press.

Instead, we will use Excel as follows:

On the *Standard Toolbar* click on f_x . From the *Function Category* select *Math & Trig*. Scroll down the function name list to locate **SQRT** and click on it. Type 17.85 in the *Number* box and observe the answer displayed as **4.22493**. We get the same answer by typing **=SQRT(17.85)**. *Computers and calculators return absolute values only*. Thus,

$$\sqrt{17.85} = \pm 4.22493$$

The *cubic (or cube) root* of a number is a number that, when cubed, produces the number under the cubic root. The cubic root of a positive number is positive, whereas the cubic root of a negative number is negative. Thus, the cubic root of 125 is 5 since $5^3 = 125$, and the cubic root of -125 is -5 since $-5^3 = -125$. The cubic root of a number is represented by the cubic root symbol or with the fractional exponent $\frac{1}{3}$. Thus, the cubic root of 27 may be written as $\sqrt[3]{27}$ or $27^{1/3}$.

Example A.44

Evaluate the following cubic roots.

a. $\sqrt[3]{64}$ b. $\sqrt[3]{-343}$ c. $\sqrt[3]{0.003375}$ d. $\sqrt[3]{38.685}$

Solution:

a. The cubic root of 64 is +4 or simply 4 since $4 \times 4 \times 4 = 64$. It cannot also be -4 since $-4 \times (-4) \times (-4) = -64$. Thus,

$$\sqrt[3]{64} = 4$$

b. The cubic root of -343 is -7 since $-7 \times (-7) \times (-7) = -343$. Thus,

$$\sqrt[3]{-343} = -7$$

c. The answer is not obvious. There is a procedure for finding the cubic root of any number, but it is very tedious and thus, we will not discuss it here. Instead, we will use Excel as follows:

On the *Standard Toolbar*, we click on f_x . From the *Function Category*, we select *Math & Trig*. We scroll down the function name list to locate **POWER** and we click on it. We type 0.003375 in the *Number* box and 1/3 in the *Power* box. Excel returns 0.15. Thus,

$$\sqrt[3]{0.003375} = 0.15$$

Appendix A Numbers and Arithmetic Operations

- d. We will also use Excel. On the *Standard Toolbar*, we click on f_x . From the *Function Category*, we select *Math & Trig*. We scroll down the function name list to locate **POWER** and we click on it. We type 38.685 in the *Number* box and 1/3 in the *Power* box. Excel returns 3.382. Thus,

$$\sqrt[3]{38.685} = 3.382$$

This answer looks reasonably correct. Let us check it out by multiplication as follows:

$$3.382 \times 3.382 \times 3.382 = 38.683$$

A.12 Common and Natural Logarithms

The *common logarithm* or *base 10 logarithm* of a number, denoted as \log_{10} , is the *power (exponent)* to which the base 10 must be raised to produce that number. Thus, if a number is denoted as M and $\log_{10}M = x$, then $10^x = M$. For example, $\log_{10}100 = 2$ since $10^2 = 100$.

Another type of logarithm is the *natural* or *base e* logarithm where $e = 271828\dots$ and it is denoted as \ln or \log_e . It is the exponent to which the base e must be raised to produce that number. Thus, if a number is denoted as N and $\log_e N = \ln N = y$, then $e^y = N$. For example,

$$\ln 2 = \log_e 2 = 0.693$$

since

$$e^{0.693} = 2$$

The natural logarithm is mostly used in science and engineering, although it appears also in the computation of some financial functions, as we will see on the next chapter.

The logarithm (common or natural) of a positive number that is less than 1 is negative. The logarithm of a negative number is an imaginary number and, as mentioned before, we will not concern with these numbers.

Note A.12

For brevity, we will use \log to represent the common (base 10) logarithm, and \ln to represent the natural (base e) logarithm.

Example A.45

Evaluate the following logarithms given that $\ln 10 = 2.3026$ and $\ln 100 = 4.6052$.

- a. $\log 725$ b. $\log 0.125$ c. $\ln 64$ d. $\ln 0.75$
-

Solution:

- a. The answer must be a number between 2 and 3 because 725 lies between 100 and 1000, and by definition, $\log 100 = 2$ whereas $\log 1000 = 3$. We will use Excel to find the log of 725.

On the *Standard Toolbar*, we click on f_x . From the *Function Category*, we select *Math & Trig*. We scroll down the function name list to locate LOG10 and we click on it. We type 725 in the *Number* box and Excel returns 2.86034. Thus,

$$\log 725 = 2.86034$$

- b. The answer must be a negative number because we are seeking for the *log* of a number that is less than 1. We will use Excel to find the log of 0.125.

On the *Standard Toolbar*, we click on f_x . From the *Function Category*, we select *Math & Trig*. We scroll down the function name list to locate LOG10 and click on it. We type 0.125 in the *Number* box and Excel returns -0.90309. Thus,

$$\log 0.125 = -0.90309$$

- c. We make use of the fact that $\ln 10 = 2.3026$ and $\ln 100 = 4.6052$. Therefore, we expect $\ln 64$ to be a number greater than 2.3 and less than 4.6. Again, we will let Excel compute the answer for us.

On the *Standard Toolbar*, we click on f_x . From the *Function Category*, we select *Math & Trig*. We scroll down the function name list to locate LN, and we click on it. We type 64 in the *Number* box, and we observe the answer to be 4.15888. Thus,

$$\ln 64 = 4.15888$$

- d. The answer must be a negative number because we are seeking for the *ln* of a number less than 1. Let us again, use Excel to find the *ln* of 0.75.

On the *Standard Toolbar*, we click on f_x . From the *Function Category*, we select *Math & Trig*. We scroll down the function name list to locate LN, and we click on it. We type 0.75 in the *Number* box and observe the answer -0.28768. Thus,

$$\ln 0.75 = -0.28768$$

A.13 Decibel

A *decibel* (dB) is a unit that is used to express the relative difference in power or intensity, usually between two acoustic or electric signals. It is equal to ten times the common logarithm of the ratio of the two levels, that is,

$$\text{dB} = 10 \log \left| \frac{\text{Level 1}}{\text{Level 2}} \right| \quad (\text{A.4})$$

Appendix A Numbers and Arithmetic Operations

Example A.46

Compute the number of decibels (dB) if 90 represents Level 1 and 20 represents Level 2.

Solution:

$$\text{dB} = 10\log\left|\frac{90}{20}\right| = 10\log\left|\frac{9}{2}\right| = 10\log 4.5$$

and using Excel, we type =LOG(9/2) and Excel returns 0.6532. Then,

$$10\log 4.5 = 10 \times 0.6532 = 6.532 \text{ dB}$$

A.14 Percentages

Often, we hear that public securities, such as stocks and bonds, yield interest at a specified percentage, or banks paying or demanding interest at a specified percentage. *Percent*, denoted as %, is a number out of each hundred. For instance, the number 0.75 or 75/100 is expressed as 75% since it means 75 out of 100.

Most often, we are interested in changes of percentages. For instance, to find the percent change between an old and a new value, we use the formula

$$\% \text{ Change} = \frac{(\text{New Value} - \text{Old Value})}{\text{Old Value}} \times 100^* \quad (\text{A.5})$$

Example A.47

Joe Smith bought common stock of ABC Corporation at \$15.00 per share. Three months later he sold his stock at \$21.50 per share. What was his profit expressed in percent?

Solution:

Here, the old (initial) value is \$15.00 and the new is \$21.50. Then, using (A.5) we get

$$\% \text{ Change} = \frac{21.50 - 15.00}{15.00} \times 100 = \frac{6.50}{15.00} \times 100 = \frac{650}{15.00} = 43.33\%$$

In the example above, the percent change turned out to be positive. Other times, it can be negative as illustrated by the following example.

* The parentheses are used as a reminder that the subtraction must be performed before the multiplication by 100.

Example A.48

Jim Brown bought common stock at \$32.00 per share. Four months later he sold it at \$19.00 per share. What was his loss expressed in percent?

Solution:

Here, the old (initial) value is 32.00 and the new is 19.00. Then, using (A.5) we get

$$\% \text{ Change} = \frac{19.00 - 32.00}{32.00} \times 100 = \frac{-13.00}{32.00} \times 100 = \frac{-1300}{32.00} = -40.625\%$$

Other times, we want to find the percent error between exact (actual) and measured (experimental) values. In such as case, we express (A.5) as

$$\% \text{ Error} = \frac{\text{Measured Value} - \text{Exact Value}}{\text{Exact Value}} \times 100 \tag{A.6}$$

Example A.49

In Example A.25, Page A-15, we found that the exact value of the sum of the given numbers is $\frac{72797}{18480}$ and the value returned by Excel is $\frac{3881}{938}$. Compute the percent error.

Solution:

Let us use Excel to simplify these numbers.

We start with a blank worksheet, and in A1 enter =72797/18480. Excel displays 3.939232. In A2, we enter =3881/938 and we observe that Excel displays 4.137527. In A3, we enter the formula =(A2-A1)*100/A1. This is the formula of (A.6) above written in Excel language. Excel now displays 5.033851 and this is the percent error. Thus,

$$\% \text{ Error} = \frac{(4.137527 - 3.939232)}{3.939232} \times 100 = 5.03\%$$

A.15 International System of Units (SI)

The *International System of Units*, denoted as *SI* in all languages, was adopted by the General Conference on Weights and Measures in 1960. It is used extensively by the international scientific community. It was formerly known as the *Metric System*. The SI system basic units are listed in Table A.1.

TABLE A.1 SI Base Units

<i>Unit of</i>	<i>Name</i>	<i>Abbreviation</i>
Length	Metre	m
Mass	Kilogram	kg
Time	Second	s
Electric Current	Ampere	A
Temperature	Degrees Kelvin	°K
Amount of Substance	Mole	mol
Luminous Intensity	Candela	cd
Plane Angle	Radian	rad
Solid Angle	Steradian	sr

The basic units of the SI system are often expressed in smaller or larger units by various powers of 10 known as *standard prefixes*. The *common prefixes* are listed in Table A.2, and the less frequently used are listed in Table A.3.

TABLE A.2 Most Commonly Used SI Prefixes

<i>Value</i>	<i>Prefix</i>	<i>Symbol</i>	<i>Example</i>
10^9	Giga	G	12 GHz (Gigahertz) = 12×10^9 Hz
10^6	Mega	M	25 MW (Megaohms) = 25×10^6 W (ohms)
10^3	Kilo	K	13.2 KV (Kilovolts) = 13.2×10^3 volts
10^{-2}	centi	c	2.8 cm (centimeters) = 2.8×10^{-2} meter
10^{-3}	milli	m	4 mH (millihenries) = 4×10^{-3} henry
10^{-6}	micro	μ	$6 \mu\text{w}$ (microwatts) = 6×10^{-6} watt
10^{-9}	nano	n	2 ns (nanoseconds) = 2×10^{-9} second
10^{-12}	pico	p	3 pF (picofarads) = 3×10^{-12} Farad

TABLE A.3 Less Frequently Used SI Prefixes

Value	Prefix	Symbol	Example
10^{18}	Exa	E	1 Em (Exameter) = 10^{18} meters
10^{15}	Peta	P	5 Pyrs (Petayears) = 5×10^{15} years
10^{12}	Tera	T	3 T\$ (Teradollars) = 3×10^{12} dollars
10^{-15}	femto	f	7 fA (femtoamperes) = 7×10^{-15} ampere
10^{-18}	atto	a	9 aC (attocoulombs) = 9×10^{-18} coulomb

Table A.4 shows some conversion factors between the SI and the English system. For other conversions, the Microsoft Excel **CONVERT** function can be used. This will be discussed next.

TABLE A.4 Conversion Factors

1 in. (inch)	2.54 cm (centimeters)
1 mi. (mile)	1.609 Km (Kilometers)
1 lb. (pound)	0.4536 Kg (Kilograms)
1 qt. (quart)	946 cm ³ (cubic centimeters)
1 cm (centimeter)	0.3937 in. (inch)
1 Km (Kilometer)	0.6214 mi. (mile)
1 Kg (Kilogram)	2.2046 lbs (pounds)
1lt. (liter) = 1000 cm ³	1.057 quarts
1 Å (Angstrom)	10^{-10} meter
1 mm (micron)	10^{-6} meter

Table A.5 shows typical temperature values in degrees Fahrenheit and the equivalent temperature values in degrees Celsius and degrees Kelvin.

Other units used in physical sciences and electronics are derived from the SI base units and the most common are listed in Table A.6.

TABLE A.5 Temperature Scales Equivalents

°F	°C	°K
-523.4	-273	0
32	0	273
0	-17.8	255.2
77	25	298
98.6	37	310
212	100	373

TABLE A.6 SI Derived Units

Unit of	Name	Formula
Force	Newton (N)	$N = \text{Kg} \cdot \text{m}/\text{s}^2$
Pressure	Pascal (Pa)	$\text{Pa} = \text{N}/\text{m}^2$
Stress	Pascal (Pa)	$\text{Pa} = \text{N}/\text{m}^2$
Work	Joule (J)	$J = \text{N} \cdot \text{m}$
Energy	Joule (J)	$J = \text{N} \cdot \text{m}$
Power	Watt (W)	$W = \text{J}/\text{s}$
Voltage	Volt (V)	$V = \text{W}/\text{A}$
Resistance	Ohm (Ω)	$\Omega = \text{V}/\text{A}$
Conductance	Siemens (S)	$S = \text{A}/\text{V}$
Capacitance	Farad (F)	$F = \text{A} \cdot \text{s}/\text{V}$
Inductance	Henry (H)	$H = \text{V} \cdot \text{s}/\text{A}$
Frequency	Hertz (Hz)	$\text{Hz} = 1/\text{s}$
Quantity of Electricity	Coulomb (C)	$C = \text{A} \cdot \text{s}$
Magnetic Flux	Weber (Wb)	$\text{Wb} = \text{V} \cdot \text{s}$
Magnetic Flux Density	Tesla (T)	$T = \text{Wb}/\text{m}^2$
Luminous Flux	Lumen (lm)	$\text{lm} = \text{cd} \cdot \text{sr}$
Illuminance	Lux (lx)	$\text{lx} = \text{lm}/\text{m}^2$
Radioactivity	Becquerel (Bq)	$\text{Bq} = \text{s}^{-1}$
Radiation Dose	Gray (Gy)	$\text{Gy} = \text{J}/\text{Kg}$
Volume	Litre (L)	$L = \text{m}^3 \times 10^{-3}$

We conclude this appendix with a few examples, to illustrate the use of the CONVERT function in Excel.

Example A.50

Use the Excel CONVERT function to convert 54 °F to degrees Celsius.

Solution:

On the *Standard Toolbar*, we click on f_x . From the *Function Category*, we select *Engineering*. We scroll down the function name list to locate CONVERT, and we click on it. We type 54 in the *Number* box, F in the *From_Unit* box and C in the *To_Unit* box. We click on OK, and we observe the answer 12.22 °C rounded to two decimal places.

Example A.51

Use the Microsoft Excel CONVERT function to convert 42 °C to degrees Fahrenheit.

Solution:

On the *Standard Toolbar*, we click on f_x . From the *Function Category*, we select *Engineering*. We scroll down the function name list to locate CONVERT, and we click on it. We type 42 in the *Number* box, C in the *From_Unit* box and F in the *To_Unit* box. We click on OK, and we observe the answer 107.6 °F.

Note A.13

To find out what entries are allowed on the *From_Unit* and *To_unit* boxes, when using the CONVERT function, we click on *Help* on the *Standard Toolbar*, and on the pull-down menu, we click on *Contents and Index*. We select *Index* on the upper left corner, and we scroll down to *CONVERT Worksheet Function*. We click on *Display*, and we see a detailed description of the CONVERT function. We also see a list of text that CONVERT accepts as valid entries for the *From_Unit* and *To_unit* boxes.

Example A.52

Use the CONVERT function to convert 55 miles to kilometers.

Solution:

On the *Standard Toolbar* we click on f_x . From the *Function Category*, we select *Engineering*. We

scroll down the function name list to locate **CONVERT**, and we click on it. We type 55 in the *Number* box, *mi* in the *From_Unit* box and *m* in the *To_Unit* box. We click on **OK** and we observe the answer 88514 meters or 88.514 Km.

A.16 Graphs

A *graph* shows two or more related quantities in pictorial form. Two common forms of graphs are the *Cartesian* or *Rectangular form* and the *Polar form*. Figure A.2 shows a Cartesian form of a graph with the horizontal axis labeled *x axis* (*abscissa*) and the vertical axis labeled *y axis* (*ordinate*). The *abscissa* and *ordinate* are the *co-ordinates* of the Cartesian graph. The 0 (zero) point is the intersection of the co-ordinates.

All values to the right of the zero point, represent positive values of *x* while those to the left, represent negative values. Likewise, all values above the zero point represent positive values of *y*, while those below are negative values. Any point on the plane, formed by the *x* and *y* axes is designated as $P(x,y)$. Thus, point *A* is identified as $A(3,2)$ since it is located 3 units to the right of the origin (*zero*), and 2 units above the origin in the *y* axis direction. Likewise, point *B* is identified as $B(-2,4)$ since it is located 2 units to the left of the origin (*zero*) in the *x* axis, and 4 units above the origin in the *y* axis direction. Points *C* and *D* are identified similarly.

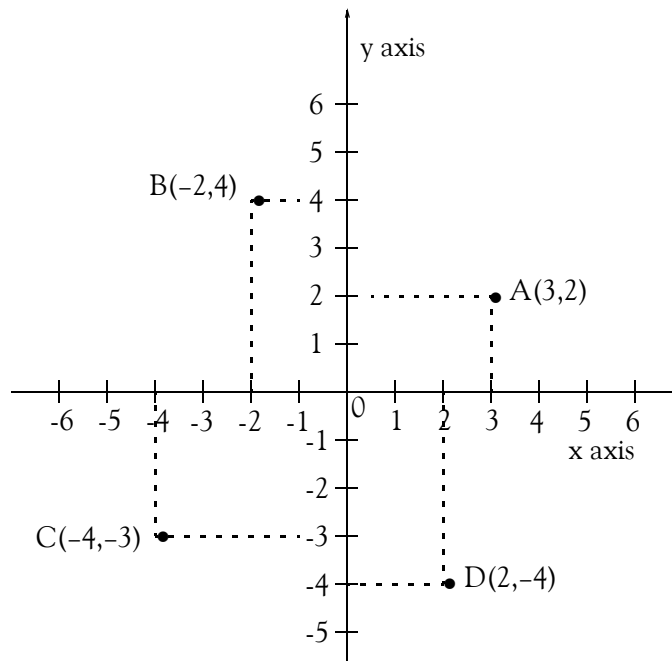
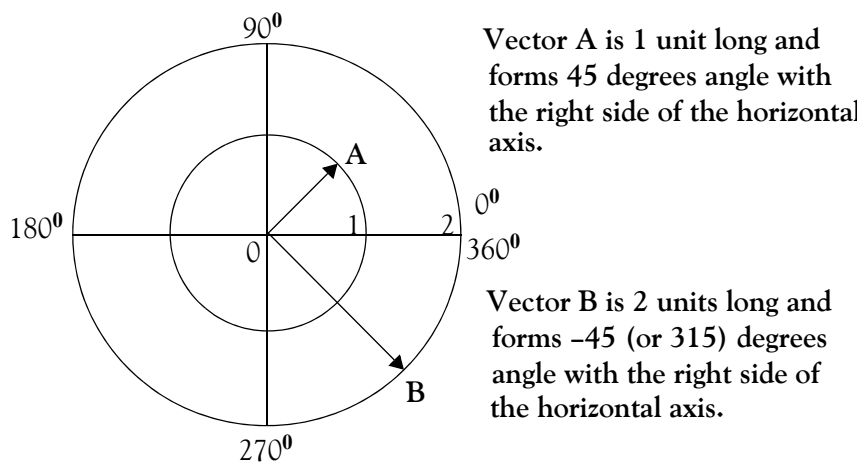


Figure A.2. Cartesian (rectangular) co-ordinates graph

Figure A.3 shows a *polar form* of a graph. Any point on this graph, is identified by its distance from the origin, and the angle that the line connecting the origin and that point, makes with the hori-

zontal line of the graph, starting at zero degrees point and rotating counterclockwise. The polar form is used in navigation and radar detection, and we will not be concerned with this type.



Direction of Rotation is Counterclockwise

Figure A.3. Polar graph

Any line from the origin to any other point is a vector. A *vector* is a line that is specified by its length from a reference point, normally the origin, and the angle it forms with a reference axis, normally the horizontal axis. Thus, **A** and **B** in Figure A.3 are vectors. A *phasor* is a rotating vector. For example, if vector **A** rotates (normally counterclockwise), then it is referred to as a *phasor*.

Figure A.4 shows a graph with a *constant value* of 12 units. This graph is referred to as a *uniform function*, since the number of units does not change with time. Thus, if the ordinate *v* represents a physical quantity such as the voltage of an automobile battery, we see that this voltage has the same value of 12 volts for all times.

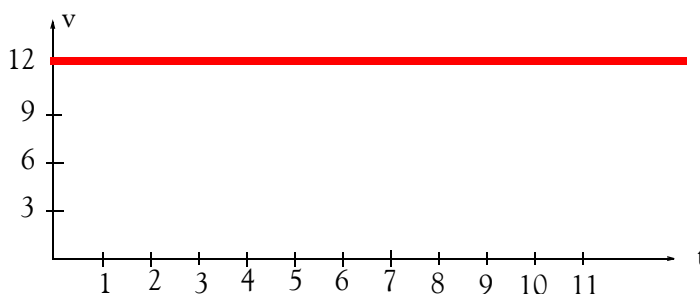


Figure A.4. Graph of a uniform function

Figure A.5 shows a *linear graph*. It is called linear because there is a linear (straight line) relationship between the quantities represented by the coordinates. Linear graphs have the property that equal increments in the abscissa produce equal increments on the ordinate axis.

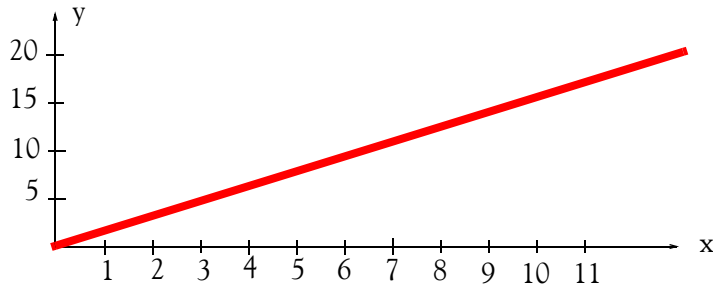


Figure A.5. Linear graph

Figure A.6 shows two *non-linear graphs* on the same set of a Cartesian form. As indicated, one is a *rising exponential*, and the other a *decaying exponential*. We will encounter a rising exponential curve in Chapter 9 when we discuss the exponential distribution. Decaying exponential curves are used in many applications such as in *radioactivity*, which is a topic discussed in physics and chemistry. For instance, the decay rate of radioactive elements can be represented by decaying exponentials. The isotope* *Thorium-234*, for example, decays to half of its original radioactivity in 25 days, and thus, it is said to have a half-life of 25 days. However, *Thorium-232* has a half-life of approximately 14 billion years.

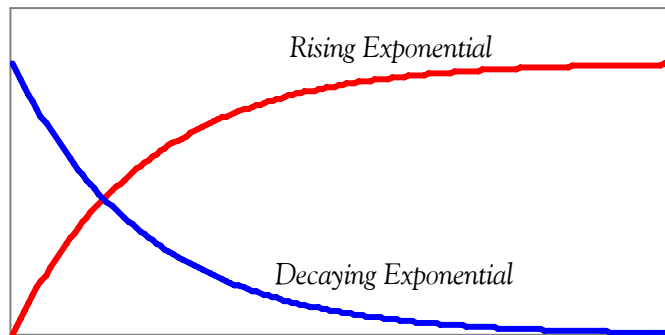


Figure A.6. Non-linear (exponential) graphs

Non-linear graphs are those that do not produce equal increments on the ordinate axis for equal increments on the abscissa.

Figure A.7 is a *sinewave graph*, which is also a non-linear graph. Sinewaves are used extensively in AC Electronics to represent time-varying quantities such as voltage and current.

* Isotopes are discussed in physics and chemistry textbooks. Briefly, an isotope is one of two or more atoms having the same atomic number but different mass numbers.

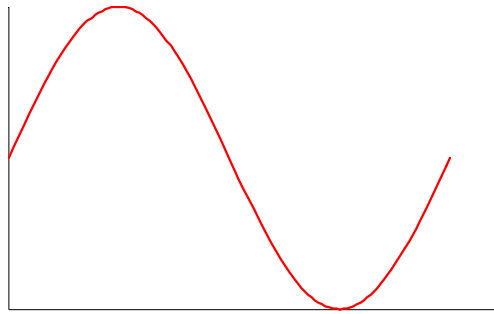


Figure A.7. Sinewave graph.

A.17 Summary

- The *decimal* (base 10) number system uses the *digits* (numbers) 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. This is the number system we use in our everyday arithmetic calculations such as the monetary transactions.
- A *positive number* is a number greater than zero and it is understood to have a plus (+) sign in front of it. The (+) sign in front of a positive number is generally omitted. Thus, any number without a sign in front of it is understood to be a positive number.
- A *negative number* is less than zero and it is written with a minus (-) sign in front of it. The minus (-) sign in front of a negative number is a must; it is necessary so that we can distinguish negative numbers from the positive numbers.
- Positive and negative numbers can be whole (integer) or fractional numbers. To avoid confusion between the addition operation (+) and positive numbers, which are also denoted with the (+) sign, we enclose positive numbers with their sign inside parentheses whenever necessary. Likewise, we enclose negative numbers in parentheses to distinguish them from the subtraction (-) symbol.
- The number 0 (*zero*) is considered neither positive nor negative; it is the number that separates the negative from the positive numbers.
- The positive and negative numbers we are familiar with, are referred to as the *real numbers*.
- The *absolute value* of a number is that number without a positive or negative sign, and is enclosed in small vertical lines. Thus, the absolute value of X is written as $|X|$.
- To add numbers with the same sign, we add the absolute values of these numbers and place the common sign (+ or -) in front of the result (sum). We can omit the plus sign in the result if positive. We must not omit the minus sign if the result is negative.
- For the subtraction $A-B$, the number A is called the *minuend* and the number B is called the *subtrahend*. The result of the subtraction is called the *difference*.
- To add two numbers with different signs, we subtract the number with the smaller absolute value from the number with the larger absolute value, and we place the sign of the larger number in front of the result (sum).
- To subtract one number (the subtrahend) from another number (the minuend) we change the sign (we replace + with - or - with +) of the subtrahend, and we perform addition instead of subtraction.
- Let X be any number; then, the following relations apply for the possible combinations of addition and subtraction of positive and negative numbers.

$$+(+X) = +X = X$$

$$+(-X) = -X$$

$$-(+X) = -X$$

$$-(-X) = X$$

- When two numbers with the same sign are multiplied, the product will be positive (+). If the numbers have different signs, the product will be negative (-).
- A *rational number* is a number that can be expressed as an integer or a quotient of integers, excluding zero as a denominator. In division, the number that is to be divided by another is called the *dividend*. For the ratio form A/B the dividend is A . The dividend is also referred to as the *numerator*. The number by which the dividend is to be divided, is called the *divisor*. For the ratio form A/B the divisor is B . The divisor is also called the *denominator*. The number obtained by dividing one quantity by another is the *quotient*.
- When two numbers with the same sign are divided, the quotient will be positive (+). If the numbers have different signs, the quotient will be negative (-).
- A *decimal point* is the dot in a number that is written in decimal form. Its purpose is to indicate where values change from positive to negative powers of 10. Every number has a decimal point although it may not be shown.
- A number is classified as an *integer*, *fractional* or *mixed* depending on the position of the decimal point in that number.
- An *integer* number is a whole number and it is understood that its decimal point is positioned after the last (rightmost) digit although not shown.
- A *fractional* number, positive or negative, is always a number less than one. It can be expressed in a rational form such as $3/4$ or in *decimal point form* as 0.75 . When a fractional number is written in decimal point form, the decimal point appears in front of the first (leftmost) digit.
- A *mixed number* consists of an integer and a fractional number separated by the decimal point. For instance, the number 409.0875 consists of the integer part 409 and the fractional part 0.0875 .
- Any number (integer, fractional or mixed) may be written in a rational form where the numerator (dividend) is the number itself and the denominator (divisor) is 1.
- The *reciprocal* of an integer number is a fraction with numerator 1 and denominator that number itself. Since any number can be considered as a fraction with denominator 1, the reciprocal of a fraction is another fraction with the numerator and denominator reversed. The product (multiplication) of any number by its reciprocal is 1.

Appendix A Numbers and Arithmetic Operations

- The value of a rational number is not changed if both numerator and denominator are multiplied by the same number.
- To add two or more fractional numbers in rational form, we first express the numbers in a common denominator; then, we add the numerators to obtain the numerator of the sum. The denominator of the sum is the common denominator. If the numbers are in decimal point form, we first align the given numbers with the decimal point, and we add the numbers as it is done in normal addition.
- The *Least Common Multiple (LCM)* is the smallest quantity that is divisible by two or more given quantities without a remainder.
- To subtract a number given in rational form from another number of the same form, we first express the numbers in a common denominator; then, we subtract the numerator of the subtrahend from the numerator of the minuend to obtain the numerator of the difference. The denominator of the difference is the common denominator. If the numbers are in decimal point form, we first align the given numbers with the decimal point; then, we subtract the subtrahend from the minuend as it is done in normal subtraction.
- To multiply a number given in a rational form by another number of the same form, we multiply the numerators to obtain the numerator of the product. Then, we multiply the denominators to obtain the denominator of the product. If the numbers are in decimal point form, we multiply the given numbers as is done in normal multiplication, and we place the decimal point at the position that is equal to the sum of the decimal places of the given numbers.
- To divide a number (dividend) given in a rational form, by another number (divisor) of the same form, we multiply the numerator of the dividend by the denominator of the divisor to obtain the numerator of the quotient. Then, we multiply the denominator of the dividend by the numerator of the divisor to obtain the denominator of the quotient. If the numbers are in decimal point form, we first express them in rational form and convert to integers by multiplying both numerator and denominator by the same smallest possible number that will make them integer numbers; then, we perform the division.
- A *complex fraction* is a fraction whose numerator or denominator or both are fractions. The general form of a complex fraction is

$$\frac{\frac{X}{Y}}{\frac{W}{Z}}$$

Any complex fraction can be replaced by a simple fraction whose numerator is the product of the outer numbers X and Z, and whose denominator is the product of the inner numbers Y and W, as shown in the sketch below.

$$\text{Outer Numbers} \left\{ \begin{array}{l} \text{Inner Numbers} \rightarrow \frac{X}{Y} \\ \text{Inner Numbers} \rightarrow \frac{W}{Z} \end{array} \right. = \frac{XZ}{YW}$$

- In the mathematical expression a^n , a is the *base*, n is the *exponent* or *power*, and a^n is said to be a *number in exponential form*. The exponent indicates the number of times the base appears in a string of multiplication operations.
- Any number written without an exponent is understood to have exponent 1.
- Any number that has exponent 0 is equal to 1.
- Any number may be written in powers of 10 multiplied by the appropriate number of digits.
- Any number can be replaced by its reciprocal if the sign of its exponent is changed.
- In any number, the digit immediately to the left of the decimal point occupies the *units* position, the second digit occupies the *tens* position, the third the *hundreds* position, the fourth the *thousands* position and so on. The leftmost digit is the *most significant digit (msd)* and the rightmost is the *least significant digit (lsd)*.
- A number is said to be written in scientific notation when the decimal point appears immediately to the right of the msd.
- To write an integer or a mixed number in scientific notation, we move the decimal point to the left until we reach the msd. We place it immediately to the right of the msd. Then, we multiply the new number by 10 raised to the power that is equal to the number of positions the decimal point was moved to the left.
- To write a fractional number in scientific notation, we move the decimal point to the right of the msd and we multiply the number by 10 raised to the negative power that is equal to the number of positions the decimal point was moved to the right.
- Multiplication of very large or very small numbers is simplified with application of the following rule:
 1. Write the numbers in scientific notation
 2. Multiply the parts of the numbers without the 10s and their exponents
 3. Multiply the product of Step 2 above by 10 with the exponent obtained by the sum of the exponents of the numbers

Appendix A Numbers and Arithmetic Operations

- Division of very large or very small numbers is simplified with application of the following rule:
 1. Write the numbers in scientific notation
 2. Divide the parts of the numbers without the 10s and their exponents
 3. Multiply the product of Step 2 above by 10 raised to the power obtained by subtracting the exponent of the divisor from the exponent of the dividend.
- The *square root* of a number is the number that, when it is squared, i.e., when multiplied by itself, produces the number under the square root. The square root of a non-negative number can be either positive or negative. *The square root of a negative number is an imaginary number.*
- The *cubic (or cube) root* of a number is a number that, when cubed, produces the number under the cubic root. The cubic root of a positive number is positive, whereas the cubic root of a negative number is negative. The cubic root of a number A is represented by the cubic root symbol $\sqrt[3]{A}$ or with the fractional exponent $\frac{1}{3}$, i.e., $A^{1/3}$.
- The *common logarithm* or *base 10 logarithm* of a number, denoted as \log_{10} , is the *power (exponent)* to which the base 10 must be raised to produce that number. Thus, if a number is denoted as M and $\log_{10}M = x$, then $10^x = M$.
- The *natural* or *base e* logarithm where $e = 271828\dots$ denoted as \ln or \log_e , is the exponent to which the base e must be raised to produce that number. Thus, if a number is denoted as N and $\log_e N = \ln N = y$, then $e^y = N$.
- The logarithm (common or natural) of a positive number that is less than 1 is negative. The logarithm of a negative number is an imaginary number.
- A *decibel (dB)* is a unit that is used to express the relative difference in power or intensity, usually between two acoustic or electric signals. It is equal to ten times the common logarithm of the ratio of the two levels, that is,

$$\text{dB} = 10 \log \left| \frac{\text{Level 1}}{\text{Level 2}} \right|$$

- *Percent*, denoted as %, is a number out of each hundred. The percent change between an old and a new value is found by application of the formula

$$\% \text{ Change} = \frac{(\text{New Value} - \text{Old Value})}{\text{Old Value}} \times 100$$

- The percent error between exact (actual) and measured (experimental) values can be found by application of the formula

$$\% \text{ Error} = \frac{\text{Measured Value} - \text{Exact Value}}{\text{Exact Value}} \times 100$$

- The *International System of Units*, denoted as *SI* in all languages, was adopted by the General Conference on Weights and Measures in 1960. It is used extensively by the international scientific community. It was formerly known as the *Metric System*.

A.18 Exercises

- Write the following numbers in scientific notation.
 - 54,000
 - 67,000,000
 - 89,000,000,000
 - 0.0045
 - 0.0000076
 - 0.00000000098
- Perform the following multiplications. The answers need not be in scientific notation.
 - $(1.2 \times 10^5) \times (2.1 \times 10^8)$
 - $(3.4 \times 10^7) \times (4.3 \times 10^{-2})$
 - $(5.6 \times 10^{-9}) \times (6.5 \times 10^6)$
 - $(7.9 \times 10^{-12}) \times (9.8 \times 10^{-6})$
- Perform the following divisions. The answers need not be in scientific notation.
 - $(5 \times 10^8) \div (1.25 \times 10^4)$
 - $(7.5 \times 10^3) \div (2.5 \times 10^9)$
 - $(8 \times 10^{-7}) \div (2 \times 10^5)$
 - $(9 \times 10^{-6}) \div (3 \times 10^{-9})$
- Perform the following temperature scale conversions. Round your answers to the nearest integer.
 - 40 °C to °F
 - 60 °F to °C
 - 50 °C to °K
 - 320 °K to °C
 - 90 °F to °K
 - 350 °K to °F

A.19 Solutions to End-of-Appendix Exercises

1.

a. $54,000 = 5.4 \times 10^4$ b. $67,000,000 = 6.7 \times 10^7$ c. $89,000,000,000 = 8.9 \times 10^{10}$

d. $0.0045 = 4.5 \times 10^{-3}$ e. $0.0000076 = 7.6 \times 10^{-6}$ f. $0.00000000098 = 9.8 \times 10^{-10}$

2.

a. $1.2 \times 2.1 = 2.52$, $5 + 8 = 13$, and thus $(1.2 \times 10^5) \times (2.1 \times 10^8) = 2.52 \times 10^{13}$

b. $3.4 \times 4.3 = 14.62$, $7 + (-2) = 5$, and thus $(3.4 \times 10^7) \times (4.3 \times 10^{-2}) = 14.62 \times 10^5$

c. $5.6 \times 6.5 = 36.4$, $-9 + 6 = 3$, and thus $(5.6 \times 10^{-9}) \times (6.5 \times 10^6) = 36.4 \times 10^3$

d. $7.9 \times 9.8 = 77.42$, $-12 + (-6) = -18$, and thus $(7.9 \times 10^{-12}) \times (9.8 \times 10^{-6}) = 77.42 \times 10^{-18}$

3.

a. $5/1.25 = 4$, $8 - 4 = 4$, and thus $(5 \times 10^8) \div (1.25 \times 10^4) = 4 \times 10^4$

b. $7.5/2.5 = 3$, $3 - 9 = -6$, and thus $(7.5 \times 10^3) \div (2.5 \times 10^9) = 3 \times 10^{-6}$

c. $8/2 = 4$, $-7 - 5 = -12$, and thus $(8 \times 10^{-7}) \div (2 \times 10^5) = 4 \times 10^{-12}$

d. $9/3 = 3$, $-6 - (-9) = 3$, and thus $(9 \times 10^{-6}) \div (3 \times 10^{-9}) = 3 \times 10^3$

4. Using the procedures that we used in Examples A.50 and A.51 we get:

a. `=CONVERT(40,"C","F")=104 °F`

b. `=CONVERT(60,"F","C")= 16 °C`

c. `=CONVERT(50,"C","K")= 323 °K`

d. `=CONVERT(320,"K","C")= 47 °C`

e. `=CONVERT(90,"F","K")=305 °K`

f. `=CONVERT(350,"K","F")=170 °F`

The solutions of Exercises 4.a and 4.b above can also be implemented with the Simulink* model of Figure A.8, and the solutions of Exercises 4.e and 4.f can be implemented with the model of Figure A.9.

* For an introduction to Simulink, please refer to *Introduction to Simulink with Engineering Applications*, ISBN 0-9744239-7-1.

Appendix A Numbers and Arithmetic Operations

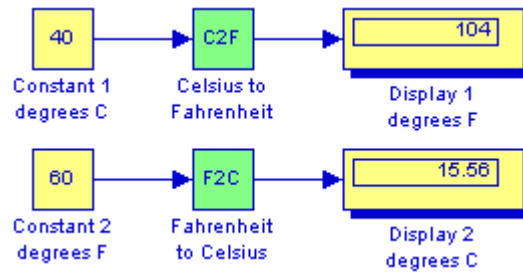


Figure A.8. Model for the solutions of Exercises 4.a and 4.b

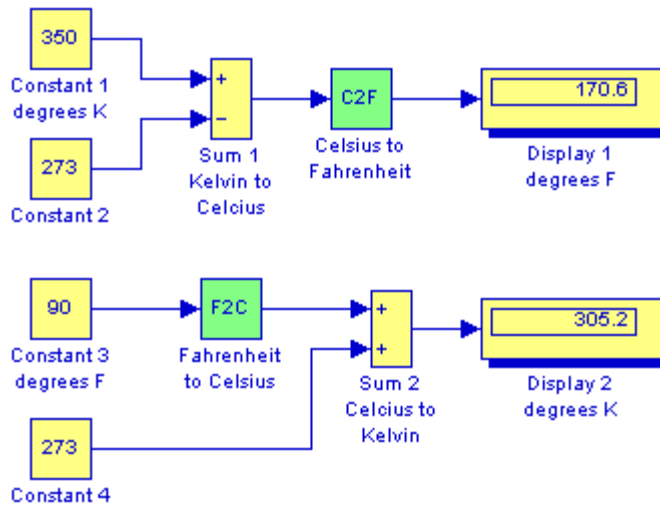


Figure A.9. Model for the solutions of Exercises 4.e and 4.f

This appendix serves as an introduction to the basic MATLAB commands and functions, procedures for naming and saving the user generated files, comment lines, access to MATLAB's Editor / Debugger, finding the roots of a polynomial, and making plots. Several examples are provided with detailed explanations.

B.1 MATLAB® and Simulink®

MATLAB and Simulink are products of The MathWorks,™ Inc. These are two outstanding software packages for scientific and engineering computations and are used in educational institutions and in industries including automotive, aerospace, electronics, telecommunications, and environmental applications. MATLAB enables us to solve many advanced numerical problems rapidly and efficiently.

B.2 Command Window

To distinguish the screen displays from the user commands, important terms, and MATLAB functions, we will use the following conventions:

Click: Click the left button of the mouse

Courier Font: Screen displays

Helvetica Font: User inputs at MATLAB's command window prompt >> or EDU>>*

Helvetica Bold: MATLAB functions

Times Bold Italic: Important terms and facts, notes and file names

When we first start MATLAB, we see various help topics and other information. Initially, we are interested in the *command screen* which can be selected from the Window drop menu. When the command screen, we see the prompt >> or EDU>>. This prompt is displayed also after execution of a command; MATLAB now waits for a new command from the user. It is highly recommended that we use the *Editor/Debugger* to write our program, save it, and return to the command screen to execute the program as explained below.

To use the Editor/Debugger:

1. From the *File* menu on the toolbar, we choose *New* and click on *M-File*. This takes us to the *Editor Window* where we can type our *script* (list of statements) for a new file, or open a previously saved file. We must save our program with a file name which starts with a letter. **Impor-**

* EDU>> is the MATLAB prompt in the Student Version

tant! MATLAB is *case sensitive*, that is, it distinguishes between upper- and lower-case letters. Thus, *t* and *T* are two different letters in MATLAB language. The files that we create are saved with the file name we use and the extension *.m*; for example, *myfile01.m*. It is a good practice to save the script in a file name that is descriptive of our script content. For instance, if the script performs some matrix operations, we ought to name and save that file as *matrices01.m* or any other similar name. We should also use a floppy disk or an external drive to backup our files.

2. Once the script is written and saved as an *m-file*, we may exit the *Editor/Debugger* window by clicking on *Exit Editor/Debugger* of the *File* menu. MATLAB then returns to the command window.
3. To execute a program, we type the file name *without* the *.m* extension at the `>>` prompt; then, we press `<enter>` and observe the execution and the values obtained from it. If we have saved our file in drive *a* or any other drive, we must make sure that it is added to the desired directory in MATLAB's search path. The MATLAB User's Guide provides more information on this topic.

Henceforth, it will be understood that each input command is typed after the `>>` prompt and followed by the `<enter>` key.

The command `help matlab\iofun` will display input/output information. To get help with other MATLAB topics, we can type `help` followed by any topic from the displayed menu. For example, to get information on graphics, we type `help matlab\graphics`. The MATLAB User's Guide contains numerous help topics.

To appreciate MATLAB's capabilities, we type `demo` and we see the MATLAB Demos menu. We can do this periodically to become familiar with them. Whenever we want to return to the command window, we click on the `Close` button.

When we are done and want to leave MATLAB, we type `quit` or `exit`. But if we want to clear all previous values, variables, and equations without exiting, we should use the command `clear`. This command erases everything; it is like exiting MATLAB and starting it again. The command `clc` clears the screen but MATLAB still remembers all values, variables and equations that we have already used. In other words, if we want to clear all previously entered commands, leaving only the `>>` prompt on the upper left of the screen, we use the `clc` command.

All text after the `%` (percent) symbol is interpreted as a *comment line* by MATLAB, and thus it is ignored during the execution of a program. A comment can be typed on the same line as the function or command or as a separate line. For instance,

```
conv(p,q) % performs multiplication of polynomials p and q
```

```
% The next statement performs partial fraction expansion of p(x) / q(x)
```

are both correct.

One of the most powerful features of MATLAB is the ability to do computations involving *complex numbers*. We can use either *i*, or *j* to denote the imaginary part of a complex number, such as $3-4i$ or $3-4j$. For example, the statement

```
z=3-4j
```

displays

```
z = 3.0000-4.0000i
```

In the above example, a multiplication (*) sign between 4 and j was not necessary because the complex number consists of numerical constants. However, if the imaginary part is a function, or variable such as $\cos(x)$, we must use the multiplication sign, that is, we must type `cos(x)*j` or `j*cos(x)` for the imaginary part of the complex number.

B.3 Roots of Polynomials

In MATLAB, a polynomial is expressed as a *row vector* of the form $[a_n \ a_{n-1} \ \dots \ a_2 \ a_1 \ a_0]$. These are the coefficients of the polynomial in descending order. **We must include terms whose coefficients are zero.**

We find the roots of any polynomial with the **roots(p)** function; **p** is a row vector containing the polynomial coefficients in descending order.

Example B.1

Find the roots of the polynomial

$$p_1(x) = x^4 - 10x^3 + 35x^2 - 50x + 24$$

Solution:

The roots are found with the following two statements where we have denoted the polynomial as **p1**, and the roots as **roots_p1**.

```
p1=[1 -10 35 -50 24] % Specify and display the coefficients of p1(x)
```

```
p1 =
     1    -10     35    -50     24
```

```
roots_p1=roots(p1) % Find the roots of p1(x)
```

```
roots_p1 =
     4.0000
     3.0000
     2.0000
     1.0000
```

We observe that MATLAB displays the polynomial coefficients as a row vector, and the roots as a column vector.

Example B.2

Find the roots of the polynomial

$$p_2(x) = x^5 - 7x^4 + 16x^2 + 25x + 52$$

Solution:

There is no cube term; therefore, we must enter zero as its coefficient. The roots are found with the statements below, where we have defined the polynomial as **p2**, and the roots of this polynomial as **roots_p2**. The result indicates that this polynomial has three real roots, and two complex roots. Of course, complex roots always occur in *complex conjugate** pairs.

```
p2=[1 -7 0 16 25 52]
```

```
p2 =  
     1     -7         0         16         25         52
```

```
roots_p2=roots(p2)
```

```
roots_p2 =  
     6.5014  
     2.7428  
    -1.5711  
    -0.3366 + 1.3202i  
    -0.3366 - 1.3202i
```

B.4 Polynomial Construction from Known Roots

We can compute the coefficients of a polynomial, from a given set of roots, with the **poly(r)** function where **r** is a row vector containing the roots.

Example B.3

It is known that the roots of a polynomial are 1, 2, 3, and 4. Compute the coefficients of this polynomial.

* By definition, the conjugate of a complex number $A = a + jb$ is $A^* = a - jb$

Solution:

We first define a row vector, say $r3$, with the given roots as elements of this vector; then, we find the coefficients with the **poly(r)** function as shown below.

```
r3=[1 2 3 4]      % Specify the roots of the polynomial
```

```
r3 =
```

```
     1     2     3     4
```

```
poly_r3=poly(r3)  % Find the polynomial coefficients
```

```
poly_r3 =
```

```
     1    -10     35    -50     24
```

We observe that these are the coefficients of the polynomial $p_1(x)$ of Example B.1.

Example B.4

It is known that the roots of a polynomial are -1 , -2 , -3 , $4 + j5$, and $4 - j5$. Find the coefficients of this polynomial.

Solution:

We form a row vector, say $r4$, with the given roots, and we find the polynomial coefficients with the **poly(r)** function as shown below.

```
r4=[-1 -2 -3 4+5j 4-5j]
```

```
r4 =
```

```
Columns 1 through 4
```

```
-1.0000    -2.0000   -3.0000   -4.0000+ 5.0000i
```

```
Column 5
```

```
-4.0000- 5.0000i
```

```
poly_r4=poly(r4)
```

```
poly_r4 =
```

```
     1     14     100     340     499     246
```

Therefore, the polynomial is

$$p_4(x) = x^5 + 14x^4 + 100x^3 + 340x^2 + 499x + 246$$

B.5 Evaluation of a Polynomial at Specified Values

The **polyval(p,x)** function evaluates a polynomial $p(x)$ at some specified value of the indepen-

dent variable x .

Example B.5

Evaluate the polynomial

$$p_5(x) = x^6 - 3x^5 + 5x^3 - 4x^2 + 3x + 2 \quad (\text{B.1})$$

at $x = -3$.

Solution:

```
p5=[1 -3 0 5 -4 3 2]; % These are the coefficients of the given polynomial
% The semicolon (;) after the right bracket suppresses the
% display of the row vector that contains the coefficients of p5.
%
val_minus3=polyval(p5, -3) % Evaluate p5 at x=-3; no semicolon is used here
% because we want the answer to be displayed

val_minus3 =
    1280
```

Other MATLAB functions used with polynomials are the following:

conv(a,b) – multiplies two polynomials **a** and **b**

[q,r]=deconv(c,d) –divides polynomial **c** by polynomial **d** and displays the quotient **q** and remainder **r**.

polyder(p) – produces the coefficients of the derivative of a polynomial **p**.

Example B.6

Let

$$p_1 = x^5 - 3x^4 + 5x^2 + 7x + 9$$

and

$$p_2 = 2x^6 - 8x^4 + 4x^2 + 10x + 12$$

Compute the product $p_1 \cdot p_2$ using the **conv(a,b)** function.

Solution:

```
p1=[1 -3 0 5 7 9]; % The coefficients of p1
p2=[2 0 -8 0 4 10 12]; % The coefficients of p2
p1p2=conv(p1,p2) % Multiply p1 by p2 to compute coefficients of the product p1p2
```

$$p_1 p_2 = \begin{matrix} 2 & -6 & -8 & 34 & 18 & -24 & -74 & -88 & 78 & 166 & 174 & 108 \end{matrix}$$

Therefore,

$$p_1 \cdot p_2 = 2x^{11} - 6x^{10} - 8x^9 + 34x^8 + 18x^7 - 24x^6 - 74x^5 - 88x^4 + 78x^3 + 166x^2 + 174x + 108$$

Example B.7

Let

$$p_3 = x^7 - 3x^5 + 5x^3 + 7x + 9$$

and

$$p_4 = 2x^6 - 8x^5 + 4x^2 + 10x + 12$$

Compute the quotient p_3/p_4 using the **[q,r]=deconv(c,d)** function.

Solution:

% It is permissible to write two or more statements in one line separated by semicolons
 p3=[1 0 -3 0 5 7 9]; p4=[2 -8 0 0 4 10 12]; [q,r]=deconv(p3,p4)

$$q = \begin{matrix} 0.5000 \\ 0 & 4 & -3 & 0 & 3 & 2 & 3 \end{matrix}$$

Therefore,

$$q = 0.5 \quad r = 4x^5 - 3x^4 + 3x^2 + 2x + 3$$

Example B.8

Let

$$p_5 = 2x^6 - 8x^4 + 4x^2 + 10x + 12$$

Compute the derivative $\frac{d}{dx}p_5$ using the **polyder(p)** function.

Solution:

p5=[2 0 -8 0 4 10 12]; % The coefficients of p5
 der_p5=polyder(p5) % Compute the coefficients of the derivative of p5

$$der_p5 = \begin{matrix} 12 & 0 & -32 & 0 & 8 & 10 \end{matrix}$$

Therefore,

$$\frac{d}{dx}p_5 = 12x^5 - 32x^3 + 4x^2 + 8x + 10$$

B.6 Rational Polynomials

Rational Polynomials are those which can be expressed in ratio form, that is, as

$$R(x) = \frac{\text{Num}(x)}{\text{Den}(x)} = \frac{b_n x^n + b_{n-1} x^{n-1} + b_{n-2} x^{n-2} + \dots + b_1 x + b_0}{a_m x^m + a_{m-1} x^{m-1} + a_{m-2} x^{m-2} + \dots + a_1 x + a_0} \quad (\text{B.2})$$

where some of the terms in the numerator and/or denominator may be zero. We can find the roots of the numerator and denominator with the **roots(p)** function as before.

As noted in the comment line of Example B.7, we can write MATLAB statements in one line, if we separate them by commas or semicolons. **Commas will display the results whereas semicolons will suppress the display.**

Example B.9

Let

$$R(x) = \frac{p_{\text{num}}}{p_{\text{den}}} = \frac{x^5 - 3x^4 + 5x^2 + 7x + 9}{x^6 - 4x^4 + 2x^2 + 5x + 6}$$

Express the numerator and denominator in factored form, using the **roots(p)** function.

Solution:

```
num=[1 -3 0 5 7 9]; den=[1 0 -4 0 2 5 6]; % Do not display num and den coefficients
roots_num=roots(num), roots_den=roots(den) % Display num and den roots
```

```
roots_num =
    2.4186 + 1.0712i    2.4186 - 1.0712i   -1.1633
   -0.3370 + 0.9961i   -0.3370 - 0.9961i
```

```
roots_den =
    1.6760 + 0.4922i    1.6760 - 0.4922i   -1.9304
   -0.2108 + 0.9870i   -0.2108 - 0.9870i   -1.0000
```

As expected, the complex roots occur in complex conjugate pairs.

For the numerator, we have the factored form

$$p_{\text{num}} = (x-2.4186 - j1.0712)(x-2.4186 + j1.0712)(x + 1.1633) \\ (x + 0.3370 - j0.9961)(x + 0.3370 + j0.9961)$$

and for the denominator, we have

$$p_{\text{den}} = (x-1.6760 - j0.4922)(x-1.6760 + j0.4922)(x + 1.9304) \\ (x + 0.2108 - j0.9870)(x + 0.2108 + j0.9870)(x + 1.0000)$$

We can also express the numerator and denominator of this rational function as a combination of *linear* and *quadratic* factors. We recall that, in a quadratic equation of the form $x^2 + bx + c = 0$ whose roots are x_1 and x_2 , the negative sum of the roots is equal to the coefficient b of the x term, that is, $-(x_1 + x_2) = b$, while the product of the roots is equal to the constant term c , that is, $x_1 \cdot x_2 = c$. Accordingly, we form the coefficient b by addition of the complex conjugate roots and this is done by inspection; then we multiply the complex conjugate roots to obtain the constant term c using MATLAB as follows:

$$(2.4186 + 1.0712i)*(2.4186 - 1.0712i)$$

$$\text{ans} = 6.9971$$

$$(-0.3370 + 0.9961i)*(-0.3370 - 0.9961i)$$

$$\text{ans} = 1.1058$$

$$(1.6760 + 0.4922i)*(1.6760 - 0.4922i)$$

$$\text{ans} = 3.0512$$

$$(-0.2108 + 0.9870i)*(-0.2108 - 0.9870i)$$

$$\text{ans} = 1.0186$$

Thus,

$$R(x) = \frac{p_{\text{num}}}{p_{\text{den}}} = \frac{(x^2 - 4.8372x + 6.9971)(x^2 + 0.6740x + 1.1058)(x + 1.1633)}{(x^2 - 3.3520x + 3.0512)(x^2 + 0.4216x + 1.0186)(x + 1.0000)(x + 1.9304)}$$

We can check this result of Example B.9 above with MATLAB's *Symbolic Math Toolbox* which is a collection of tools (functions) used in solving symbolic expressions. They are discussed in detail in MATLAB's Users Manual. For the present, our interest is in using the **collect(s)** function that is used to multiply two or more symbolic expressions to obtain the result in polynomial form. We must remember that the **conv(p,q)** function is used with numeric expressions only, that is, polynomial coefficients.

Before using a symbolic expression, we must create one or more symbolic variables such as x, y, t ,

and so on. For our example, we use the following script:

```
syms x % Define a symbolic variable and use collect(s) to express numerator in polynomial form
collect((x^2-4.8372*x+6.9971)*(x^2+0.6740*x+1.1058)*(x+1.1633))
```

```
ans =
x^5-29999/10000*x^4-1323/3125000*x^3+7813277909/
1562500000*x^2+1750276323053/250000000000*x+4500454743147/
500000000000
```

and if we simplify this, we find that is the same as the numerator of the given rational expression in polynomial form. We can use the same procedure to verify the denominator.

B.7 Using MATLAB to Make Plots

Quite often, we want to plot a set of ordered pairs. This is a very easy task with the MATLAB **plot(x,y)** command that plots y versus x , where x is the horizontal axis (abscissa) and y is the vertical axis (ordinate).

Example B.10

Consider the electric circuit of Figure B.1, where the radian frequency ω (radians/second) of the applied voltage was varied from 300 to 3000 in steps of 100 radians/second, while the amplitude was held constant.

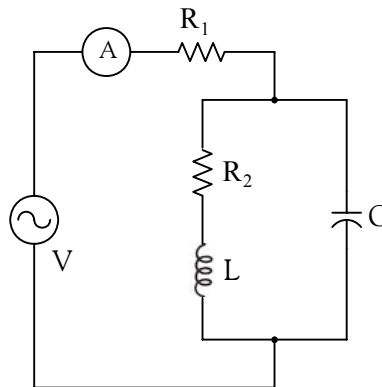


Figure B.1. Electric circuit for Example B.10

The ammeter readings were then recorded for each frequency. The magnitude of the impedance $|Z|$ was computed as $|Z| = |V/A|$ and the data were tabulated in Table B.1.

TABLE B.1 Table for Example B.10

ω (rads/s)	Z Ohms	ω (rads/s)	Z Ohms
300	39.339	1700	90.603
400	52.589	1800	81.088
500	71.184	1900	73.588
600	97.665	2000	67.513
700	140.437	2100	62.481
800	222.182	2200	58.240
900	436.056	2300	54.611
1000	1014.938	2400	51.428
1100	469.83	2500	48.717
1200	266.032	2600	46.286
1300	187.052	2700	44.122
1400	145.751	2800	42.182
1500	120.353	2900	40.432
1600	103.111	3000	38.845

Plot the magnitude of the impedance, that is, $|Z|$ versus radian frequency ω .

Solution:

We cannot type ω (omega) in the MATLAB Command prompt, so we will use the English letter w instead.

If a statement, or a row vector is too long to fit in one line, it can be continued to the next line by typing three or more periods, then pressing `<enter>` to start a new line, and continue to enter data. This is illustrated below for the data of w and z . Also, as mentioned before, we use the semicolon (;) to suppress the display of numbers that we do not care to see on the screen.

The data are entered as follows:

```
w=[300 400 500 600 700 800 900 1000 1100 1200 1300 1400 1500 1600 1700 1800 1900...
2000 2100 2200 2300 2400 2500 2600 2700 2800 2900 3000];
%
z=[39.339 52.789 71.104 97.665 140.437 222.182 436.056...
1014.938 469.830 266.032 187.052 145.751 120.353 103.111...
90.603 81.088 73.588 67.513 62.481 58.240 54.611 51.468...
48.717 46.286 44.122 42.182 40.432 38.845];
```

Of course, if we want to see the values of w or z or both, we simply type w or z , and we press `<enter>`. To plot z (y-axis) versus w (x-axis), we use the `plot(x,y)` command. For this example, we use `plot(w,z)`. When this command is executed, MATLAB displays the plot on MATLAB's *graph screen* and MATLAB denotes this plot as Figure 1. This plot is shown in Figure B.2.

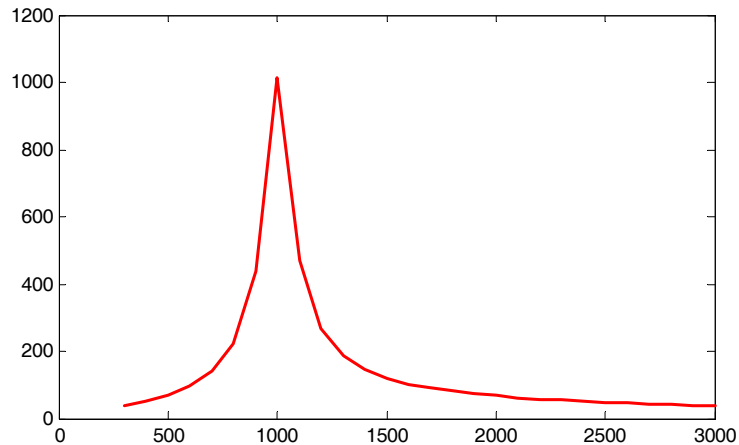


Figure B.2. Plot of impedance $|z|$ versus frequency ω for Example B.10

This plot is referred to as the *magnitude frequency response* of the circuit.

To return to the command window, we press any key, or from the *Window* pull-down menu, we select *MATLAB Command Window*. To see the graph again, we click on the *Window* pull-down menu, and we choose *Figure 1*.

We can make the above, or any plot, more presentable with the following commands:

grid on: This command adds grid lines to the plot. The **grid off** command removes the grid. The command **grid** toggles them, that is, changes from off to on or vice versa. The default* is off.

box off: This command removes the box (the solid lines which enclose the plot), and **box on** restores the box. The command **box** toggles them. The default is on.

title('string'): This command adds a line of the text **string** (label) at the top of the plot.

xlabel('string') and **ylabel('string')** are used to label the x- and y-axis respectively.

The magnitude frequency response is usually represented with the x-axis in a logarithmic scale. We can use the **semilogx(x,y)** command which is similar to the **plot(x,y)** command, except that the x-axis is represented as a log scale, and the y-axis as a linear scale. Likewise, the **semilogy(x,y)** command is similar to the **plot(x,y)** command, except that the y-axis is represented as a log scale, and the x-axis as a linear scale. The **loglog(x,y)** command uses logarithmic scales for both axes.

Throughout this text it will be understood that **log** is the common (base 10) logarithm, and **ln** is

* A default is a particular value for a variable that is assigned automatically by an operating system and remains in effect unless canceled or overridden by the operator.

the natural (base e) logarithm. We must remember, however, the function **log(x)** in MATLAB is the natural logarithm, whereas the common logarithm is expressed as **log10(x)**, and the logarithm to the base 2 as **log2(x)**.

Let us now redraw the plot with the above options by adding the following statements:

```
semilogx(w,z); grid;           % Replaces the plot(w,z) command
title('Magnitude of Impedance vs. Radian Frequency');
xlabel('w in rads/sec'); ylabel('|Z| in Ohms')
```

After execution of these commands, the plot is as shown in Figure B.3.

If the y-axis represents power, voltage or current, the x-axis of the frequency response is more often shown in a logarithmic scale, and the y-axis in dB (decibels).

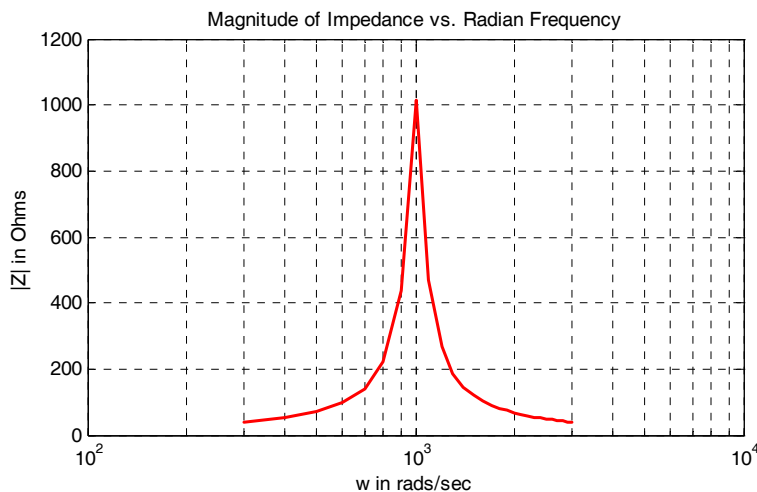


Figure B.3. Modified frequency response plot of Figure B.2.

To display the voltage v in a dB scale on the y-axis, we add the relation $\text{dB}=20*\log_{10}(v)$, and we replace the **semilogx(w,z)** command with **semilogx(w,dB)**.

The command **gtext('string')*** switches to the current *Figure Window*, and displays a cross-hair that can be moved around with the mouse. For instance, we can use the command **gtext('Impedance |Z| versus Frequency')**, and this will place a cross-hair in the *Figure* window. Then, using the mouse, we can move the cross-hair to the position where we want our label to begin, and we press <enter>.

* With the latest MATLAB Versions 6 and 7 (Student Editions 13 and 14), we can add text, lines and arrows directly into the graph using the tools provided on the *Figure Window*. For advanced MATLAB graphics, please refer to *The Math-Works Using MATLAB Graphics* documentation.

The command **text(x,y,'string')** is similar to **gtext('string')**. It places a label on a plot in some specific location specified by **x** and **y**, and **string** is the label which we want to place at that location. We will illustrate its use with the following example which plots a 3-phase sinusoidal waveform.

The first line of the script below has the form

linspace(first_value, last_value, number_of_values)

This function specifies *the number of data points* but not the increments between data points. An alternate function is

x=first: increment: last

and this specifies *the increments between points* but not the number of data points.

The script for the 3-phase plot is as follows:

```
x=linspace(0, 2*pi, 60); % pi is a built-in function in MATLAB;  
% we could have used x=0:0.02*pi:2*pi or x = (0: 0.02: 2)*pi instead;  
y=sin(x); u=sin(x+2*pi/3); v=sin(x+4*pi/3);  
plot(x,y,x,u,x,v); % The x-axis must be specified for each function  
grid on, box on, % turn grid and axes box on  
text(0.75, 0.65, 'sin(x)'); text(2.85, 0.65, 'sin(x+2*pi/3)'); text(4.95, 0.65, 'sin(x+4*pi/3)')
```

These three waveforms are shown on the same plot of Figure B.4.

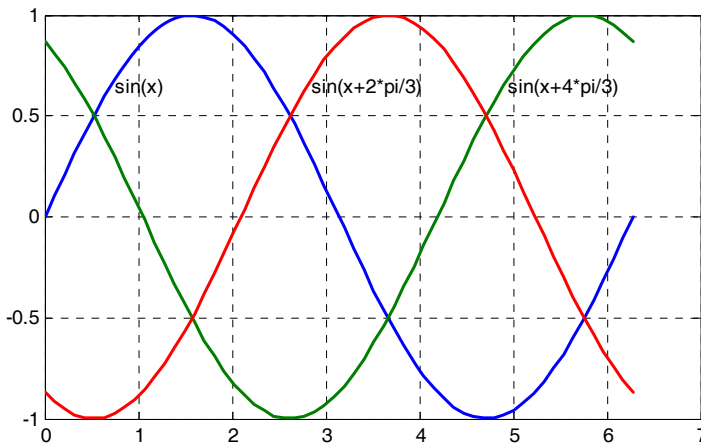


Figure B.4. Three-phase waveforms

In our previous examples, we did not specify line styles, markers, and colors for our plots. However, MATLAB allows us to specify various line types, plot symbols, and colors. These, or a combination of these, can be added with the **plot(x,y,s)** command, where **s** is a character string containing one or more characters shown on the three columns of Table B.2. MATLAB has no default color; it starts with blue and cycles through the first seven colors listed in Table B.2 for each additional line in the plot. Also, there is no default marker; no markers are drawn unless they are selected. The default line is the solid line. But with the latest MATLAB versions, we can select the line color, line width, and other options directly from the *Figure Window*.

TABLE B.2 Styles, colors, and markets used in MATLAB

Symbol	Color	Symbol	Marker	Symbol	Line Style
b	blue	.	point	—	solid line
g	green	o	circle	:	dotted line
r	red	x	x-mark	-.	dash-dot line
c	cyan	+	plus	---	dashed line
m	magenta	*	star		
y	yellow	s	square		
k	black	d	diamond		
w	white	v	triangle down		
		^	triangle up		
		<	triangle left		
		>	triangle right		
		p	pentagram		
		h	hexagram		

For example, **plot(x,y,'m*:')** plots a magenta dotted line with a star at each data point, and **plot(x,y,'rs')** plots a red square at each data point, but does not draw any line because no line was selected. If we want to connect the data points with a solid line, we must type **plot(x,y,'rs-')**. For additional information we can type **help plot** in MATLAB's command screen.

The plots we have discussed thus far are two-dimensional, that is, they are drawn on two axes. MATLAB has also a three-dimensional (three-axes) capability and this is discussed next.

The **plot3(x,y,z)** command plots a line in 3-space through the points whose coordinates are the elements of *x*, *y* and *z*, where *x*, *y* and *z* are three vectors of the same length.

The general format is **plot3(x₁,y₁,z₁,s₁,x₂,y₂,z₂,s₂,x₃,y₃,z₃,s₃,...)** where **x_n**, **y_n** and **z_n** are vectors or matrices, and **s_n** are strings specifying color, marker symbol, or line style. These strings are the same as those of the two-dimensional plots.

Example B.11

Plot the function

$$z = -2x^3 + x + 3y^2 - 1 \quad (\text{B.3})$$

Solution:

We arbitrarily choose the interval (length) shown on the script below.

```
x = -10: 0.5: 10;           % Length of vector x
y = x;                     % Length of vector y must be same as x
z = -2.*x.^3+x+3.*y.^2-1; % Vector z is function of both x and y*
plot3(x,y,z); grid
```

The three-dimensional plot is shown in Figure B.5.

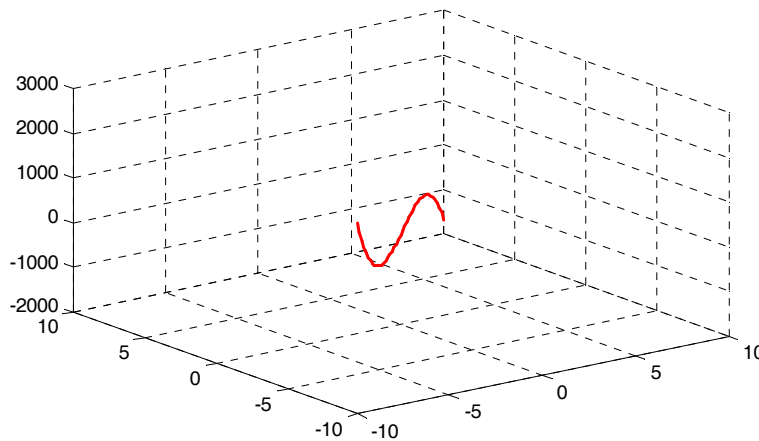


Figure B.5. Three dimensional plot for Example B.11

In a two-dimensional plot, we can set the limits of the x - and y -axes with the **axis([xmin xmax ymin ymax])** command. Likewise, in a three-dimensional plot we can set the limits of all three axes with the **axis([xmin xmax ymin ymax zmin zmax])** command. It must be placed after the **plot(x,y)** or **plot3(x,y,z)** commands, or on the same line without first executing the **plot** command. This must be done for each plot. The three-dimensional **text(x,y,z,'string')** command will place **string** beginning at the co-ordinate (x,y,z) on the plot.

For three-dimensional plots, **grid on** and **box off** are the default states.

* This statement uses the so called dot multiplication, dot division, and dot exponentiation where the multiplication, division, and exponential operators are preceded by a dot. These important operations will be explained in Section B.9.

We can also use the **mesh(x,y,z)** command with two vector arguments. These must be defined as $\text{length}(x) = n$ and $\text{length}(y) = m$ where $[m, n] = \text{size}(Z)$. In this case, the vertices of the mesh lines are the triples $\{x(j), y(i), Z(i, j)\}$. We observe that **x** corresponds to the columns of *Z*, and **y** corresponds to the rows.

To produce a mesh plot of a function of two variables, say $z = f(x, y)$, we must first generate the *X* and *Y* matrices that consist of repeated rows and columns over the range of the variables *x* and *y*. We can generate the matrices *X* and *Y* with the **[X,Y]=meshgrid(x,y)** function that creates the matrix *X* whose rows are copies of the vector **x**, and the matrix *Y* whose columns are copies of the vector **y**.

Example B.12

The volume *V* of a right circular cone of radius *r* and height *h* is given by

$$V = \frac{1}{3}\pi r^2 h \tag{B.4}$$

Plot the volume of the cone as *r* and *h* vary on the intervals $0 \leq r \leq 4$ and $0 \leq h \leq 6$ meters.

Solution:

The volume of the cone is a function of both the radius *r* and the height *h*, that is,

$$V = f(r, h)$$

The three-dimensional plot is created with the following MATLAB script where, as in the previous example, in the second line we have used the dot multiplication, dot division, and dot exponentiation. This will be explained in Section B.9.

```
[R,H]=meshgrid(0: 4, 0: 6); % Creates R and H matrices from vectors r and h;...
V=(pi .* R.^ 2 .* H) ./ 3; mesh(R, H, V);...
xlabel('x-axis, radius r (meters)'); ylabel('y-axis, altitude h (meters)');...
zlabel('z-axis, volume (cubic meters)'); title('Volume of Right Circular Cone'); box on
```

The three-dimensional plot of Figure B.6 shows how the volume of the cone increases as the radius and height are increased.

The plots of Figure B.5 and B.6 are rudimentary; MATLAB can generate very sophisticated three-dimensional plots. The MATLAB User's Manual and the Using MATLAB Graphics Manual contain numerous examples.

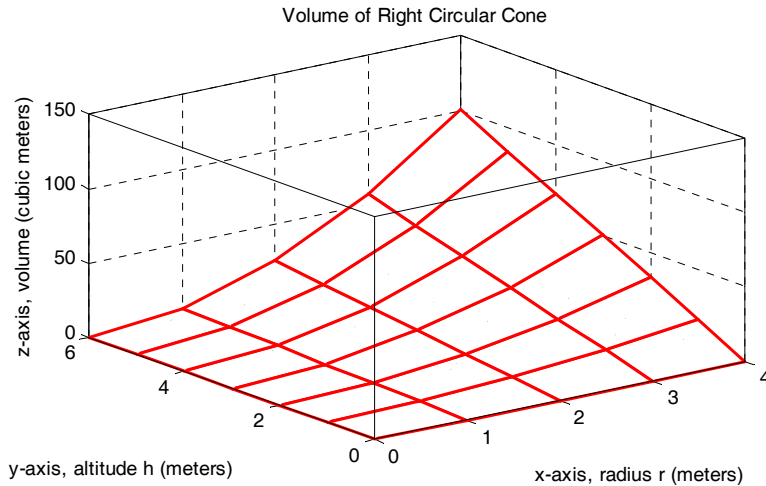


Figure B.6. Volume of a right circular cone.

B.8 Subplots

MATLAB can display up to four windows of different plots on the *Figure* window using the command **subplot(m,n,p)**. This command divides the window into an $m \times n$ matrix of plotting areas and chooses the p th area to be active. No spaces or commas are required between the three integers m , n and p . The possible combinations are shown in Figure B.7.

We will illustrate the use of the **subplot(m,n,p)** command following the discussion on multiplication, division and exponentiation that follows.

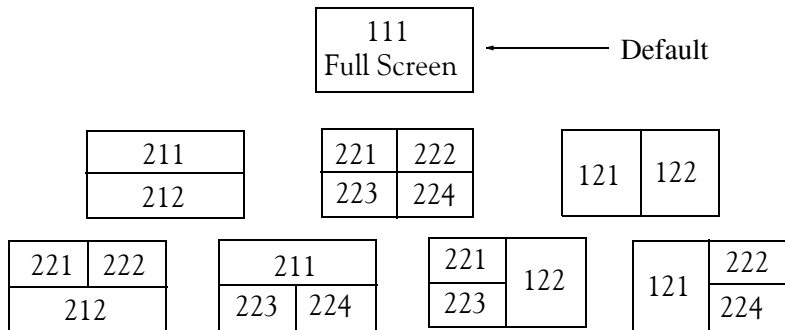


Figure B.7. Possible subplot arrangements in MATLAB

B.9 Multiplication, Division, and Exponentiation

MATLAB recognizes two types of multiplication, division, and exponentiation. These are the **matrix** multiplication, division, and exponentiation, and the **element-by-element** multiplication, division, and exponentiation. They are explained in the following paragraphs.

In Section B.2, the arrays [a b c ...], such as those that contained the coefficients of polynomials, consisted of one row and multiple columns, and thus are called **row vectors**. If an array has one column and multiple rows, it is called a **column vector**. We recall that the elements of a row vector are separated by spaces. To distinguish between row and column vectors, the elements of a column vector must be separated by semicolons. An easier way to construct a column vector, is to write it first as a row vector, and then transpose it into a column vector. MATLAB uses the single quotation character (') to transpose a vector. Thus, a column vector can be written either as

```
b=[-1; 3; 6; 11]
```

or as

```
b=[-1 3 6 11]'
```

As shown below, MATLAB produces the same display with either format.

```
b=[-1; 3; 6; 11]
```

```
b =  
    -1  
     3  
     6  
    11
```

```
b=[-1 3 6 11]'      % Observe the single quotation character (')
```

```
b =  
    -1  
     3  
     6  
    11
```

We will now define Matrix Multiplication and Element-by-Element multiplication.

1. Matrix Multiplication (multiplication of row by column vectors)

Let

$$\mathbf{A} = [a_1 \ a_2 \ a_3 \ \dots \ a_n]$$

and

$$\mathbf{B} = [b_1 \ b_2 \ b_3 \ \dots \ b_n]'$$

be two vectors. We observe that \mathbf{A} is defined as a row vector whereas \mathbf{B} is defined as a column vector, as indicated by the transpose operator ('). Here, multiplication of the row vector \mathbf{A} by the column vector \mathbf{B} , is performed with the matrix multiplication operator (*). Then,

$$\mathbf{A} * \mathbf{B} = [a_1 b_1 + a_2 b_2 + a_3 b_3 + \dots + a_n b_n] = \text{single value} \quad (\text{B.5})$$

For example, if

$$\mathbf{A} = [1 \ 2 \ 3 \ 4 \ 5]$$

and

$$\mathbf{B} = [-2 \ 6 \ -3 \ 8 \ 7]'$$

the matrix multiplication $\mathbf{A}*\mathbf{B}$ produces the single value 68, that is,

$$\mathbf{A}*\mathbf{B} = 1 \times (-2) + 2 \times 6 + 3 \times (-3) + 4 \times 8 + 5 \times 7 = 68$$

and this is verified with the MATLAB script

```
A=[1 2 3 4 5]; B=[-2 6 -3 8 7]'; A*B % Observe transpose operator (') in B
ans =
    68
```

Now, let us suppose that both \mathbf{A} and \mathbf{B} are row vectors, and we attempt to perform a row-by-row multiplication with the following MATLAB statements.

```
A=[1 2 3 4 5]; B=[-2 6 -3 8 7]; A*B % No transpose operator (') here
```

When these statements are executed, MATLAB displays the following message:

```
??? Error using ==> *

```

```
Inner matrix dimensions must agree.
```

Here, because we have used the matrix multiplication operator (*) in $\mathbf{A}*\mathbf{B}$, MATLAB expects vector \mathbf{B} to be a column vector, not a row vector. It recognizes that \mathbf{B} is a row vector, and warns us that we cannot perform this multiplication using the matrix multiplication operator (*). Accordingly, we must perform this type of multiplication with a different operator. This operator is defined below.

2. Element-by-Element Multiplication (multiplication of a row vector by another row vector)

Let

$$\mathbf{C} = [c_1 \ c_2 \ c_3 \ \dots \ c_n]$$

and

$$\mathbf{D} = [d_1 \ d_2 \ d_3 \ \dots \ d_n]$$

be two row vectors. Here, multiplication of the row vector \mathbf{C} by the row vector \mathbf{D} is performed with the *dot multiplication operator* ($.*$). There is no space between the dot and the multiplication symbol. Thus,

$$\mathbf{C}.*\mathbf{D} = [c_1d_1 \ c_2d_2 \ c_3d_3 \ \dots \ c_nd_n] \quad (\text{B.6})$$

This product is another row vector with the same number of elements, as the elements of \mathbf{C}

and **D**.

As an example, let

$$\mathbf{C} = [1 \ 2 \ 3 \ 4 \ 5]$$

and

$$\mathbf{D} = [-2 \ 6 \ -3 \ 8 \ 7]$$

Dot multiplication of these two row vectors produce the following result.

$$\mathbf{C}.*\mathbf{D} = 1 \times (-2) \quad 2 \times 6 \quad 3 \times (-3) \quad 4 \times 8 \quad 5 \times 7 = -2 \quad 12 \quad -9 \quad 32 \quad 35$$

Check with MATLAB:

```
C=[1 2 3 4 5]; % Vectors C and D must have
D=[-2 6 -3 8 7]; % same number of elements
C.*D % We observe that this is a dot multiplication
ans =
    -2    12    -9    32    35
```

Similarly, the division (/) and exponentiation (^) operators, are used for matrix division and exponentiation, whereas dot division (./) and dot exponentiation (.^) are used for element-by-element division and exponentiation, as illustrated in Examples B.11 and B.12 above.

We must remember that *no space is allowed between the dot (.) and the multiplication, division, and exponentiation operators.*

Note: A dot (.) is never required with the plus (+) and minus (-) operators.

Example B.13

Write the MATLAB script that produces a simple plot for the waveform defined as

$$y = f(t) = 3e^{-4t} \cos 5t - 2e^{-3t} \sin 2t + \frac{t^2}{t+1} \tag{B.7}$$

in the $0 \leq t \leq 5$ seconds interval.

Solution:

The MATLAB script for this example is as follows:

```
t=0: 0.01: 5; % Define t-axis in 0.01 increments
y=3 .* exp(-4 .* t) .* cos(5 .* t)-2 .* exp(-3 .* t) .* sin(2 .* t) + t.^2 ./ (t+1);
plot(t,y); grid; xlabel('t'); ylabel('y=f(t)'); title('Plot for Example B.13')
```

The plot for this example is shown in Figure B.8.

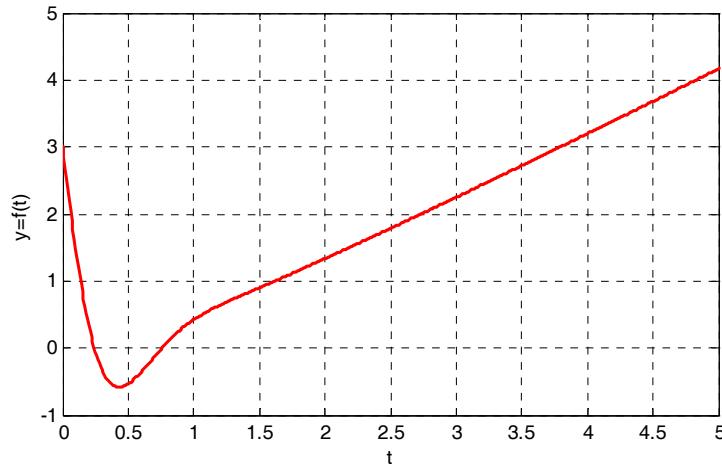


Figure B.8. Plot for Example B.13

Had we, in this example, defined the time interval starting with a negative value equal to or less than -1 , say as $-3 \leq t \leq 3$, MATLAB would have displayed the following message:

Warning: Divide by zero.

This is because the last term (the rational fraction) of the given expression, is divided by zero when $t = -1$. To avoid division by zero, we use the special MATLAB function **eps**, which is a number approximately equal to 2.2×10^{-16} . It will be used with the next example.

The command **axis([xmin xmax ymin ymax])** scales the current plot to the values specified by the arguments **xmin**, **xmax**, **ymin** and **ymax**. There are no commas between these four arguments. This command must be placed *after* the plot command and must be repeated for each plot. The following example illustrates the use of the dot multiplication, division, and exponentiation, the **eps** number, the **axis([xmin xmax ymin ymax])** command, and also MATLAB's capability of displaying up to four windows of different plots.

Example B.14

Plot the functions

$$y = \sin^2x, \quad z = \cos^2x, \quad w = \sin^2x \cdot \cos^2x, \quad v = \sin^2x / \cos^2x$$

in the interval $0 \leq x \leq 2\pi$ using 100 data points. Use the **subplot** command to display these functions on four windows on the same graph.

Solution:

The MATLAB script to produce the four subplots is as follows:

```
x=linspace(0,2*pi,100);           % Interval with 100 data points
y=(sin(x).^ 2); z=(cos(x).^ 2);
w=y.* z;
v=y./ (z+eps);% add eps to avoid division by zero
subplot(221);% upper left of four subplots
plot(x,y); axis([0 2*pi 0 1]);
title('y= (sinx) ^ 2');
subplot(222);                       % upper right of four subplots
plot(x,z); axis([0 2*pi 0 1]);
title('z= (cosx) ^ 2');
subplot(223);                       % lower left of four subplots
plot(x,w); axis([0 2*pi 0 0.3]);
title('w= (sinx) ^ 2*(cosx) ^ 2');
subplot(224);                       % lower right of four subplots
plot(x,v); axis([0 2*pi 0 400]);
title('v= (sinx) ^ 2/(cosx) ^ 2');
```

These subplots are shown in Figure B.9.

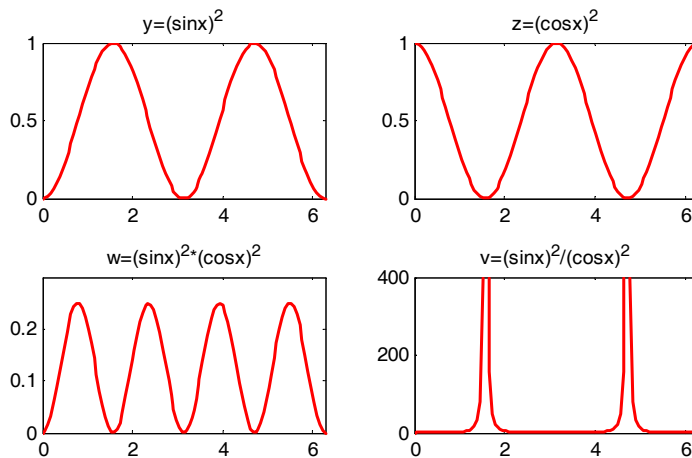


Figure B.9. Subplots for the functions of Example B.14

The next example illustrates MATLAB's capabilities with imaginary numbers. We will introduce the **real(z)** and **imag(z)** functions that display the real and imaginary parts of the complex quantity $z = x + iy$, the **abs(z)**, and the **angle(z)** functions that compute the absolute value (magnitude) and phase angle of the complex quantity $z = x + iy = r\angle\theta$. We will also use the

polar(theta,r) function that produces a plot in polar coordinates, where **r** is the magnitude, **theta** is the angle in radians, and the **round(n)** function that rounds a number to its nearest integer.

Example B.15

Consider the electric circuit of Figure B.10.

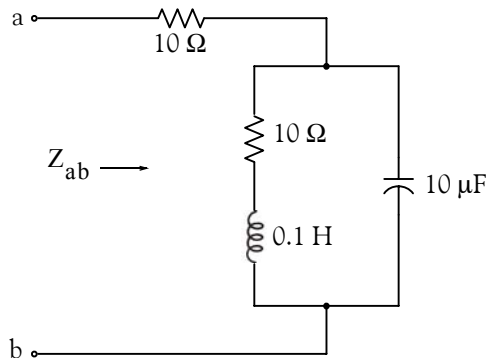


Figure B.10. Electric circuit for Example B.15

With the given values of resistance, inductance, and capacitance, the impedance Z_{ab} as a function of the radian frequency ω can be computed from the following expression:

$$Z_{ab} = Z = 10 + \frac{10^4 - j(10^6/\omega)}{10 + j(0.1\omega - 10^5/\omega)} \quad (\text{B.8})$$

- Plot $\text{Re}\{Z\}$ (the real part of the impedance Z) versus frequency ω .
- Plot $\text{Im}\{Z\}$ (the imaginary part of the impedance Z) versus frequency ω .
- Plot the impedance Z versus frequency ω in polar coordinates.

Solution:

The MATLAB script below computes the real and imaginary parts of Z_{ab} which, for simplicity, are denoted as z , and plots these as two separate graphs (parts a & b). It also produces a polar plot (part c).

```
w=0: 1: 2000;           % Define interval with one radian interval;...
z=(10+(10.^4 -j.* 10.^6 ./ (w+eps)) ./ (10 + j.* (0.1 .* w -10.^5./ (w+eps)))));...
%
% The first five statements (next two lines) compute and plot Re{z}
real_part=real(z); plot(w,real_part);...
xlabel('radian frequency w'); ylabel('Real part of Z'); grid
```

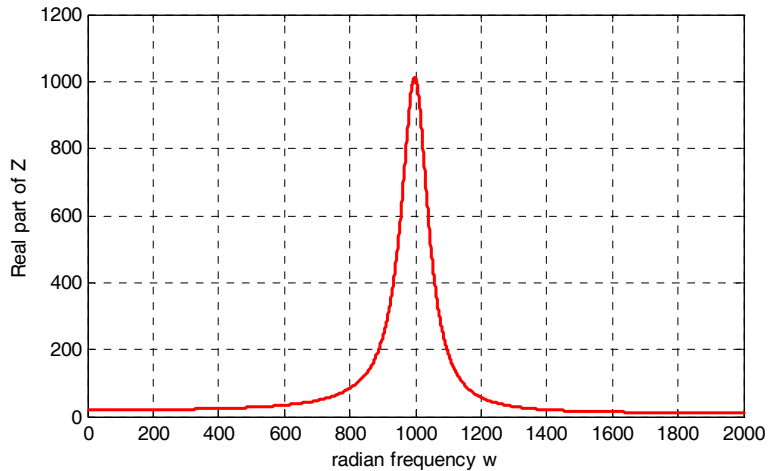


Figure B.11. Plot for the real part of the impedance in Example B.15

```
% The next five statements (next two lines) compute and plot Im{z}
imag_part=imag(z); plot(w,imag_part);...
xlabel('radian frequency w'); ylabel('Imaginary part of Z'); grid
```

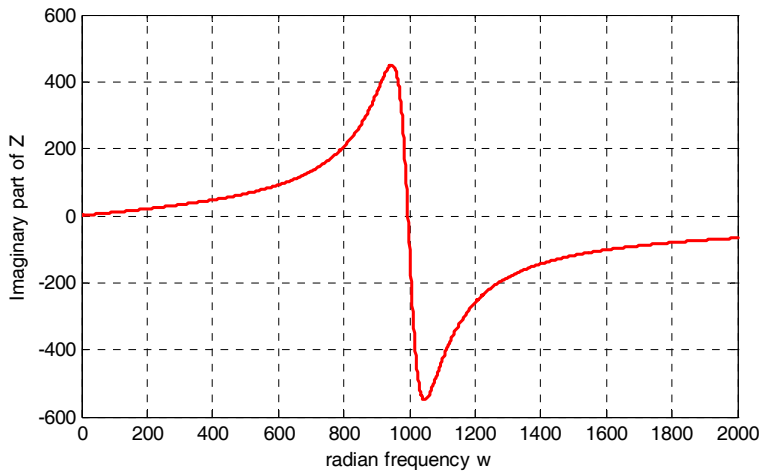


Figure B.12. Plot for the imaginary part of the impedance in Example B.15

```
% The last six statements (next five lines) below produce the polar plot of z
mag=abs(z);           % Computes |Z|;...
rndz=round(abs(z));   % Rounds |Z| to read polar plot easier;...
theta=angle(z);      % Computes the phase angle of impedance Z;...
polar(theta,rndz);    % Angle is the first argument
ylabel('Polar Plot of Z'); grid
```

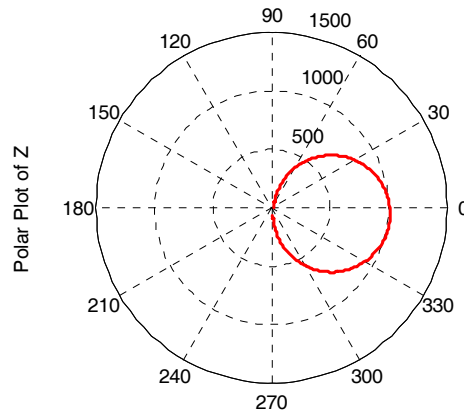


Figure B.13. Polar plot of the impedance in Example B.15

Example B.15 clearly illustrates how powerful, fast, accurate, and flexible MATLAB is.

B.10 Script and Function Files

MATLAB recognizes two types of files: *script files* and *function files*. Both types are referred to as *m-files* since both require the *.m* extension.

A *script file* consists of two or more built-in functions such as those we have discussed thus far. Thus, the script for each of the examples we discussed earlier, make up a script file. Generally, a script file is one which was generated and saved as an *m-file* with an editor such as the MATLAB's Editor/Debugger.

A *function file* is a user-defined function using MATLAB. We use function files for repetitive tasks. The first line of a function file must contain the word *function*, followed by the output argument, the equal sign (*=*), and the input argument enclosed in parentheses. The function name and file name must be the same, but the file name must have the extension *.m*. For example, the function file consisting of the two lines below

```
function y = myfunction(x)
y=x.^ 3 + cos(3.* x)
```

is a function file and must be saved as *myfunction.m*

For the next example, we will use the following MATLAB functions:

fzero(f,x) – attempts to find a zero of a function of one variable, where **f** is a string containing the name of a real-valued function of a single real variable. MATLAB searches for a value near a

point where the function **f** changes sign, and returns that value, or returns NaN if the search fails. **Important:** We must remember that we use **roots(p)** to find the roots of polynomials only, such as those in Examples B.1 and B.2.

fplot(fcn,lims) – plots the function specified by the string **fcn** between the x-axis limits specified by **lims = [xmin xmax]**. Using **lims = [xmin xmax ymin ymax]** also controls the y-axis limits. The string **fcn** must be the name of an *m-file* function or a string with variable *x*.

NaN (Not-a-Number) is not a function; it is MATLAB's response to an undefined expression such as $0/0$, ∞/∞ , or inability to produce a result as described on the next paragraph. We can avoid division by zero using the **eps** number, which we mentioned earlier.

Example B.16

Find the zeros, the minimum, and the maximum values of the function

$$f(x) = \frac{1}{(x-0.1)^2 + 0.01} - \frac{1}{(x-1.2)^2 + 0.04} - 10 \quad (\text{B.9})$$

in the interval $-1.5 \leq x \leq 1.5$

Solution:

We first plot this function to observe the approximate zeros, maxima, and minima using the following script.

```
x=-1.5:0.01:1.5;
y=1./((x-0.1).^2+0.01)-1./((x-1.2).^2+0.04)-10;
plot(x,y); grid
```

The plot is shown in Figure B.14.

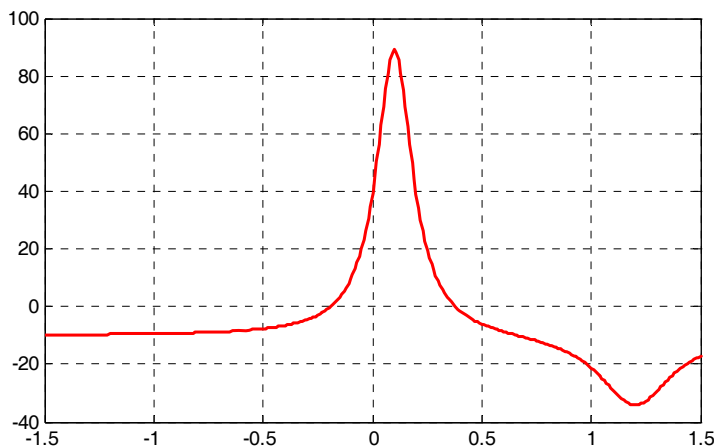


Figure B.14. Plot for Example B.16 using the `plot` command

The roots (zeros) of this function appear to be in the neighborhood of $x = -0.2$ and $x = 0.3$. The maximum occurs at approximately $x = 0.1$ where, approximately, $y_{\max} = 90$, and the minimum occurs at approximately $x = 1.2$ where, approximately, $y_{\min} = -34$.

Next, we define and save $f(x)$ as the **funczero01.m** function m-file with the following script:

```
function y=funczero01(x)
% Finding the zeros of the function shown below
y=1/((x-0.1)^2+0.01)-1/((x-1.2)^2+0.04)-10;
```

To save this file, from the File drop menu on the Command Window, we choose New, and when the Editor Window appears, we type the script above and we save it as **funczero01**. MATLAB appends the extension **.m** to it.

Now, we can use the **fplot(fcn,lims)** command to plot $f(x)$ as follows:

```
fplot('funczero01', [-1.5 1.5]); grid
```

This plot is shown in Figure B.15. As expected, this plot is identical to the plot of Figure B.14 which was obtained with the **plot(x,y)** command as shown in Figure B.14.

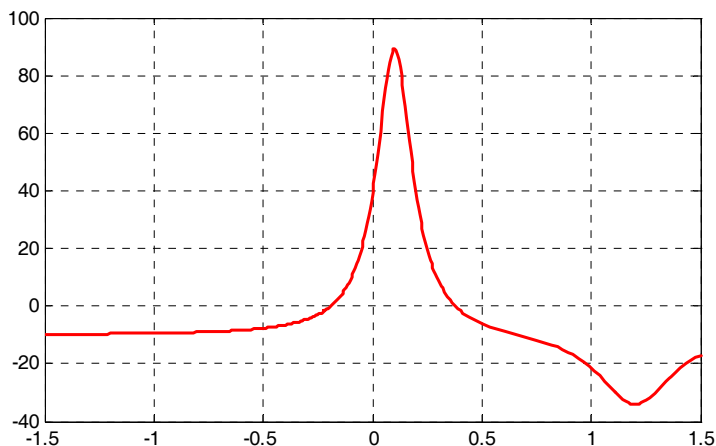


Figure B.15. Plot for Example B.16 using the **fplot** command

We will use the **fzero(f,x)** function to compute the roots of $f(x)$ in Equation (B.9) more precisely. The MATLAB script below will accomplish this.

```
x1= fzero('funczero01', -0.2);
x2= fzero('funczero01', 0.3);
fprintf('The roots (zeros) of this function are r1= %3.4f', x1);
fprintf(' and r2= %3.4f\n', x2)
```

MATLAB displays the following:

The roots (zeros) of this function are $r_1 = -0.1919$ and $r_2 = 0.3788$

The earlier MATLAB versions included the function **fmin(f,x1,x2)** and with this function we could compute both a minimum of some function $f(x)$ or a maximum of $f(x)$ since a maximum of $f(x)$ is equal to a minimum of $-f(x)$. This can be visualized by flipping the plot of a function $f(x)$ upside-down. This function is no longer used in MATLAB and thus we will compute the maxima and minima from the derivative of the given function.

From elementary calculus, we recall that the maxima or minima of a function $y = f(x)$ can be found by setting the first derivative of a function equal to zero and solving for the independent variable x . For this example we use the **diff(x)** function which produces the approximate derivative of a function. Thus, we use the following MATLAB script:

```
syms x ymin zmin; ymin=1/((x-0.1)^2+0.01)-1/((x-1.2)^2+0.04)-10;...
zmin=diff(ymin)
```

```
zmin =
-1 / ( (x-1/10) ^2+1/100) ^2 * (2*x-1/5) +1 / ( (x-6/5) ^2+1/25) ^2 * (2*x-12/5)
```

When the command

```
solve(zmin)
```

is executed, MATLAB displays a very long expression which when copied at the command prompt and executed, produces the following:

```
ans =
    0.6585 + 0.3437i
ans =
    0.6585 - 0.3437i
ans =
    1.2012
```

The real value 1.2012 above is the value of x at which the function y has its minimum value as we observe also in the plot of Figure B.15.

To find the value of y corresponding to this value of x , we substitute it into $f(x)$, that is,

```
x=1.2012; ymin=1 / ((x-0.1) ^ 2 + 0.01) -1 / ((x-1.2) ^ 2 + 0.04) -10
ymin = -34.1812
```

We can find the maximum value from $-f(x)$ whose plot is produced with the script

```
x=-1.5:0.01:1.5; ymax=-1./((x-0.1).^2+0.01)+1./((x-1.2).^2+0.04)+10; plot(x,ymax); grid
```

and the plot is shown in Figure B.16.

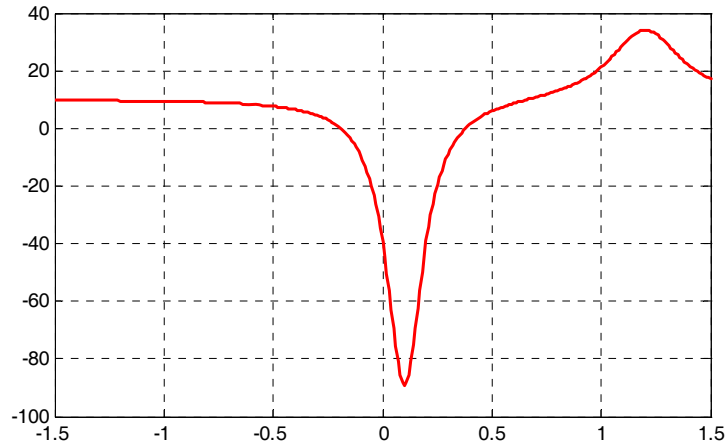


Figure B.16. Plot of $-f(x)$ for Example B.16

Next we compute the first derivative of $-f(x)$ and we solve for x to find the value where the maximum of y_{max} occurs. This is accomplished with the MATLAB script below.

```
syms x ymax zmax; ymax=-(1/((x-0.1)^2+0.01))-1/((x-1.2)^2+0.04)-10; zmax=diff(ymax)
zmax =
 1 / ((x-1/10)^2+1/100)^2*(2*x-1/5) - 1 / ((x-6/5)^2+1/25)^2*(2*x-12/5)
solve(zmax)
```

When the command

```
solve(zmax)
```

is executed, MATLAB displays a very long expression which when copied at the command prompt and executed, produces the following:

```
ans =
 0.6585 + 0.3437i
ans =
 0.6585 - 0.3437i
ans =
 1.2012
ans =
 0.0999
```

From the values above we choose $x = 0.0999$ which is consistent with the plots of Figures B.15 and B.16. Accordingly, we execute the following script to obtain the value of y_{min} .

```
x=0.0999;           % Using this value find the corresponding value of ymax
ymax=1 / ((x-0.1) ^ 2 + 0.01) - 1 / ((x-1.2) ^ 2 + 0.04) - 10

ymax = 89.2000
```

B.11 Display Formats

MATLAB displays the results on the screen in integer format without decimals if the result is an integer number, or in short floating point format with four decimals if it a fractional number. The format displayed has nothing to do with the accuracy in the computations. MATLAB performs all computations with accuracy up to 16 decimal places.

The output format can be changed with the **format** command. The available MATLAB formats can be displayed with the **help format** command as follows:

help format

```
FORMAT Set output format.
All computations in MATLAB are done in double precision.
FORMAT may be used to switch between different output display formats
as follows:
```

```
FORMAT Default. Same as SHORT.
FORMAT SHORT Scaled fixed point format with 5 digits.
FORMAT LONG Scaled fixed point format with 15 digits.
FORMAT SHORT E Floating point format with 5 digits.
FORMAT LONG E Floating point format with 15 digits.
FORMAT SHORT G Best of fixed or floating point format with 5 digits.
FORMAT LONG G Best of fixed or floating point format with 15 digits.
FORMAT HEX Hexadecimal format.
FORMAT + The symbols +, - and blank are printed for positive, negative,
and zero elements. Imaginary parts are ignored.
FORMAT BANK Fixed format for dollars and cents.
FORMAT RAT Approximation by ratio of small integers.
```

Spacing:

```
FORMAT COMPACT Suppress extra line-feeds.
FORMAT LOOSE Puts the extra line-feeds back in.
```

Some examples with different format displays are given below.

```
format short 33.3335 Four decimal digits (default)
format long 33.33333333333334 16 digits
format short e 3.3333e+01 Four decimal digits plus exponent
format short g 33.333 Better of format short or format short e
format bank 33.33 two decimal digits
format + only + or - or zero are printed
```

```
format rat 100/3 rational approximation
```

The **disp(X)** command displays the array **X** without printing the array name. If **X** is a string, the text is displayed.

The **fprintf(format,array)** command displays and prints both text and arrays. It uses specifiers to indicate where and in which format the values would be displayed and printed. Thus, if **%f** is used, the values will be displayed and printed in fixed decimal format, and if **%e** is used, the values will be displayed and printed in scientific notation format. With this command only the real part of each parameter is processed.

This appendix is just an introduction to MATLAB.* This outstanding software package consists of many applications known as *Toolboxes*. The MATLAB Student Version contains just a few of these Toolboxes. Others can be bought directly from The MathWorks,™ Inc., as add-ons.

* For more MATLAB applications, please refer to *Numerical Analysis Using MATLAB and Spreadsheets*, ISBN 0-9709511-1-6.

Appendix C

The Gamma and Beta Functions and Distributions

This appendix serves as an introduction to the gamma and beta functions. These are special functions that find wide applications in science and engineering. They are also used in probability and in the computation of certain integrals.

C.1 The Gamma Function

The *gamma function*, denoted as $\Gamma(n)$, is also known as *generalized factorial function*. It is defined as

$$\Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx \quad (C.1)$$

and this improper* integral converges (approaches a limit) for all $n > 0$.

We will derive the basic properties of the gamma function and its relation to the well known factorial function

$$n! = n(n-1)(n-2)\dots 3 \cdot 2 \cdot 1 \quad (C.2)$$

We will evaluate the integral of (C.1) by performing integration by parts using the relation

$$\int u dv = uv - \int v du \quad (C.3)$$

Letting

$$u = e^{-x} \quad \text{and} \quad dv = x^{n-1} \quad (C.4)$$

we obtain

$$du = -e^{-x} dx \quad \text{and} \quad v = \frac{x^n}{n} \quad (C.5)$$

Then, with (C.3), we write (C.1) as

* Improper integrals are two types and these are:

a. $\int_a^b f(x) dx$ where the limits of integration a or b or both are infinite

b. $\int_a^b f(x) dx$ where $f(x)$ becomes infinite at a value x between the lower and upper limits of integration inclusive.

Appendix C The Gamma and Beta Functions and Distributions

$$\Gamma(n) = \frac{x^n e^{-x}}{n} \Big|_{x=0}^{\infty} + \frac{1}{n} \int_0^{\infty} x^n e^{-x} dx \quad (C.6)$$

With the condition that $n > 0$, the first term on the right side of (C.6) vanishes at the lower limit, that is, for $x = 0$. It also vanishes at the upper limit as $x \rightarrow \infty$. This can be proved with *L' Hôpital's rule** by differentiating both numerator and denominator m times, where $m \geq n$. Then,

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{x^n e^{-x}}{n} &= \lim_{x \rightarrow \infty} \frac{x^n}{n e^x} = \lim_{x \rightarrow \infty} \frac{\frac{d^m}{dx^m} x^n}{\frac{d^m}{dx^m} n e^x} = \lim_{x \rightarrow \infty} \frac{\frac{d^{m-1}}{dx^{m-1}} n x^{n-1}}{\frac{d^{m-1}}{dx^{m-1}} n e^x} = \dots \\ &= \lim_{x \rightarrow \infty} \frac{n(n-1)(n-2)\dots(n-m+1)x^{n-m}}{n e^x} \\ &= \lim_{x \rightarrow \infty} \frac{(n-1)(n-2)\dots(n-m+1)}{x^{m-n} e^x} = 0 \end{aligned} \quad (C.7)$$

Therefore, (C.6) reduces to

$$\Gamma(n) = \frac{1}{n} \int_0^{\infty} x^n e^{-x} dx \quad (C.8)$$

and with (C.1) we have

$$\Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx = \frac{1}{n} \int_0^{\infty} x^n e^{-x} dx \quad (C.9)$$

By comparing the two integrals of (C.9), we observe that

$$\boxed{\Gamma(n) = \frac{\Gamma(n+1)}{n}} \quad (C.10)$$

or

$$\boxed{n\Gamma(n) = \Gamma(n+1)} \quad (C.11)$$

* Often, the ratio of two functions, such as $\frac{f(x)}{g(x)}$, for some value of x , say a , results in the indeterminate form $\frac{f(a)}{g(a)} = \frac{0}{0}$. To work around this problem, we consider the limit $\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$, and we wish to find this limit, if it exists. *L'Hôpital's rule* states that if $f(a) = g(a) = 0$, and if the limit $\frac{d}{dx} f(x) / \frac{d}{dx} g(x)$ as x approaches a exists, then,

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \left(\frac{\frac{d}{dx} f(x)}{\frac{d}{dx} g(x)} \right)$$

It is convenient to use (C.10) for $n < 0$, and (C.11) for $n > 0$.

From (C.10), we see that $\Gamma(n)$ becomes infinite as $n \rightarrow 0$.

For $n = 1$, (C.1) yields

$$\Gamma(1) = \int_0^{\infty} e^{-x} dx = -e^{-x} \Big|_0^{\infty} = 1 \quad (\text{C.12})$$

Thus, we have derived the important relation,

$$\Gamma(1) = 1 \quad (\text{C.13})$$

From the recurring relation of (C.11), we obtain

$$\begin{aligned} \Gamma(2) &= 1 \cdot \Gamma(1) = 1 \\ \Gamma(3) &= 2 \cdot \Gamma(2) = 2 \cdot 1 = 2! \\ \Gamma(4) &= 3 \cdot \Gamma(3) = 3 \cdot 2 = 3! \end{aligned} \quad (\text{C.14})$$

and in general

$\Gamma(n + 1) = n! \quad \text{for } n = 1, 2, 3, \dots$

(C.15)

The formula of (C.15) is a very useful relation; it establishes the relationship between the $\Gamma(n)$ function and the factorial $n!$.

We must remember that, whereas the factorial $n!$ is defined only for zero (recall that $0! = 1$) and positive integer values, the gamma function exists (is continuous) everywhere *except* at 0 and *negative integer numbers*, that is, $-1, -2, -3$, and so on. For instance, when $n = -0.5$, we can find $\Gamma(-0.5)$ in terms of $\Gamma(0.5)$, but if we substitute the numbers 0, $-1, -2, -3$ and so on in (C.11), we obtain values which are not consistent with the definition of the $\Gamma(n)$ function, as defined in that relation.

Stated in other words, *the $\Gamma(n)$ function is defined for all positive integers and positive fractional values, and for all negative fractional, but not negative integer values.*

We can use the MATLAB **gamma(n)** function to plot $\Gamma(n)$ versus n . This is done with the code below that produces the plot shown in Figure C.1.

```
n=-4: 0.05: 4; g=gamma(n); plot(n,g); axis([-4 4 -6 6]); grid;
title('The Gamma Function'); xlabel('n'); ylabel('Gamma(n)')
```

Figure C.1 shows the plot of the function $\Gamma(n)$ versus n .

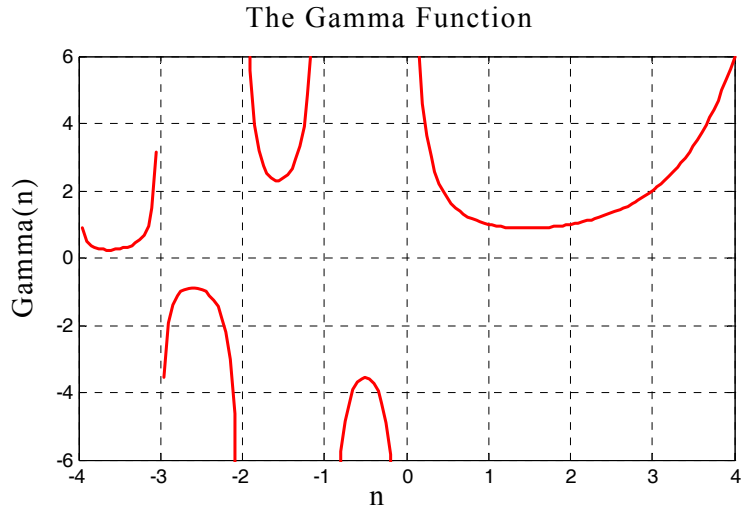


Figure C.1. Plot of the gamma function

Numerical values of $\Gamma(n)$ for $1 \leq n \leq 2$, can be found in math tables, but we can use (C.10) or (C.11) to compute values outside this range. Of course, we can use MATLAB to find any valid values of n .

Example C.1

Compute:

- a. $\Gamma(3.6)$ b. $\Gamma(0.5)$ c. $\Gamma(-0.5)$

Solution:

a. From (C.11),

$$\Gamma(n + 1) = n\Gamma(n)$$

Then,

$$\Gamma(3.6) = 2.6\Gamma(2.6) = (2.6)(1.6)\Gamma(1.6)$$

and from math tables,

$$\Gamma(1.6) = 0.8953$$

Therefore,

$$\Gamma(3.6) = (2.6)(1.6)(0.8953) = 3.717$$

b. From (C.10),

$$\Gamma(n) = \frac{\Gamma(n + 1)}{n}$$

Then,

$$\Gamma(0.5) = \frac{\Gamma(0.5 + 1)}{0.5} = \frac{\Gamma(1.5)}{0.5}$$

and from math tables,

$$\Gamma(1.5) = 0.8862$$

Therefore,

$$\Gamma(0.5) = (2)(0.8862) = 1.772$$

c. From (C.10),

$$\Gamma(n) = \frac{\Gamma(n+1)}{n}$$

Then,

$$\Gamma(-0.5) = \frac{\Gamma(-0.5+1)}{-0.5} = \frac{\Gamma(0.5)}{-0.5} = -2\Gamma(0.5)$$

and using the result of (b),

$$\Gamma(-0.5) = -2\Gamma(0.5) = (-2)(1.772) = -3.544$$

We can verify these answers with MATLAB as follows:

a=gamma(3.6), b=gamma(0.5), c=gamma(-0.5)

a =
3.7170
b =
1.7725
c =
-3.5449

Excel does not have a function which evaluates $\Gamma(n)$ directly. It does, however, have the **GAMMALN(x)** function. Therefore, we can use the **=EXP(GAMMALN(n))** function to evaluate $\Gamma(n)$ at some positive value of n . But because it first computes the natural log, it does not produce an answer if n is negative as shown in Figure C.2.

	A	B	C
1	Example B.1 using Excel		
2			exp(gammaln(n))=
3	n	gammaln(n)	gamma(n)
4			
5	3.6	1.3129	3.7170
6			
7	0.5	0.5724	1.7725
8			
9	-0.5	#NUM!	#NUM!

Figure C.2. Using Excel to find $\Gamma(n)$.

Appendix C The Gamma and Beta Functions and Distributions

Example C.2

Prove that when n is a positive integer, the relation

$$\boxed{\Gamma(n) = (n-1)!} \quad (\text{C.16})$$

is true.

Proof:

From (C.11),

$$\Gamma(n+1) = n\Gamma(n) \quad (\text{C.17})$$

Then,

$$\Gamma(n) = (n-1)\Gamma(n-1) \quad (\text{C.18})$$

Next, replacing n with $n-1$ on the left side of (C.18), we obtain

$$\Gamma(n-1) = (n-2)\Gamma(n-2) \quad (\text{C.19})$$

Substitution of (C.19) into (C.18) yields

$$\Gamma(n) = (n-1)(n-2)\Gamma(n-2) \quad (\text{C.20})$$

By n repeated substitutions, we obtain

$$\Gamma(n) = (n-1)(n-2)(n-3)\dots 1\Gamma(1) \quad (\text{C.21})$$

and since $\Gamma(1) = 1$, we have

$$\Gamma(n) = (n-1)(n-2)(n-3)\dots 1 \quad (\text{C.22})$$

or

$$\Gamma(n) = (n-1)! \quad (\text{C.23})$$

Example C.3

Use the definition of the $\Gamma(n)$ function to compute the exact value of $\Gamma(1/2)$

Solution:

From (C.1),

$$\Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx \quad (\text{C.24})$$

Then,

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} x^{0.5-1} e^{-x} dx = \int_0^{\infty} x^{-0.5} e^{-x} dx \quad (\text{C.25})$$

Letting

$$x = y^2$$

we obtain

$$\frac{dx}{dy} = 2y$$

or

$$dx = 2ydy$$

By substitution of the last three relations into (C.25), we obtain

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} y^{2(-0.5)} e^{-y^2} 2ydy = 2 \int_0^{\infty} y^{-1} ye^{-y^2} dy = 2 \int_0^{\infty} e^{-y^2} dy \quad (C.26)$$

Next, we define $\Gamma(1/2)$ as a function of both x and y , that is, we let

$$\Gamma\left(\frac{1}{2}\right) = 2 \int_0^{\infty} e^{-x^2} dx \quad (C.27)$$

$$\Gamma\left(\frac{1}{2}\right) = 2 \int_0^{\infty} e^{-y^2} dy \quad (C.28)$$

Multiplication of (C.27) by (C.28) yields

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^2 = 4 \int_0^{\infty} e^{-x^2} dx \int_0^{\infty} e^{-y^2} dy = 4 \int_0^{\infty} \int_0^{\infty} e^{-(x^2+y^2)} dx dy \quad (C.29)$$

Now, we convert (C.29) to polar coordinates by making the substitution

$$\rho^2 = x^2 + y^2 \quad (C.30)$$

and by recalling that:

1. the total area of a region is found by either one of the double integrals

$$A = \iint dx dy = \iint r dr d\theta \quad (C.31)$$

2. from differential calculus

$$\frac{d}{du} e^{u^2} = e^{u^2} \frac{d}{du} u^2 = 2ue^{u^2} \quad (C.32)$$

Then,

$$\int_{\rho_1}^{\rho_2} \rho e^{-\rho^2} d\rho = -\frac{1}{2} e^{-\rho^2} \quad (C.33)$$

We observe that as $x \rightarrow \infty$ and $y \rightarrow \infty$,

Appendix C The Gamma and Beta Functions and Distributions

$$\rho \rightarrow \infty \quad \text{and} \quad \theta \rightarrow \frac{\pi}{2} \quad (\text{C.34})$$

Substitution of (C.30), (C.33) and (C.34) into (C.29) yields

$$\left[\Gamma\left(\frac{1}{2}\right) \right]^2 = -2 \int_0^{\pi/2} \left(e^{-\rho^2} \Big|_{\rho=0}^{\infty} \right) d\theta = -2 \int_0^{\pi/2} (0 - 1) d\theta = 2 \int_0^{\pi/2} d\theta = 2\theta \Big|_0^{\pi/2} = \pi$$

and thus, we have obtained the exact value

$$\boxed{\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}} \quad (\text{C.35})$$

Example C.4

Compute:

a. $\Gamma(-0.5)$ b. $\Gamma(-1.5)$ c. $\Gamma(-2.5)$

Solution:

Using the relations

$$\Gamma(n) = \frac{\Gamma(n+1)}{n} \quad \text{and} \quad \Gamma(0.5) = \sqrt{\pi}$$

we obtain:

a. for $n = -0.5$,

$$\Gamma(-0.5) = \frac{\Gamma(0.5)}{-0.5} = \frac{\sqrt{\pi}}{-0.5} = -2\sqrt{\pi}$$

b. for $n = -1.5$,

$$\Gamma(-1.5) = \frac{\Gamma(-1.5+1)}{-1.5} = \frac{\Gamma(-0.5)}{-1.5} = \frac{-2\sqrt{\pi}}{-1.5} = \frac{4}{3}\sqrt{\pi}$$

c. for $n = -2.5$,

$$\Gamma(-2.5) = \frac{\Gamma(-2.5+1)}{-2.5} = \frac{\Gamma(-1.5)}{-2.5} = \frac{(4/3)\sqrt{\pi}}{-2.5} = -\frac{8}{15}\sqrt{\pi}$$

Other interesting relations involving the $\Gamma(n)$ function are:

$$\boxed{\Gamma(n)\Gamma(1-n) = \frac{\pi}{\sin n\pi}} \quad (\text{C.36})$$

for $0 < n < 1$

$$2^{2n-1}\Gamma(n)\sqrt{\pi} = \sqrt{\pi}\Gamma(2n)$$

for any $n \neq$ negative integer

(C.37)

$$\Gamma(n+1) = n!$$

$$= \sqrt{2\pi n} n^n e^{-n} \left\{ 1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3} - \frac{571}{2488320n^4} + \dots \right\}$$
(C.38)

Relation (C.38) is referred to as *Stirling's asymptotic series for the $\Gamma(n)$ function*. If n is a positive integer, the factorial $n!$ can be approximated as

$$n! \approx \sqrt{2\pi n} n^n e^{-n}$$
(C.39)

Example C.5

Use (C.36) to prove that

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

Proof:

$$\Gamma\left(\frac{1}{2}\right)\Gamma\left(1 - \frac{1}{2}\right) = \Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right) = \frac{\pi}{\sin\frac{\pi}{2}}$$

or

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^2 = \pi$$

Therefore,

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

Example C.6

Compute the product

$$\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right)$$

Solution:

Using (C.36), we obtain

Appendix C The Gamma and Beta Functions and Distributions

$$\Gamma\left(\frac{1}{3}\right)\Gamma\left(1-\frac{1}{3}\right) = \frac{\pi}{\sin(\pi/3)}$$

or

$$\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right) = \frac{\pi}{\sqrt{3}/2} = \frac{2\pi}{\sqrt{3}} = \frac{2\sqrt{3}\pi}{3}$$

Example C.7

Use (C.37) to find the value of

$$\Gamma\left(\frac{3}{2}\right)$$

Solution:

$$2^{3-1}\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{3}{2} + \frac{1}{2}\right) = \sqrt{\pi}\Gamma\left(2 \cdot \frac{3}{2}\right)$$

or

$$2^2\Gamma\left(\frac{3}{2}\right)\Gamma(2) = \sqrt{\pi}\Gamma(3)$$

or

$$\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}\Gamma(3)}{4\Gamma(2)} = \frac{2!\sqrt{\pi}}{4 \cdot 1} = \frac{\sqrt{\pi}}{2}$$

Example C.8

Use (C.39) to compute 50!

Solution:

$$50! \approx \sqrt{2\pi \times 50} \times 50^{50} \times e^{-50}$$

We can use MATLAB or Excel as a calculator to evaluate this expression. With MATLAB we type and execute the expression

```
sqrt(2*pi*50)*50^50*exp(-50)
```

```
ans =  
3.0363e+064
```

This is an approximation. To find the exact value, we use the relation $\Gamma(n+1) = n!$ and the MATLAB **gamma(n)** function. Then,

gamma(50+1)

ans =
3.0414e+064

We can check this answer with the Excel FACT(n) function, that is, =FACT(50) and Excel displays 3.04141E+64

The $\Gamma(n)$ function is very useful in integrating some improper integrals. Some examples follow.

Example C.9

Using the definition of the $\Gamma(n)$ function, evaluate the integrals

$$\text{a. } \int_0^{\infty} x^4 e^{-x} dx \quad \text{b. } \int_0^{\infty} x^5 e^{-2x} dx$$

Solution:

By definition,

$$\int_0^{\infty} x^{n-1} e^{-x} dx = \Gamma(n)$$

Then,

a.

$$\int_0^{\infty} x^4 e^{-x} dx = \Gamma(5) = 4! = 24$$

b.

Let $2x = y$; then, $dx = dy/2$, and by substitution,

$$\int_0^{\infty} x^5 e^{-2x} dx = \int_0^{\infty} \left(\frac{y}{2}\right)^5 e^{-y} \frac{dy}{2} = \frac{1}{2^6} \int_0^{\infty} y^5 e^{-y} dy = \frac{\Gamma(6)}{64} = \frac{5!}{64} = \frac{120}{64} = \frac{15}{8}$$

Example C.10

A negatively charged particle is α meters apart from the positively charged side of an electric field. It is initially at rest, and then moves towards the positively charged side with a force inversely proportional to its distance from it. Assuming that the particle moves towards the center of the positively charged side, considered to be the center of attraction 0, derive an expression for the time required the negatively charged particle to reach 0 in terms of the distance α and its mass m .

Appendix C The Gamma and Beta Functions and Distributions

Solution:

Let the center of attraction 0 be the point zero on the x -axis, as indicated in Figure C.3.

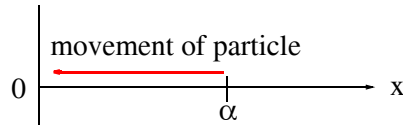


Figure C.3. Sketch for Example C.10

By Newton's law,

$$m \frac{dx^2}{dt^2} = -\frac{k}{x} \quad (\text{C.40})$$

where

m = mass of particle

x = distance (varies with time)

k = positive constant of proportionality and the

–sign indicates that the distance x decreases as time t increases.

At $t = 0$, the particle is assumed to be located on the x -axis at point $x = \alpha$, and moves towards the origin at $x = 0$. Let the velocity of the particle be v . Then,

$$\frac{dx}{dt} = v \quad (\text{C.41})$$

and

$$\frac{dx^2}{dt^2} = \frac{dv}{dt} = \frac{dv dx}{dx dt} = v \frac{dv}{dx} \quad (\text{C.42})$$

Substitution of (C.42) into (C.40) yields

$$mv \frac{dv}{dx} = -\frac{k}{x} \quad (\text{C.43})$$

or

$$mvdv = -\frac{k}{x}(dx) \quad (\text{C.44})$$

Integrating both sides of (C.44), we obtain

$$\frac{mv^2}{2} = -k \ln x + C \quad (\text{C.45})$$

where C represents the constants of integration of both sides, and it is evaluated from the initial condition that $v = 0$ when $x = \alpha$. Then,

$$C = k \ln \alpha \quad (\text{C.46})$$

and by substitution into (C.45),

$$\frac{mv^2}{2} = k \ln \alpha - k \ln x = k \ln \frac{\alpha}{x} \quad (\text{C.47})$$

Solving for v^2 and taking the square root of both sides we obtain

$$v = \frac{dx}{dt} = \pm \sqrt{\frac{2k}{m} \ln \frac{\alpha}{x}} \quad (\text{C.48})$$

Since x decreases as t increases, we choose the negative sign, that is,

$$\frac{dx}{dt} = -\sqrt{\frac{2k}{m} \ln \frac{\alpha}{x}} \quad (\text{C.49})$$

Solving (C.49) for dt we obtain

$$dt = -\sqrt{\frac{m}{2k}} \frac{dx}{\sqrt{\ln(\alpha/x)}} \quad (\text{C.50})$$

We are interested in the time required for the particle to reach the origin 0. We denote this time as T ; it is found from the relation (C.51) below, noting that the integration on the right side is with respect to the distance x where at $t = 0$, $x = \alpha$, and at $\tau = t$, $x = 0$. Then,

$$T = \int_0^t d\tau = -\sqrt{\frac{m}{2k}} \int_{\alpha}^0 \frac{dx}{\sqrt{\ln(\alpha/x)}} \quad (\text{C.51})$$

To simplify (C.51), we let

$$y = \ln\left(\frac{\alpha}{x}\right), \text{ then } e^y = \frac{\alpha}{x} \quad (\text{C.52})$$

or

$$x = \alpha e^{-y}, \text{ and } dx = -\alpha e^{-y} dy \quad (\text{C.53})$$

Also, since

$$\lim_{x \rightarrow \alpha} \ln\left(\frac{\alpha}{x}\right) = 0 \text{ and } \lim_{x \rightarrow 0} \ln\left(\frac{\alpha}{x}\right) = \infty$$

the lower and upper limits of integration in (C.51), are being replaced with 0 and ∞ respectively. Therefore, we express (C.51) as

$$T = -\sqrt{\frac{m}{2k}} \int_0^{\infty} \frac{-\alpha e^{-y} dy}{\sqrt{y}} = \alpha \sqrt{\frac{m}{2k}} \int_0^{\infty} y^{-1/2} e^{-y} dy$$

Appendix C The Gamma and Beta Functions and Distributions

Finally, using the definition of the $\Gamma(n)$ function, we obtain

$$\Gamma = \alpha \Gamma\left(\frac{1}{2}\right) \sqrt{\frac{m}{2k}} = \alpha \sqrt{\pi} \sqrt{\frac{m}{2k}} = \alpha \sqrt{\frac{\pi m}{2k}} \quad (\text{C.54})$$

Example C.11

Evaluate the integrals

$$\int_0^{\pi/2} \cos^n \theta d\theta \quad \text{and} \quad \int_0^{\pi/2} \sin^n \theta d\theta \quad (\text{C.55})$$

Solution:

From the definition of the $\Gamma(n)$ function,

$$\Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx \quad (\text{C.56})$$

Also,

$$\Gamma(m) = \int_0^{\infty} x^{m-1} e^{-x} dx \quad (\text{C.57})$$

For $m > 0$ and $n > 0$, multiplication of (C.56) by (C.57) yields

$$\Gamma(m)\Gamma(n) = \int_0^{\infty} u^{m-1} e^{-u} du \int_0^{\infty} v^{n-1} e^{-v} dv \quad (\text{C.58})$$

where u and v are dummy variables of integration. Next, letting $u = x^2$ and $v = y^2$, we obtain $du = 2x dx$ and $dv = 2y dy$. Then, with these substitutions, relation (C.58) is written as

$$\begin{aligned} \Gamma(m)\Gamma(n) &= \int_0^{\infty} x^{2(m-1)} 2x e^{-x^2} dx \int_0^{\infty} y^{2(n-1)} 2y e^{-y^2} dy = 4 \int_0^{\infty} x^{2m-2} x e^{-x^2} dx \int_0^{\infty} y^{2n-2} y e^{-y^2} dy \\ &= 4 \int_0^{\infty} \int_0^{\infty} x^{2m-1} y^{2n-1} e^{-(x^2+y^2)} dx dy \end{aligned} \quad (\text{C.59})$$

Next, we convert (C.59) to polar coordinates by letting $x = \rho \cos \theta$ and $y = \rho \sin \theta$. Then,

$$\begin{aligned} \Gamma(m)\Gamma(n) &= 4 \int_0^{\pi/2} \int_0^{\infty} (\rho \cos \theta)^{2m-1} (\rho \sin \theta)^{2n-1} e^{-\rho^2} \rho d\rho d\theta \\ &= 2 \int_0^{\pi/2} \cos^{2m-1} \theta \cdot \sin^{2n-1} \theta d\theta \int_0^{\infty} \rho^{2m+2n-2} e^{-\rho^2} 2\rho d\rho \end{aligned} \quad (\text{C.60})$$

To simplify (C.60), we let $\rho^2 = w$; then, $dw = 2\rho d\rho$ and thus relation (C.60) is written as

$$\begin{aligned}\Gamma(m)\Gamma(n) &= 2 \int_0^{\pi/2} \cos^{2m-1}\theta \cdot \sin^{2n-1}\theta d\theta \int_0^\infty w^{m+n-1} e^{-w} dw \\ &= 2 \int_0^{\pi/2} \cos^{2m-1}\theta \cdot \sin^{2n-1}\theta d\theta \cdot \Gamma(m+n)\end{aligned}\tag{C.61}$$

Rearranging (C.61) we obtain

$$\int_0^{\pi/2} \cos^{2m-1}\theta \cdot \sin^{2n-1}\theta d\theta = \frac{\Gamma(m)\Gamma(n)}{2\Gamma(m+n)}\tag{C.62}$$

and this expression can be simplified by replacing $2m - 1$ with n , that is, $m = (n + 1)/2$, and $2n - 1$ with 0 , that is, $n = 1/2$. Then, we obtain the special case of (C.62) as

$$\int_0^{\pi/2} \cos^n\theta d\theta = \frac{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{2\Gamma\left(\frac{n+1}{2} + \frac{1}{2}\right)} = \frac{\Gamma\left(\frac{n+1}{2}\right)\sqrt{\pi}}{\Gamma\left(\frac{n}{2} + 1\right)2}\tag{C.63}$$

If, in (C.62), we replace $2m - 1$ with 0 and $2n - 1$ with m , we obtain the integral of the $\sin^n\theta$ function as

$$\int_0^{\pi/2} \sin^m\theta d\theta = \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{m+1}{2}\right)}{2\Gamma\left(\frac{1}{2} + \frac{m+1}{2}\right)} = \frac{\Gamma\left(\frac{m+1}{2}\right)\sqrt{\pi}}{\Gamma\left(\frac{m}{2} + 1\right)2}\tag{C.64}$$

We observe that (C.63) and (C.64) are equal since m and n can be interchanged. Therefore,

$$\boxed{\int_0^{\pi/2} \cos^n\theta d\theta = \int_0^{\pi/2} \sin^n\theta d\theta = \frac{\Gamma\left(\frac{n+1}{2}\right)\sqrt{\pi}}{\Gamma\left(\frac{n}{2} + 1\right)2} \quad n > -1}\tag{C.65}$$

The relations of (C.65) are known as *Wallis's formulas*.

C.2 The Gamma Distribution

One of the most common probability distributions is the *gamma distribution* which is defined as

Appendix C The Gamma and Beta Functions and Distributions

$$f(x, n, \beta) = \frac{x^{n-1} e^{-x/\beta}}{\beta^n \Gamma(n)} \quad x > 0, \quad n, \beta > 0 \quad (\text{C.66})$$

A detailed discussion of this probability distribution is beyond the scope of this book; it will suffice to say that it is used in reliability and queuing theory. When n is a positive integer, it is referred to as *Erlang distribution*. Figure C.4 shows the probability density function (pdf) of the gamma distribution for $n = 3$ and $\beta = 2$.

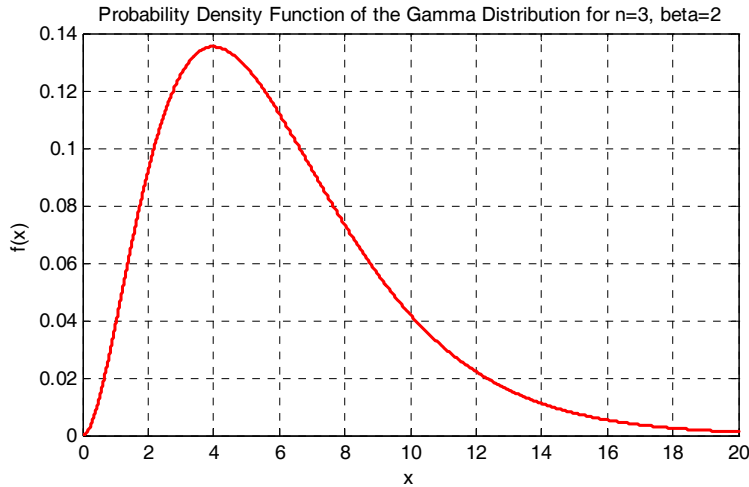


Figure C.4. The pdf for the gamma distribution

The plot of Figure C.4 was constructed with the MATLAB script below.

```
x=0:0.005:20; n=3; b=2; g=x.^(n-1).*exp(-x./b)/(b.^n.*gamma(n));...
plot(x,g); xlabel('x'); ylabel('f(x)');...
title('Probability Density Function of the Gamma Distribution for n=3, beta=2'); grid
```

We can also evaluate the gamma distribution with the Excel **GAMMADIST** function whose syntax is

=GAMMADIST(x,alpha,beta,cumulative)

where:

x = value at which the distribution is to be evaluated

alpha = the parameter n in (C.66)

beta = the parameter β in (C.66)

cumulative = a TRUE / FALSE logical value; if TRUE, **GAMMADIST** returns the cumulative distribution function (cdf), and if FALSE, it returns the probability density function (pdf).

Example C.12

Use Excel's =GAMMADIST function to evaluate $f(x)$, that is, the pdf of the *gamma distribution* if:

- a. $x = 4, n = 3, \text{ and } \beta = 2$
- b. $x = 7, n = 3, \text{ and } \beta = 2$

Solution:

Since we are interested in the probability density function (pdf) values, we specify the FALSE condition. Then,

- a. =GAMMADIST(4,3,2,FALSE) returns 0.1353
- b. =GAMMADIST(7,3,2,FALSE) returns 0.0925

We observe that these values are consistent with the plot of Figure C.4.

C.3 The Beta Function

The *beta function*, denoted as $B(m, n)$, is defined as

$$B(m, n) = \int_0^1 x^{m-1} (1-x)^{n-1} dx \tag{C.67}$$

where $m > 0$ and $n > 0$.

Example C.13

Prove that

$$B(m, n) = B(n, m) \tag{C.68}$$

Proof:

Let $x = 1 - y$; then, $dx = -dy$. We observe that as $x \rightarrow 0, y \rightarrow 1$ and as $x \rightarrow 1, y \rightarrow 0$. Therefore,

$$\begin{aligned} B(m, n) &= \int_0^1 x^{m-1} (1-x)^{n-1} dx = -\int_1^0 (1-y)^{m-1} [1-(1-y)]^{n-1} dy \\ &= \int_0^1 (1-y)^{m-1} y^{n-1} dy = \int_0^1 y^{n-1} (1-y)^{m-1} dy = B(n, m) \end{aligned}$$

and thus (C.68) is proved.

Appendix C The Gamma and Beta Functions and Distributions

Example C.14

Prove that

$$B(m, n) = 2 \int_0^{\pi/2} \cos^{2m-1}\theta \cdot \sin^{2n-1}\theta d\theta \quad (\text{C.69})$$

Proof:

We let $x = \sin^2\theta$; then, $dx = 2\sin\theta\cos\theta d\theta$. We observe that as $x \rightarrow 0$, $\theta \rightarrow 0$ and as $x \rightarrow 1$, $\theta \rightarrow \pi/2$. Then,

$$\begin{aligned} B(m, n) &= \int_0^1 x^{m-1}(1-x)^{n-1} dx = 2 \int_0^{\pi/2} (\sin^2\theta)^{m-1} (\cos^2\theta)^{n-1} \sin\theta\cos\theta d\theta \\ &= 2 \int_0^{\pi/2} (\sin^{2m-1}\theta)(\cos^{2n-1}\theta) d\theta \end{aligned} \quad (\text{C.70})$$

Example C.15

Prove that

$$B(m, n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} \quad (\text{C.71})$$

Proof:

The proof is evident from (C.62) and (C.70).

The $B(m, n)$ function is also useful in evaluating certain integrals as illustrated by the following examples.

Example C.16

Evaluate the integral

$$\int_0^1 x^4(1-x)^3 dx \quad (\text{C.72})$$

Solution:

By definition

$$B(m, n) = \int_0^1 x^{m-1}(1-x)^{n-1} dx$$

and thus for this example,

$$\int_0^1 x^4(1-x)^3 dx = B(5, 4)$$

Using (C.71) we obtain

$$B(5, 4) = \frac{\Gamma(5)\Gamma(4)}{\Gamma(9)} = \frac{4!3!}{8!} = \frac{24 \times 6}{40320} = \frac{144}{40320} = \frac{1}{280} \tag{C.73}$$

We can also use the MATLAB **beta(m,n)** function. For this example,

```
format rat; % display answer in rational format
z=beta(5,4)
```

```
z =
    1/280
```

Excel does not provide a function that computes the B(m, n) function directly. However, we can use (C.71) for its computation as shown in Figure C.5.

	$\Gamma(m)$ exp(gammaaln(m))	$\Gamma(n)$ exp(gammaaln(n))	$\Gamma(m+n)$ exp(gammaaln(m+n))	Beta(m,n)= $\Gamma(m) \times \Gamma(n) / \Gamma(m+n)$
m= 5	24.00	6.00	40320.00	1/280
n= 4				

Figure C.5. Computation of the beta function with Excel.

Example C.17

Evaluate the integral

$$\int_0^2 \frac{x^2}{\sqrt{2-x}} dx \tag{C.74}$$

Solution:

Let $x = 2v$; then $x^2 = 4v^2$, and $dx = 2dv$. We observe that as $x \rightarrow 0$, $v \rightarrow 0$, and as $x \rightarrow 2$, $v \rightarrow 1$. Then, (C.74) becomes

$$\begin{aligned} \int_0^1 \frac{4v^2}{\sqrt{2-2v}} 2dv &= \frac{8}{\sqrt{2}} \int_0^1 \frac{v^2}{\sqrt{1-v}} dv = 4\sqrt{2} \int_0^1 v^2(1-v)^{-1/2} dv \\ &= 4\sqrt{2} \cdot B\left(3, \frac{1}{2}\right) = 4\sqrt{2} \frac{\Gamma(3)\Gamma(1/2)}{\Gamma(7/2)} \end{aligned} \tag{C.75}$$

where

$$\begin{aligned} \Gamma(3) &= 2! \\ \Gamma\left(\frac{1}{2}\right) &= \sqrt{\pi} \\ \Gamma\left(\frac{7}{2}\right) &= \left(\frac{7}{2}-1\right)\Gamma\left(\frac{7}{2}-1\right) = \frac{5}{2}\Gamma\left(\frac{5}{2}\right) = \frac{5}{2}\left(\frac{5}{2}-1\right)\Gamma\left(\frac{5}{2}-1\right) \\ &= \frac{5}{2} \cdot \frac{3}{2} \cdot \left(\frac{3}{2}-1\right)\Gamma\left(\frac{3}{2}-1\right) = \frac{15}{8}\Gamma\left(\frac{1}{2}\right) = \frac{15}{8}\sqrt{\pi} \end{aligned} \tag{C.76}$$

Then, from (C.74), (C.75) and (C.76) we obtain

$$\int_0^2 \frac{x^2}{\sqrt{2-x}} dx = \frac{4\sqrt{2} \cdot 2! \cdot \sqrt{\pi}}{(15/8)\sqrt{\pi}} = \frac{64\sqrt{2}}{15} \tag{C.77}$$

C.4 The Beta Distribution

The *beta distribution* is defined as

$$f(x, m, n) = \frac{x^{m-1}(1-x)^{n-1}}{B(m, n)} \quad x < 0 < 1, \quad m, n > 0 \tag{C.78}$$

A plot of the beta probability density function (pdf) for $m = 3$ and $n = 2$, is shown in Figure C.6.

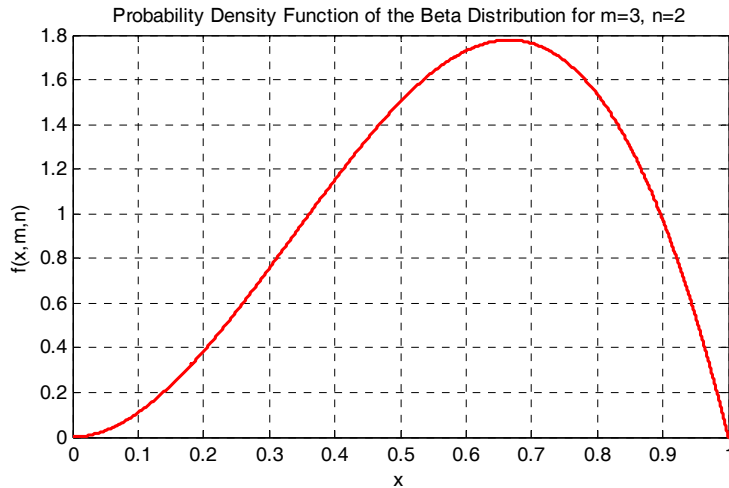


Figure C.6. The pdf of the beta distribution

The plot of Figure C.6 was constructed with the MATLAB script below.

```
x=0:0.0005:1; m=3; n=2; b=x.^(m-1).*(1-x).^(n-1)./beta(3,2);...
plot(x,b); xlabel('x'); ylabel('f(x,m,n)');...
title('Probability Density Function of the Beta Distribution for m=3, n=2'); grid
```

As with the gamma probability distribution, a detailed discussion of the beta probability distribution is beyond the scope of this book; it will suffice to say that it is used in computing variations in percentages of samples such as the percentage of the time in a day people spent at work, driving habits, eating times and places, etc.

Using (C.71) we can express the beta distribution as

$$f(x, m, n) = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} \cdot x^{m-1}(1-x)^{n-1} \quad x < 0 < 1, \quad m, n > 0 \quad (\text{C.79})$$

We can evaluate the beta *cumulative* distribution function (cdf) with Excel's BETADIST function whose syntax is

=BETADIST(x,alpha,beta,A,B)

where:

x = value between **A** and **B** at which the distribution is to be evaluated

alpha = the parameter **m** in (C.79)

beta = the parameter **n** in (C.79)

A = the lower bound to the interval of **x**

B = the upper bound to the interval of **x**

From the plot of Figure C.6, we see that when $x = 1$, $f(x, m, n)$ which represents the probability density function, is zero. However, the cumulative distribution (the *area* under the curve) at this point is 100% or unity since this is the upper limit of the x -range. This value can be verified by

=BETADIST(1,3,2,0,1)

which returns 1.0000

Appendix D

Introduction to Markov Chains

This appendix introduces the concept of Markov chains as applied to probability and statistics. It is defined in terms of probability vectors and stochastic matrices. These concepts are illustrated with several examples.

D.1 Stochastic Processes

A *stochastic process* is a finite sequence of experiments in which each experiment has a finite number of outcomes with given probabilities. Alternately, a stochastic process is a family of random variables.

Definition 1

The vector

$$u = \{u_1, u_2, \dots, u_n\} \quad (\text{D.1})$$

is called a *probability vector* if all of its components are non-negative and their sum is unity.

Example D.1

Determine which of the following are probability vectors.

- a. $x = \{0.75 \quad 0 \quad -0.25 \quad 0.25\}$
- b. $y = \{0.75 \quad 0.25 \quad 0 \quad 0.25\}$
- c. $z = \{0.25 \quad 0.25 \quad 0 \quad 0.50\}$

Solution:

According to Definition 1, only z is a probability vector since all of its components are non-negative and their sum is unity.

D.2 Stochastic Matrices

Definition 2

A *stochastic matrix* is a square matrix where each of its rows is a probability vector.

Example D.2

Determine which of the following are stochastic matrices.

$$X = \begin{bmatrix} 1/4 & 3/4 \\ 1/3 & 1/3 \end{bmatrix} \quad Y = \begin{bmatrix} 1/3 & 0 & 2/3 \\ 3/4 & 1/2 & -1/4 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \quad Z = \begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 1/6 & 1/3 \\ 1/3 & 2/3 & 0 \end{bmatrix}$$

Solution:

According to Definition 2, only Z is a stochastic matrix, since each of its rows is a probability vector.

Theorem 1

If A and B are stochastic matrices, their product AB results in a stochastic matrix also.

Theorem 2

If A is a stochastic matrix, all powers of A , that is, all A^n matrices are stochastic matrices.

Definition 3

A stochastic process which satisfies the following two properties, in a sequence of trials whose outcomes are X_1, X_2, \dots, X_n , is called a finite *Markov chain*.

1. Each outcome belongs to a finite set of outcomes x_1, x_2, \dots, x_n that is referred to as the *state space* of the system. If the outcome on the i th trial is x_i , then we say that the system is in state x_i at step i or it is at the i th step.
2. The outcome of any trial depends, at most, upon the outcome of the immediately preceding trial, and not upon any other previous outcome. With each pair of states (x_i, x_j) , there is a given probability p_{ij} that x_j occurs immediately after x_i occurs.

The numbers p_{ij} are called transition probabilities and are arranged in a matrix that is called *transition matrix* P , and it is arranged as shown in (D.2) below.

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2n} \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{m1} & p_{m2} & p_{m3} & \cdots & p_{mn} \end{bmatrix} \quad (\text{D.2})$$

Theorem 3

The transition matrix P of a Markov chain is a stochastic matrix.

Example D.3

An employee either brings his brown bag (lunch), or goes out to eat at a nearby restaurant during the work week. He never goes out to eat two consecutive days, but if he brings his brown bag, the next day he may again bring his brown bag, or go out to eat with equal probability.

- Determine the state space of this situation.
- Determine whether this stochastic process is a Markov chain.
- If it is a Markov chain, express it as a transition matrix.

Solution:

- The state space of this situation is $\{\mathbf{b}$ (brown bag), \mathbf{r} (restaurant) $\}$
- Since the outcome of any day depends only on what happened the preceding day, this stochastic process is a Markov chain.
- The transition matrix of this Markov chain is

$$P = \begin{bmatrix} & \mathbf{b} & \mathbf{r} \\ \mathbf{b} & 1/2 & 1/2 \\ \mathbf{r} & 1 & 0 \end{bmatrix} \quad (\text{D.3})$$

The \mathbf{b} row of the matrix of (D.3) represents the fact that if he brings his brown bag in any particular day, the next day he will bring his brown bag, or eat out at a restaurant with equal probability. The \mathbf{r} row indicates the condition that if he eats out in any particular day, he definitely will not go out to a restaurant on the next day.

Definition 4

A state x_i y_i in a Markov chain is called *absorbing state*, if the system remains in that state once it enters it. The state x_i y_i is absorbing if and only if the i th row of a transition matrix P , has 1 on the main diagonal and zeros everywhere else in that row.

Example D.4

Determine whether or not, the transition matrix of the Markov chain of (D.4) below has any absorbing states.

$$P = \begin{bmatrix} & \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \mathbf{x}_4 & \mathbf{x}_5 \\ \mathbf{y}_1 & 1/4 & 0 & 1/4 & 1/4 & 1/4 \\ \mathbf{y}_2 & 0 & 1 & 0 & 0 & 0 \\ \mathbf{y}_3 & 1/2 & 0 & 1/4 & 1/4 & 0 \\ \mathbf{y}_4 & 0 & 1 & 0 & 0 & 0 \\ \mathbf{y}_5 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{D.4})$$

Solution:

According to Definition 4, the transition matrix of (D.4) has two absorbing states, x_2 y_2 and x_5 y_5 since they appear in the main diagonal, and the rows in which they appear have zeros everywhere else.

D.3 Transition Diagrams

The transition probabilities of a Markov chain are often represented by a *transition diagram* such as that shown in Figure D.1.

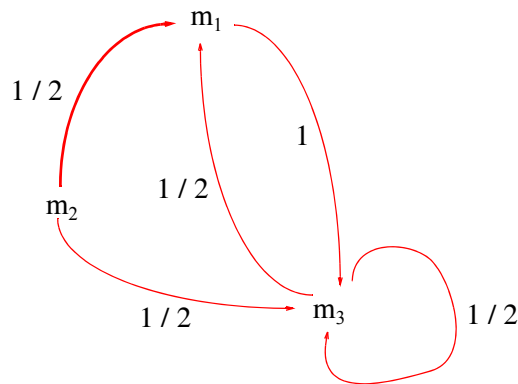


Figure D.1. Transition diagram showing the transition probabilities of a Markov chain.

For this transition diagram, the state space is $\{m_1, m_2, m_3\}$ and the transition matrix is

$$P = \begin{bmatrix} & \mathbf{m}_1 & \mathbf{m}_2 & \mathbf{m}_3 \\ \mathbf{m}_1 & 0 & 0 & 1 \\ \mathbf{m}_2 & 1/2 & 0 & 1/2 \\ \mathbf{m}_3 & 1/2 & 0 & 1/2 \end{bmatrix} \quad (\text{D.5})$$

D.4 Regular Stochastic Matrices

Definition 5

A stochastic matrix P is said to be *regular*, if all elements of a power matrix P^m are greater than zero.

Example D.5

Determine whether or not the matrices A and B below are regular.

$$A = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} \quad (D.6)$$

Solution:

For Matrix A

$$A^2 = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 1/4 \end{bmatrix}$$

Since all elements of the matrix A^2 are greater than zero, Matrix A is said to be regular.

For Matrix B

$$B^2 = \begin{bmatrix} 1 & 0 \\ 3/4 & 1/4 \end{bmatrix} \quad B^3 = \begin{bmatrix} 1 & 0 \\ 7/8 & 1/8 \end{bmatrix} \quad B^4 = \begin{bmatrix} 1 & 0 \\ 15/16 & 1/16 \end{bmatrix}$$

Since B^m has a 0 in the first row for $m \geq 1$, Matrix B is not regular.

Theorem 4

If P is a regular stochastic matrix, then

- P has a unique fixed probability vector t , and the components of t are all greater than zero.
- the sequence P, P^2, P^3, \dots, P^m approaches the matrix T which has the same dimension as P and whose rows are each the fixed probability vector t .
- if p is any probability vector, the sequence of vectors $pP, pP^2, pP^3, \dots, pP^m$ approaches the fixed probability vector t .

Example D.6

In Example D.5 we found that the matrix A is a regular stochastic matrix. Let this same matrix be denoted as P , that is,

$$P = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$$

We want to find a probability vector t with two components x and $1 - x$, that is, $t = \{x, 1 - x\}$ so that $tP = t$.

Solution:

In matrix form, the condition $tP = t$ is

$$\begin{bmatrix} x & 1 - x \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} x & 1 - x \end{bmatrix}$$

or

$$\begin{bmatrix} 0 + \frac{1}{2}(1 - x) & x + \frac{1}{2}(1 - x) \end{bmatrix} = \begin{bmatrix} x & 1 - x \end{bmatrix}$$

or

$$\begin{bmatrix} \frac{1}{2} - \frac{1}{2}x & \frac{1}{2} + \frac{1}{2}x \end{bmatrix} = \begin{bmatrix} x & 1 - x \end{bmatrix}$$

Equating corresponding elements of these matrices, we get

$$\frac{1}{2} - \frac{1}{2}x = x$$

$$\frac{1}{2} + \frac{1}{2}x = 1 - x$$

and either of these equations yields $x = 1/3$. Therefore,

$$t = \begin{bmatrix} x & 1 - x \end{bmatrix} = \begin{bmatrix} 1/3 & 1 - 1/3 \end{bmatrix} = \begin{bmatrix} 1/3 & 2/3 \end{bmatrix}$$

and thus, t is the unique fixed probability vector of the regular stochastic matrix P .

Next, by the second property of Theorem 4, the sequence P, P^2, P^3, \dots, P^m approaches the matrix T that has the same dimension as P , and whose rows are each the fixed probability vector t . Thus,

$$T = \begin{bmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{bmatrix} = \begin{bmatrix} 0.33 & 0.67 \\ 0.33 & 0.67 \end{bmatrix} \quad (D.7)$$

The same result can be obtained by the sequence P, P^2, P^3, \dots, P^m , that is,

$$P^2 = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$$

$$P^3 = P^2 \cdot P = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/4 & 3/4 \\ 3/8 & 5/8 \end{bmatrix}$$

$$P^4 = P^3 \cdot P = \begin{bmatrix} 1/4 & 3/4 \\ 3/8 & 5/8 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 3/8 & 5/8 \\ 5/16 & 11/16 \end{bmatrix}$$

$$P^5 = P^4 \cdot P = \begin{bmatrix} 3/8 & 5/8 \\ 5/16 & 11/16 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 5/16 & 11/16 \\ 11/32 & 21/32 \end{bmatrix}$$

or

$$P^5 = \begin{bmatrix} 0.31 & 0.69 \\ 0.34 & 0.66 \end{bmatrix} \quad (D.8)$$

Comparison of (D.8) with (D.7), reveals that for higher powers of P , we obtain values closer to T .

D.5 Some Practical Examples

Example D.7

A sales engineer sells computers and accessories in three cities, San Francisco (SF), San Jose (SJ), and Oakland (Oak), and spends the entire day in any of these cities. If he spends one day in SF, he will spend the next day in SJ, and if he spends his day in SJ or Oak, he is twice as likely to spend the next day in SF as in the other two cities. Compute the percentages of the time he spends in each of these cities over an extended period of time.

Solution:

We can model this situation as a three-state Markov chain. The transition matrix for this example is

$$P = \begin{bmatrix} & \mathbf{SF} & \mathbf{SJ} & \mathbf{Oak} \\ \mathbf{SF} & 0 & 1 & 0 \\ \mathbf{SJ} & 2/3 & 0 & 1/3 \\ \mathbf{Oak} & 2/3 & 1/3 & 0 \end{bmatrix}$$

A unique probability vector t of the matrix P can be any vector of the form $u = \{x \ y \ z\}$, i.e.,

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

$$\begin{bmatrix} \frac{2}{3}y + \frac{2}{3}z & x + \frac{1}{3}z & \frac{1}{3}y \end{bmatrix} = \begin{bmatrix} x & y & z \end{bmatrix}$$

Equating corresponding elements of these matrices we get

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

These equations yield the trivial solution $x = y = z = 0$. To obtain a meaningful solution, we group any two of the equations of (D.9) with the condition $x + y + z = 1$. This condition, grouped with the first two equations of (D.9) yields

$$x + y + z = 1 \quad -x + \frac{2}{3}y + \frac{2}{3}z = 0 \quad x - y + \frac{1}{3}z = 0 \quad (\text{D.9})$$

We will use Excel to find the values of the unknowns. The spreadsheet is shown in Figure D.2.

	A	B	C	D	E	F	G	H
1	Spreadsheet for Matrix Inversion and Matrix Multiplication							
2	Example 11.7							
3								
4		1	1	1			1	
5	A=	-1	2/3	2/3		B=	0	
6		1	-1	1/3			0	
7								
8		2/5	-3/5	0			0.400	
9	A ⁻¹ =	4/9	-1/3	-3/4		X=	0.450	
10		1/7	1	3/4			0.150	

Figure D.2. Spreadsheet for the solution of the system of equations of (D.10)

The step-by-step procedure is provided in Chapter 2, Example 2.18, Page 2–23.

Thus, the desired probability vector t is $t = \{0.40 \quad 0.45 \quad 0.15\}$ and this indicates that, over an extended period of time he spends 40% of the time in San Francisco, 45% in San Jose, and 15% in Oakland.

Example D.8

An electronics test technician, henceforth referred to as board tester, spends all his time on the test bench testing personal computer motherboards and, on the average, it takes 30 minutes to complete the testing of each motherboard. A pickup/delivery man, working for the same company, comes to the test bench every 30 minutes to pick up the boards that the board tester has completed, and delivers new boards to be tested. The number of untested boards brought to the test board tester’s bench varies, but it is estimated that 30% of the time, the pickup/delivery man brings no boards, 50% of the time brings one board, and 20% of the time he brings two boards. Accordingly, at no time there should be more than three untested boards on the test bench, including the board which is currently being tested.

Determine the percentage of time that the test board tester is idle, assuming that any untested boards on the bench at the end of the work shift remain there for the next working day.

Solution:

We can model this process as a three-state Markov chain, where the states represent the number of untested boards on the bench just before the pickup/delivery man arrives.

Let s_0 represent the state that no untested boards remain on the bench, s_1 the state that there is one untested board, and s_2 the state that there are two untested boards just before the pickup/delivery man arrives. We want to find out the percentage of the time the board tester is idle. As a first step, we construct the state transition diagram shown in Figure D.3.

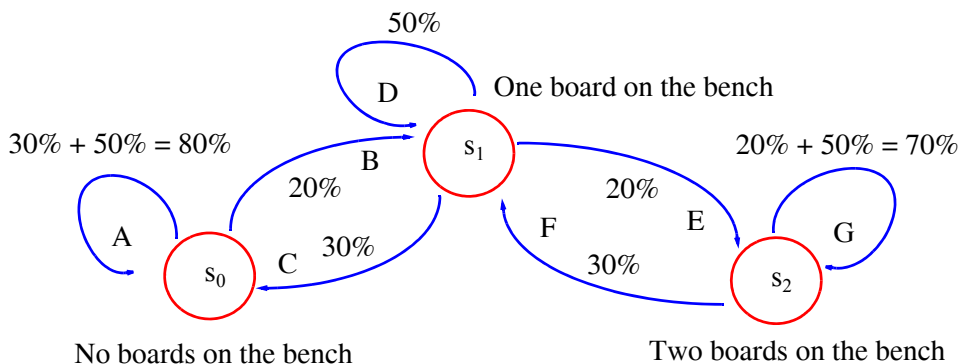


Figure D.3. State transition diagram for Example D.8.

Appendix D Introduction to Markov Chains

The state transition diagram was constructed with the assumption that the untested boards are on the bench *just before the pickup/delivery man arrives*. Let us assume that the board tester is idle (no testing is performed and no boards are on the bench), and the pickup/delivery man arrives. He may bring no boards (30% probability), one board (50% probability) or two boards (20% probability). In the case where the pickup/delivery man brings no boards (30%), or one board (50%), the next time he comes there will be no untested boards on the bench. The probability of this happening is $30\% + 50\% = 80\%$; this is denoted as Path A on the state transition diagram. But if the pickup/delivery man leaves two boards, while the board tester is idled, there will be one board untested on the bench when the pickup/delivery man arrives again. The probability of this happening is 20% and it is denoted as Path B.

Next, suppose that the board tester is not idled, that is, he is testing a board when the pickup/delivery man arrives. If he brings no boards (probability 30%), there will be no boards untested when he comes again, and thus the board tester will be idled. This condition is shown by Path C. However, if the pickup/delivery man brings one board (probability 50%), there will be one board untested when he comes again; this condition is shown by Path D. The other paths are drawn using similar reasoning.

Now, using the state transition diagram, we construct the following transition matrix.

$$P = \begin{bmatrix} s_0 & s_1 & s_3 \\ s_0 & 0.8 & 0.2 & 0 \\ s_1 & 0.3 & 0.5 & 0.2 \\ s_3 & 0 & 0.3 & 0.7 \end{bmatrix} \quad (\text{D.10})$$

The next step is to find the probability vector t from

$$[x \ y \ z] \cdot \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.3 & 0.5 & 0.2 \\ 0 & 0.3 & 0.7 \end{bmatrix} = [x \ y \ z]$$

Multiplication of the matrices on the left side yields

$$[0.8x + 0.3y + 0 \quad 0.2x + 0.5y + 0.3z \quad 0 + 0.2y + 0.7z] = [x \ y \ z]$$

Equating corresponding elements of these matrices we get

$$\begin{aligned} 0.8x + 0.3y &= x \\ 0.2x + 0.5y + 0.3z &= y \\ 0.2y + 0.7z &= z \end{aligned} \quad (\text{D.11})$$

or

$$\begin{aligned}
 -0.2x + 0.3y &= 0 \\
 0.2x - 0.5y + 0.3z &= 0 \\
 0.2y - 0.3z &= 0
 \end{aligned}
 \tag{D.12}$$

These equations yield the trivial solution $x = y = z = 0$. To obtain a meaningful solution, we group any two of the equations of (D.13) with the condition $x + y + z = 1$. Grouping this condition with the first two equations of (D.13) we get

$$\begin{aligned}
 x + y + z &= 1 \\
 -0.2x + 0.3y &= 0 \\
 0.2x - 0.5y + 0.3z &= 0
 \end{aligned}
 \tag{D.13}$$

We will use Excel to find the values of the unknowns. The step-by-step procedure was given in Chapter 2, Example 2.18. The spreadsheet is shown in Figure D.4.

	A	B	C	D	E	F	G	H
1	Spreadsheet for Matrix Inversion and Matrix Multiplication							
2	Example 11.8							
3								
4		1.00	1.00	1.00			1.00	
5	A=	-0.20	0.30	0.00		B=	0.00	
6		0.20	-0.50	0.30			0.00	
7								
8		0.47	-4.21	-1.58			0.47	
9	A ⁻¹ =	0.32	0.53	-1.05		X=	0.32	
10		0.21	3.68	2.63			0.21	

Figure D.4. Spreadsheet for the solution of the system of equations of (D.13).

Thus, the desired probability vector t is $t = \{0.47 \quad 0.32 \quad 0.21\}$ and the first value represents the percentage of state s_0 (no untested boards on the bench). This indicates that over an extended period, the board tester starts a 30 minute interval in state s_0 0.47 of the time. Since the probability that the pickup/delivery man brings no untested boards to him is 30%, the board tester is idle for $0.47 \times 0.3 = 0.141$ or 14.1% of his daily shift.

Example D.9

ABC Technologies is a contract manufacturer for personal computer modems. The three major steps in the manufacturing process in succession are:

1. Attaching small devices (resistors, capacitors, diodes, etc.) to the modem by the soldering process.

2. Inserting integrated circuit (IC) components to the modem.
3. Perform a final test to verify proper operation.

When each of the above three steps is completed, the units are tested to determine whether they will be forwarded to the next step, if they pass the test, or to the rework area if they fail the test. The testing procedure is shown by the flow chart of Figure D.5.

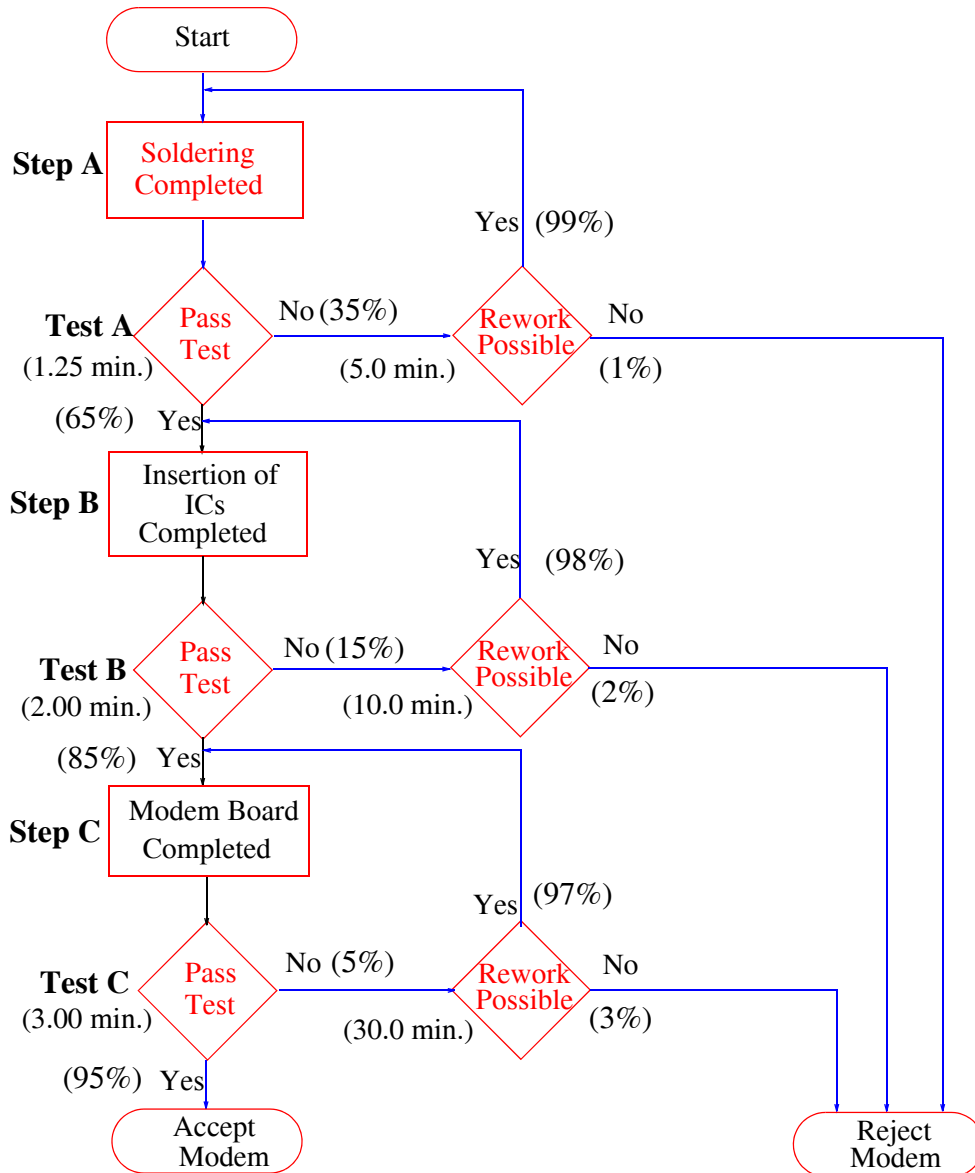


Figure D.5. Flow chart for Example D.9

Shown also are the probability percentages that a modem will go from one stage to the next. For instance, the flow chart indicates that there is a 65% probability that the modem will pass the test and thus go from Stage A to Stage B. Likewise, it is shown that there is a 99% probability that rework is possible at Stage A. These probabilities are normally provided by production engineers. The nominal times for a modem to go through each test and the rework times are also shown on the flow chart. The nominal times are also provided by production engineers.

Using a Markov chain procedure, compute total modem reject rate.

Solution:

We start the Markov chain with the construction of a matrix of probabilities that a modem will move from one test to the next.

Table D.1 shows the probabilities given on the flow chart of Figure D.5.

TABLE D.1 Probabilities that a modem will move from one test to the next

1	2	3	Probability from 1 to 2	Probability from 1 to 3
Test A	Step B	Rework and retest A	0.65	0.35
Test B	Test C	Rework and retest B	0.85	0.15
Test C	Accept	Rework and retest C	0.95	0.05
Rework, return to A	Test A	Reject	0.99	0.01
Rework, return to B	Test B	Reject	0.98	0.02
Rework, return to C	Test C	Reject	0.97	0.03

Table D.2 shows the information of Table D.1 with the probabilities into a matrix form.

TABLE D.2 Array of probabilities

	Test A	Test B	Test C	Rework, return to A	Rework, return to B	Rework, return to C	Accept	Reject
Test A	0.00	0.65	0.00	0.35	0.00	0.00	0.00	0.00
Test B	0.00	0.00	0.85	0.00	0.15	0.00	0.00	0.00
Test C	0.00	0.00	0.00	0.00	0.00	0.05	0.95	0.00
Rework, return to A	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Rework, return to B	0.00	0.98	0.00	0.00	0.00	0.00	0.00	0.02
Rework, return to C	0.00	0.00	0.97	0.00	0.00	0.00	0.00	0.03

The conditions of the left most column are also listed on top of the table, since these can be either a source or a destination. The two right most columns show the probabilities for the two absorbing states, accept or reject. These are the states at which there is no exit once they have entered these states.

Appendix D Introduction to Markov Chains

Next, we construct the 6×6 matrix shown on Table D.3, where there is a 1.00 in the main diagonal, that is, where the row and column headings are the same, and the other values are the negative values corresponding to the values of Table D.2. This procedure is based on eigenvalues, eigenvectors, and matrix decomposition methods. These topics are discussed in advanced theory of Markov chain textbooks.

TABLE D.3 The 6 by 6 matrix to be inverted

	Test A	Test B	Test C	Rework, return to A	Rework, return to B	Rework, return to C
Test A	1.00	-0.65	0.00	-0.35	0.00	0.00
Test B	0.00	1.00	-0.85	0.00	-0.15	0.00
Test C	0.00	0.00	1.00	0.00	0.00	-0.05
Rework, return to A	-0.99	0.00	0.00	1.00	0.00	0.00
Rework, return to B	0.00	-0.98	0.00	0.00	1.00	0.00
Rework, return to C	0.00	0.00	-0.97	0.00	0.00	1.00

Now, using an Excel spreadsheet, we invert the matrix of Table D.3, and we multiply the inverted matrix by the 6×2 matrix of the last two columns (Accept and Reject) of Table D.2. The result is shown in the spreadsheet of Figure D.6. The values in J12:J17 show the probabilities that a modem will be accepted, whereas the values in K12:K17 indicate the probabilities that a modem will be rejected. For instance, the probability that a modem will go to the acceptance area following Test A is 98.96%, and the probability that it will be rejected is 1.04%

We can also determine which production steps contribute most to the reject area.

	A	B	C	D	E	F	G	H	I	J	K
1	Spreadsheet for Matrix Inversion and Matrix Multiplication										
2	Example 11.8										
3			Test A	Test B	Test C	Redo A	Redo B	Redo C		Accept	Reject
4		Test A	1.00	-0.65	0.00	-0.35	0.00	0.00		0.00	0.00
5		Test B	0.00	1.00	-0.85	0.00	-0.15	0.00		0.00	0.00
6	Matrix	Test C	0.00	0.00	1.00	0.00	0.00	-0.05		0.95	0.00
7	P	Redo A	-0.99	0.00	0.00	1.00	0.00	0.00		0.00	0.01
8		Redo B	0.00	-0.98	0.00	0.00	1.00	0.00		0.00	0.02
9		Redo C	0.00	0.00	-0.97	0.00	0.00	1.00		0.00	0.03
10											
11			Test A	Test B	Test C	Redo A	Redo B	Redo C		Accept	Reject
12		Test A	1.530	1.166	1.042	0.536	0.175	0.052		0.9896	0.0104
13		Test B	0.000	1.172	1.047	0.000	0.176	0.052		0.9949	0.0051
14	Matrix	Test C	0.000	0.000	1.051	0.000	0.000	0.053		0.9984	0.0016
15	P ⁻¹	Redo A	1.515	1.154	1.031	1.530	0.173	0.052		0.9797	0.0203
16		Redo B	0.000	1.149	1.026	0.000	1.172	0.051		0.9750	0.0250
17		Redo C	0.000	0.000	1.019	0.000	0.000	1.051		0.9685	0.0315

Figure D.6. Matrix Inversion and Multiplication

For example, to find out how much of the 1.04% rejection rate originates at Test A, we divide the percentage of units which go from Test A to acceptance (98.96%) by the percentage that goes from Test B to acceptance (99.49%). The result is $98.96/99.49 = 0.9947 = 99.47\%$. This is the percentage of units that go from Test A to Test B.

To find out the percentage of the rejected modems from Test A, we subtract this value from unity, that is, $1 - 0.9947 = 0.0053 = 0.53\%$

To find out how much of the rejection rate originates at Test B, we divide the percentage of units which go from Test B to acceptance (99.49%) by the percentage of units which go from Test C to acceptance (99.84%). The result is $99.49/99.84 = 0.9965 = 99.65\%$

To determine the percentage of the rejected modems from Test B, we subtract this value from unity, that is, $1 - 0.9947 = 0.0035 = 0.35\%$

The percentage of units which are rejected from Test C is found in K14, i.e., 0.16%. We can verify that these percentages are correct by adding them. Thus, $0.53\% + 0.35\% + 0.16\% = 1.04\%$.

Row 12 of Figure D.6 contains some additional and useful information. These values can be used to compute the time it takes for a modem to pass through each step. Let us assume, for example, that the nominal time (furnished by production engineers) for a modem to go through Test A is 1.25 minutes. But since some of the units are going through the rework stage, the actual time is longer by the factor shown in C12 of the spreadsheet of Figure D.6, i.e., $1.25 \times 1.53 = 1.91$ min.

Similarly, if we assume that the nominal times for Tests B and C are 2.00 and 3.00 minutes respectively, these are increased by the factors shown in D12 and E12 and thus the actual times are $2.000 \times 1.116 = 2.23$ min. for Test B and $3.000 \times 1.042 = 3.13$ min. for Test D.

Let us also assume that the nominal times for the three rework stages are 5, 10, and 30 minutes for the rework A, B, and C steps respectively. Then, the actual times are found by multiplication of these nominal times by the values of F12, G12, and H12. Therefore, the actual times for these tasks are:

$$5.000 \times 0.536 = 2.68 \text{ min.}$$

$$10.000 \times 0.175 = 1.75 \text{ min.}$$

and

$$30.000 \times 0.052 = 1.56 \text{ min.}$$

Under ideal conditions, no modems would go through the rework stages and thus the total nominal time is $1.25 + 2.00 + 3.00 = 6.25$ minutes per product. However, the total actual time includes the time for rework and therefore, it is $1.91 + 2.23 + 3.13 + 2.68 + 1.75 + 1.56 = 13.36$ minutes per product. In other words, the actual time is $13.36/6.25 = 2.14$ times greater than the

Appendix D Introduction to Markov Chains

nominal. Accordingly, the manufacturing division of this company must make an effort to improve in this area.

Appendix E

The Lambda Index

This appendix introduces the Lambda index. Unlike standard forecasts that rely on traditional balance–sheet relationships, this index predicts liquidity. A spreadsheet example is provided to forecast liquidity of a hypothetical company.

E.1 Introduction

Briefly, the Lambda index predicts whether a company will have sufficient cash and credit to survive. It is a ratio that was developed by professor Gary Emery of the University of Oklahoma to pinpoint relevant components of liquidity, i.e., short–term balances and available credit. It provides a measure of the likelihood that a company would become insolvent.

Unlike standard cash forecasts that rely on traditional balance sheet relationships where cash and non–cash items are not differentiated and ignore cash flows and unused credit, the Lambda index incorporates uncertainty about cash flows in the form of a sample standard deviation from those cash flows. Accordingly, the Lambda index can be used like a z value from the standard normal distribution table. The higher the Lambda index, the less the probability that cash requirements will exceed cash on hand. Thus, a Lambda index of 1.64 represents a chance of one in twenty that cash requirements will exceed cash on hand, a Lambda index of 3.00 indicates that there is about one chance in a thousand, a Lambda index of 3.29 indicates that there is about one chance in two thousand, and a Lambda index of 3.90 indicates that there is about one chance in a twenty thousand.

E.2 The Lambda Index Defined

For a given period, the Lambda index is defined as

$$\frac{\text{Initial liquid reserve} + \text{Total anticipated net cash flow during the analysis period}}{\text{Uncertainty about net cash flow during the analysis period}}$$

In the equation above, the initial liquid reserve consists of cash balances, short term investments, and unused lines of credit. The anticipated net cash flow represents the difference between cash receipts and disbursements. The denominator represents the standard deviation of net cash flow.

The Lambda index calculated by the above equation is the equivalent to the z value in one tail of the normal distribution. The second tail is not relevant because a liquidity crisis results only from a negative deviation from the mean. To find out the odds that pertain to any given Lambda (z value), we look up the value in the standard normal distribution and add 0.5. Thus, the Lambda

Appendix E The Lambda Index

of 2.33 corresponds to a value of 0.4901 in the standard normal distribution table. Adding 0.5 results in the value of 0.9901, meaning that there is a 99% chance that problems will not occur, or a 1% chance that they will. We recall, from Chapter 10 that the area (probability), corresponding to this normalized z value, can be found from tables of the standard normal distribution or with the Excel's NORMSDIST function. In this case, we seek the area to the left of $z = 2.33$. Then, =NORMSDIST(2.33) returns 0.9901.

E.3 Spreadsheet Construction

In this section we will construct a worksheet that keeps a running monthly tally to show how changes affect future cash balances for a hypothetical company. Refer to the MS Excel spreadsheet below.

1	Twelve-Month Cash Forecast Summary												
2													
3		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
4		2008	2008	2008	2008	2008	2008	2008	2008	2008	2008	2008	2008
5													
6	Short-term line of credit	\$100,000	\$100,000	\$100,000	\$100,000	\$100,000	\$100,000	\$100,000	\$100,000	\$100,000	\$100,000	\$100,000	\$100,000
7	Beginning liquid assets	\$50,000	\$72,344	\$79,944	\$103,188	\$119,028	\$112,468	\$81,216	\$29,628	-\$11,232	\$2,612	\$48,964	\$106,332
8	Adjustments	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0
9	Initial liquid reserve	\$150,000	\$172,344	\$179,944	\$203,188	\$219,028	\$212,468	\$181,216	\$129,628	\$88,768	\$102,612	\$148,964	\$206,332
10													
11	Total sources of funds	\$203,800	\$196,800	\$222,000	\$200,800	\$164,800	\$122,800	\$100,400	\$116,744	\$178,800	\$214,476	\$234,800	\$258,800
12	Total uses of funds	\$181,456	\$189,200	\$198,756	\$184,960	\$171,360	\$154,052	\$151,988	\$157,604	\$164,956	\$168,124	\$177,432	\$182,488
13													
14	Net cash flow	\$22,344	\$7,600	\$23,244	\$15,840	-\$6,560	-\$31,252	-\$51,588	-\$40,860	\$13,844	\$46,352	\$57,368	\$76,312
15													
16	Ending liquid assets	\$72,344	\$79,944	\$103,188	\$119,028	\$112,468	\$81,216	\$29,628	-\$11,232	\$2,612	\$48,964	\$106,332	\$182,644
17	short-term debt, and												
18	adjustments (net)												
19													
20	Ending liquid reserve	\$172,344	\$179,944	\$203,188	\$219,028	\$212,468	\$181,216	\$129,628	\$88,768	\$102,612	\$148,964	\$206,332	\$282,644
21													
22	Standard deviation	N/A	\$10,426	\$8,784	\$7,234	\$12,357	\$21,002	\$28,781	\$29,833	\$28,812	\$31,691	\$34,701	\$38,949
23													
24	Lambda Index	N/A	17.26	23.13	30.28	17.19	8.63	4.50	2.98	3.56	4.70	5.95	7.26
25													
26	Additional cash required	N/A	-\$148,667	-\$176,837	-\$197,326	-\$175,397	-\$118,209	-\$43,286	\$731	-\$16,177	-\$53,891	-\$102,229	-\$165,797
27	to maintain Lambda of 3												

To construct the spreadsheet model above, perform the following steps:

1. Enter all labels and values shown in highlighted areas.
2. Enter the following formulas in the indicated cells:

C7: =B16 Copy cell C7 to range D7:M7

B9: =SUM(B6:B8) Copy cell B9 to range C9:M9

B14: =B11-B12

B16: =B7+B8+B14

B20: =B6+B16 Copy range B14:B20 to range C14:M14

B22: N/A Copy cell B22 to cells B24 and B26

C22: =STDEV(\$B14:C14)

C24: =C20/C22

C26: =(C22*3)-(C9+C14) Copy range C22:C26 to range D22:M22

When the spreadsheet is complete, it is highly recommended that the non-highlighted cells be protected from accidental changes or erasures.

It is assumed that the values in rows 11 and 12 have been derived from other worksheets of supporting data, covering items such as debt and equity financing, payroll, taxes, insurance, interest, dividends, and other sources and uses of funds. Row 6 contains the company's short-term line of credit, which management expects to remain at \$100,000 throughout the year. The value \$50,000 in cell B7 represents the company's short-term liquid assets, including cash and marketable securities, at the beginning of the analysis. In the following months, this number fluctuates with the net cash flow.

The adjustment cells in row 8 allow the cash manager to gauge the effect of adding or subtracting cash from the worksheet after the cash forecast has been entered. For simplicity, these cells initially contain zeros. Rows 11 and 12 contain revenue and expense forecasts for each month of the calendar year. Once these figures are entered, the formulas do the rest. The formulas in row 14 calculate net cash flow from the forecasted sources and uses of funds in rows 11 and 12. As long as net cash flow is positive and the short-term line of credit in row 6 remains stable, Lambda should increase, indicating improving liquidity.

The formulas in row 16 add the beginning liquid assets in row 7, the adjustments in row 8, and the net cash flows in row 14 to obtain the ending net short-term assets for the period. The formulas in row 20 calculate the liquid reserve at the end of the period by combining the short-term line of credit in row 6 with the ending liquid assets in row 16. The formulas in row 22 calculate the standard deviation of the net cash flows in row 14, the denominator for the Lambda index in row 24.

The formulas in row 26 calculate liquidity imbalance using a target Lambda of 3. A negative figure means that there may be extra cash on hand that should be invested elsewhere—depending on how comfortable management is with the accuracy of the cash forecast. A positive value indicates that more money or additional credit is required to keep liquidity at a comfortable level.

The spreadsheet shows that in February, the company will have \$148,667 more cash on hand than it needs to cover February's cash outlays and maintain a Lambda of 3 (cell C26). If management decides to put this cash to other use and enters an adjustment of -148,667 in cell C8, then February's Lambda would equal 3.

However, the worksheet clearly shows that unless the money is replaced later in the year, the

Appendix E The Lambda Index

long-term effects of such a move would be disastrous. The initial forecast for later months indicates that cash balances will already be on shaky ground in August, before recovering in the fall season. Therefore, subtracting \$148,667 in February would exacerbate the situation.

The Lambda index points out erratic cash flows, such as the one in this example, by including the standard deviation of cash flows in its denominator. The greater the fluctuation in cash flows, the smaller the Lambda for any given liquid reserve. Even though the company's liquid assets have more than doubled by the end of the year, December's Lambda is much smaller than February's. Thus, with a 12-month track record, Lambda becomes more reliable, that is, it reflects the kind of fluctuations that can be expected from the company's cash flow.

As the year 2008 progresses, management should substitute actual cash flow numbers for the monthly forecasts to fine-tune the Lambda index. The denominator hinges upon standard deviations in net cash flow, and actual numbers will produce a more accurate picture than forecast numbers. If, at the end of 2008, the company decides to apply Lambda to its cash forecast for 2009, it should extend the current model (in columns N through Y), using the actual cash flows reported in 2008 as the basis for calculating standard deviations in 2009.

The values in this example clearly indicate that the Lambda index is extremely sensitive to forecast accuracy. For example, if June's revenues were \$80,000 rather than the projected \$122,800, it would dramatically affect the Lambda index in the following months. Typically, a company is considered safe if it maintains a Lambda of 15 or higher.

Appendix F

The Black–Scholes Stock Options Model

This appendix is an introduction to the Black–Scholes stock options model. This model is used to determine the most probable value of a stock option based on volatility and time until expiration. It was developed by Fischer Black and Myron Scholes and has become the industry standard. This discussion is for educational purposes and academic interest only, and the reader should never invest in the options market unless he / she understands the risks involved. It is not an infallible guide in making investment decisions. A guidance from an experienced stock broker is always a must.

F.1 Stock Options

There are two kinds of options. The first is known as a *call option* which is an option to buy. When we purchase a call option, we obtain the right to buy a specified number of shares of a security (common stock), treasury bonds, or commodity futures* at some point within a specified time period at a fixed price, regardless of price changes within the specified time period.† This fixed price is referred to as the *exercise price*, or *strike price*.

Example F.1

Presently common stock of the XYZ Corporation trades at \$100 a share. Our stock broker tells us that we can buy shares of this stock at \$105 a share at any time in the next 100 days, and for this buying option the brokerage house charges us a premium of \$5 per share. Suppose that we buy 200 shares and at some time within the 100–day period this stock goes up to \$125 a share and we exercise the right to buy at this price. Since we bought this stock at \$110 (\$105 option price plus \$5 premium), our gain would be

$$\text{Gain} = \$125 \times 200 - \$110 \times 200 = 25,000 - 22,000 = \$3,000$$

However, if at the expiration date (end of 100-day period), the price of this stock drops to \$85 a share, we would have incurred a loss of $\$5 \times 200 = \$1,000$.

Next, let us suppose that instead of buying options, we bought 200 shares of this stock at \$100 a share, and we sold all shares when the stock went up to \$125 a share. In this case, our gain would be

* Stocks, treasury bonds, and commodity futures are often referred to as financial instruments.

† The American-style options differ from the European-style options in that European options can be exercised only on the day of expiration.

Appendix F The Black–Scholes Stock Options Model

$$\text{Gain} = \$125 \times 200 - \$100 \times 200 = 25,000 - 20,000 = \$5,000$$

and if the stock had gone down to \$85 a share and we sold at this price, our loss would be

$$\text{Loss} = \$100 \times 200 - \$85 \times 200 = 20,000 - 17,000 = \$3,000$$

The second kind of option, which is traded less frequently, is known as *put option* which is an option to sell. When we purchase a put option, we obtain the right to sell a specified number of shares of stock, treasury bonds, or commodity futures at some point within a specified time period at a fixed price, regardless of price changes within the specified time period. This fixed price is also referred to as the exercise price, or strike price. Obviously, a call option would be more likely to result in a profit during a rising (bull) market, whereas a put option would be more likely to result in a profit during a falling (bear) market.

Example F.2

Presently common stock of the ABC Corporation trades at \$55 a share. Our stock broker tells us that we can sell shares of this stock which we do not own, at \$50 a share at any time in the next 100 days, and for this selling option the brokerage house charges us a premium of \$5 per share. Suppose that we sell 200 shares and at some time within the 100–day period this stock goes down to \$40. At this time, we buy 200 shares at \$40 a share and we sell it to a buyer at \$50 a share. Since we sold this stock at \$45 a share (\$50 option price minus \$5 premium), our gain would be

$$\text{Gain} = \$45 \times 200 - \$40 \times 200 = 9,000 - 8,000 = \$1,000$$

and if the stock had gone down up \$60 a share and we bought at this price, our loss would be

$$\text{Loss} = \$60 \times 200 - \$45 \times 200 = 12,000 - 9,000 = \$3,000$$

Examples F.1 and F.2 above, clearly illustrate that even though it is possible to realize handsome gains with small dollar amounts, it is a risky business since we do not have any inside information as to how a corporation will perform in the near future. Again, this material is presented for educational purposes only.

F.2 The Black–Scholes Model Equations

The Black–Scholes model is based on the following equation:

$$C = SN(d_1) - Ke^{-rt}N(d_2) \quad (\text{F.1})$$

where:

C = call premium

S = current stock price

t = time until option expiration

K = option striking price

r = risk-free interest rate – usually represented by the discount rate on U.S. Government Treasury Bills with 30 days left until maturity

N = cumulative standard normal distribution

s = standard deviation of stock returns

$$d_1 = \frac{\ln(S/K) + (r + s^2/2)t}{s\sqrt{t}}$$

$$d_2 = d_1 - s\sqrt{t}$$

In equation (F.1), the factor $N(d_1)$ is a measure of the sensitivity the calculated option value has to small changes in the share price. Thus, the term $SN(d_1)$ represents the expected benefit from acquiring a security outright. The term $Ke^{-rt}N(d_2)$ in (F.1) represents the present value of paying the exercise (striking) price on the expiration day. The difference between these two terms represents the fair market value of the call option.

F.3 Spreadsheet for the Black–Scholes Model

In this section, we will construct a spreadsheet model based on the option–pricing formulas developed by Fischer Black and Myron Scholes. The price volatility, the exercise price, and the expiration date can be obtained from The Wall Street Journal, Barrons, Value Line, or a stock broker. Briefly, volatility is a statistical measure of the dispersion of returns for a given security or market index. Volatility can either be measured by using the standard deviation or variance between returns from that same security or market index. Typically, the higher the volatility, the riskier the security.

The spreadsheet is based on formulas where the terms are multiplied by decimal numbers to the sixth and seventh place. These are obtained from time–tested equations based on the normal distribution curve, and can be found in investment related texts.

The MS Excel spreadsheet is shown in Figure F.1 where the labels in Rows 2, 4, 5, 11, 13, and 20 are entered in Column A and are left–justified. The labels in Rows 6 through 10 and 14 through 17 are also entered in Column A but are right–justified. The width of Column A can be adjusted accordingly; in this model spreadsheet it is set to 40, and the width of Column B is set to 9.

Appendix F The Black–Scholes Stock Options Model

	A	B	C	D	E
2	Options Pricing Using the Black-Scholes Model				
3					0.200000
4	Instrument				0.989060
5	XYZ Corp Common Stock				0.089443
6	Price Volatility	20.00			81.4987
7	Short-term Interest Rate	5.50%			-0.425573
8	Days to Expiration	73			0.364403
9	Exercise (Striking) Price	\$85.00			0.910266
10	Instrument Current Price	\$81.25			0.335210
11	Dividend	\$3.25			0.335210
12					-0.515016
13	Calculated Premiums and Deltas				0.349393
14	Call Premium	\$1.52			0.893416
15	Call Delta	\$0.34			0.303271
16	Put Premium	\$4.99			0.303271
17	Put Delta	\$0.66			
18					
19	Option Type	3			
20	(1 = Cash Bonds 2 = Futures 3 = Securities)				

Figure F.1. Spreadsheet for options pricing using the Black-Scholes model

In cell B6 we enter the volatility, in our case 20, which we have obtained from our stock broker. In B7 we enter the short-term interest rate of 5.5%, also known as a free-risk investment. In B8 we enter the number of days left until the option expires, in our case 73. In B9 we enter the striking price \$85.00, in B10 we enter the current price \$81.25, in B11 we enter the dividend for this common stock, in our case, \$3.25, and in B19 we enter the number 3 for option type 3, i.e., securities.

Column E in the spreadsheet shows the calculated values obtained from the formulas below. The formula in E3 represents the number of days left until the expiration date in fraction of a year. Do not round the decimal numbers in cells E8 through E10.

E3: =B8/365

E4: =EXP(-B7*E3)

E5: =B6*SQRT(E3)/100

E6: =CHOOSE(B19,B10/E4+B11*E3,B10,B10/E4-B11*E3)

E7: =LN(E6/B9)/E5+0.5*E5

E8: =0.3989423*EXP(-0.5*E7^2)

E9: =1/(1+0.2316419*ABS(E7))

E10: =((((1.330274*E9-1.821256)*E9+1.781478)*E9-0.3565638)*E9+0.3193815)*E9*E8

E11: =IF(E7>0,1-E10,E10)

E12: =E7-E5

Copy range E8:E11 to E13. This fills cells E13 through E16.

Enter the formulas below to complete the spreadsheet.

B14: =E4*(E6*E11-B9*E16)

B15: =IF(B19=2,E11*E4,E11)

B16: =B14-E4*(E6-B9)

B17: =IF(B19=2,1-E11*E4,1-E11)

The value in B14 tells us that if we think that the XYZ Corp. stock will go up substantially in the next 73 days, then an option to buy at \$85 at a premium of \$1.52 or less may be a worthwhile investment. But if we think that the price of this stock will drop substantially, an option to sell at \$85 at a premium of \$4.99 (shown in B16) or less may also be a good investment.

The values in B15 and B17 show the call and put deltas, and these indicate the change that is likely to occur in one day if there is a change of \$1 in the current price of the stock. For instance, if we change the days to expiration (cell B8) to 72, and the current price (cell B10) to \$82.25, the call premium will increase to \$1.86, i.e., $\$1.52 + \$0.34 = \$1.86$.

This appendix discusses a method that we can use to predict the probability that a company will go bankrupt. It was introduced by Edward Altman, a financial economist at New York University's Graduate School of Business, and it is now known as The Altman Z-score.

G.1 The Financial Ratios

The Altman Z-score* combines a set of five financial ratios A, B, C, D, and E defined as

1.

$$A = \frac{\text{Earnings Before Interest and Taxes}}{\text{Total Assets}}$$

This ratio is a measure of the return, before interest and taxes, that a firm is earning on its investment in assets. Earnings Before Interest and Taxes (EBIT) is equal to Net Income plus Income Taxes and Interest paid.

2.

$$B = \frac{\text{Net Sales}}{\text{Total Assets}}$$

This ratio is a measure of the sales-generating of the corporation's assets.

3.

$$C = \frac{\text{Market Value of Equity}}{\text{Book Value of Total Liabilities}}$$

In this ratio, the Market Value of Equity is used whenever the common stock price is not available. The Book Value of Total Liabilities is the book value of the current debt.

4.

$$D = \frac{\text{Working Capital}}{\text{Total Assets}}$$

This ratio is a measure of the net liquid assets relative to the total capitalization. Obviously, this ratio will be decreasing for a company with consistent operating losses. The Working Capital is equal to current assets minus current liabilities.

* This Z-score is not the same as the Z-score discussed in Chapter 10, Page 10–65. To avoid confusion, we will refer this score as The Altman Z-score.

Appendix G Forecasting Bankruptcy

5.

$$E = \frac{\text{Retained Earnings}}{\text{Total Assets}}$$

This ratio is a measure of cumulative profitability with respect to total assets. Since new companies, in general, have a greater likelihood of failure, this measure is very critical for new companies.

Once the financial ratios A, B, C, D, and E above are defined, The Altman Z-score is computed as

$$Z = 3.3 \times A + 0.999 \times B + 0.6 \times C + 1.2 \times D + 1.4 \times E$$

G.2 Interpretation of The Altman Z-Score

The likelihood that a company will go bankrupt is based on the following criteria based on financial figures only.

$$Z \geq 3.00 \quad \text{Safe from bankruptcy}$$

$$2.71 \leq Z \leq 2.99 \quad \text{On alert}$$

$$1.81 \leq Z \leq 2.70 \quad \text{Very likely that the company will go bankrupt within 2 years}$$

$$Z \leq 1.80 \quad \text{Destined for bankruptcy}$$

G.3 Spreadsheet for The Altman Z-Score

Figure G.1 is an Excel spreadsheet showing the 10-year financial history and the Altman Z-Scores for each year of the XYZ Corporation. This spreadsheet is constructed as follows:

1. Enter the labels in Rows 1 through 4, and 16 through 20.
2. Enter the years 1997 through 2006 in A5 through A14 in General format
3. Enter the values shown in B5 through D14 in Number format, 0 decimal places
4. In E5 enter =SUM(B5:D5) and copy down to E6 through E14
5. Enter the values shown in F5 through H14 in Number format, 0 decimal places
6. In I5 enter =SUM(F5:H5) and copy down to I6 through I14
7. Enter the values shown in J5 through K14 in Number format, 0 decimal places
8. In L5 enter =SUM(J5:K5) and copy down to L6 through L14
9. Highlight Range B5 through L14, and from the **Insert** drop menu, choose **Name>Create**, and place a check mark on **Top row**.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Z Score History			(All values in millions)									
2	XYZ Corp												
3													
4	Year	CA	P+E	OA	TA	CL	LTL	OL	TL	Stock	RE	NW	TL+NW
5	1997	2353	658	523	3534	1126	938	25	2089	601	843	1444	3533
6	1998	2273	710	592	3575	1042	923	47	2012	615	948	1563	3575
7	1999	2324	771	693	3788	1048	926	80	2054	641	1093	1734	3788
8	2000	2649	890	777	4316	1439	933	69	2441	664	1212	1876	4317
9	2001	3266	1039	942	5247	1873	948	227	3048	694	1505	2199	5247
10	2002	3427	1277	1139	5843	2433	1327	187	3947	872	1024	1896	5843
11	2003	2672	1361	1313	5346	1808	1985	71	3864	882	600	1482	5346
12	2004	1656	966	1077	3699	1135	2015	515	3665	1112	-1078	34	3699
13	2005	1388	837	1138	3363	1367	1289	464	3120	1730	-1487	243	3363
14	2006	1412	716	1121	3249	1257	1238	452	2947	1839	-1537	302	3249

	N	O	P	Q	R	S	T	U	V	W
1										
2										
3										
4	WC	Sales	EBIT	MV	WC/TA	RE/TA	EBIT/TA	MV/TL	Sales/TA	Z-Score
5	1227	5246	289	622	0.35	0.24	0.08	0.30	1.48	2.68
6	1231	5537	434	845	0.34	0.27	0.12	0.42	1.55	2.99
7	1276	6003	494	821	0.34	0.29	0.13	0.40	1.58	3.06
8	1210	6698	466	1001	0.28	0.28	0.11	0.41	1.55	2.88
9	1393	8426	719	1122	0.27	0.29	0.14	0.37	1.61	3.00
10	994	6000	-402	1147	0.17	0.18	-0.07	0.29	1.03	1.42
11	864	7018	-16	376	0.16	0.11	0.00	0.10	1.31	1.71
12	521	4322	-1274	151	0.14	-0.29	-0.34	0.04	1.17	-0.18
13	21	3600	-231	835	0.01	-0.44	-0.07	0.27	1.07	0.39
14	155	4861	120	575	0.05	-0.47	0.04	0.20	1.50	1.13

Figure G.1. A 10–year history of the XYZ Corp financial performance

10. In M5 enter TL+NW, highlight Range M5 through M14, and from the **Insert** drop memo, choose **Name>Create**, and place a check mark on **Top row**.
11. In N5 enter CA–CL, highlight Range N5 through N14, and from the **Insert** drop memo, choose **Name>Create**, and place a check mark on **Top row**.
12. Enter the values shown in P5 through Q14 in Number format, 0 decimal places
13. In R5 enter WC/TA, highlight Range R5 through R14, and from the **Insert** drop memo, choose **Name>Create**, and place a check mark on **Top row**.
14. In S5 enter RE/TA, highlight Range S5 through S14, and from the **Insert** drop memo, choose **Name>Create**, and place a check mark on **Top row**.

Appendix G Forecasting Bankruptcy

14. In T5 enter EBIT/TA, highlight Range T5 through T14, and from the **Insert** drop memo, choose **Name>Create**, and place a check mark on **Top** row.
15. In U5 enter MV/TL, highlight Range U5 through U14, and from the **Insert** drop memo, choose **Name>Create**, and place a check mark on **Top** row.
16. In V5 enter EBIT/TA, highlight Range V5 through V14, and from the **Insert** drop memo, choose **Name>Create**, and place a check mark on **Top** row.
17. In W5 enter

$$=1.2*(WC/TA)+1.4*(RE/TA)+3.3*(EBIT/TA)+0.6*(MV/TL)+0.999*(Sales/TA)$$

then, highlight Range W5 through W14, and from the **Insert** drop memo, choose **Name>Create**, and place a check mark on **Top** row.

18. Copy the formulas down the columns.

Figure G2 shows the 10-year Z-score history for the XYZ Corp. We observe that it bounced around the top part of the On Alert area during the 1997–2001 period, but went downwards during the 2002–2004 period. It shows some improvement at the end of 2006. It could survive if it sells additional shares of stock, disposes some of its assets, or reducing the payroll significantly.

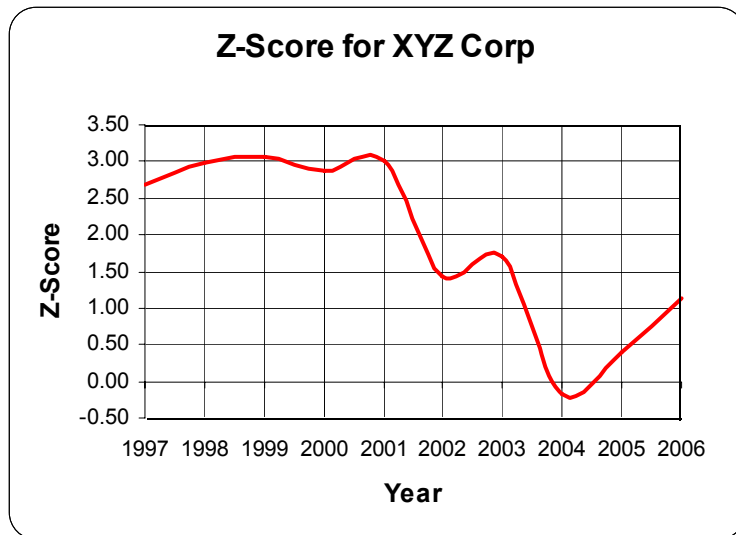


Figure G.2. A graph of the XYZ Corp 10-year Z-score history

The Altman Z-score can be used for credit analysis to accept or reject a loan application, for investment analysis so that investors can determine whether a company is headed for bankruptcy, for auditing analysis to aid accounting firms in assessing a company's financial condition, legal analysis to reveal possible fraud, and merger analysis as an aid to assess whether a prospective merger would result in a firm going bankrupt.

References and Suggestions for Further Study

A. The following publications by The MathWorks, are highly recommended for further study. They are available from The MathWorks, 3 Apple Hill Drive, Natick, MA, 01760, www.mathworks.com.

1. *Getting Started with MATLAB*
2. *Using MATLAB*
3. *Using MATLAB Graphics*
4. *Financial Toolbox*
5. *Statistics Toolbox*

B. Other references indicated in footnotes throughout this text, are listed below.

1. *Modern Investment Theory*, ISBN 0-1301917-0-1
 2. *Numerical Analysis Using MATLAB and Spreadsheets*, ISBN 0-9709511-1-6
 3. *Corporate Financial Distress and Bankruptcy*, ISBN 0-4716918-9-5
 4. *Handbook of Mathematical Functions*, ISBN 0-4866127-2-4
 5. *CRC Standard Mathematical Tables*, ISBN 0-8493-0626-4
-

Index

Symbols

% (percent) symbol in MATLAB B-2

A

a priori probabilities 8-11

abs(z) MATLAB function B-23

abscissa 2-2, A-40

absolute value A-2

absorbing state D-3

acceleration 5-3

actual frequencies 10-76

acute angle - see angle

Add Trendline Excel feature 11-4

adjoint of a matrix - see matrix

algebra defined 1-1

algebraic equation 1-2

alternative hypothesis 10-65

Altman Z-score G-1

amortize MATLAB function 6-67

Analysis of Variance - see ANOVA

analytic geometry 3-1

analytical solution 2-3

angle

acute angle

interior 3-9

obtuse 4-1

angle(z) MATLAB function B-23

angle-difference relations 4-8

angle-sum relations 4-8

annuity 6-5

ordinary 6-5

annurate MATLAB function 6-66

ANOVA

one-way 12-1

two-way 12-8

two-factor without replication

12-8, 12-14

apothem 3-11

arbitrary constant of integration 5-15

area of a polygon 3-2

arithmetic

mean 1-15

progression 1-23

series 1-23

asymptote 3-16

average defined 1-15

AVERAGE Excel function

1-16, 10-16, 12-9

average rate of change 5-1

axis MATLAB command B-16

B

base

10 logarithm A-32

10 number A-1

Bayes' rule 8-11

Bernoulli distribution 10-2

beta distribution C-20

beta function C-17

BETADIST Excel function C-21

between groups

degrees of freedom 12-4, 12-19

sum of squares 12-4, 12-18

variance 12-4, 12-19

variation 12-1

bimodal 1-17

binary number A-1

binomial

coefficients 10-1

distribution 10-2

expansion 10-2

series 10-2

theorem 10-2

bivariate normal distribution 10-56

Black-Scholes pricing model F-1

bond

convertible 6-2

corporate 6-1

coupon 6-3

defined 6-1

junk 6-3

municipal 6-1

perpetual 6-1

zero coupon 6-3

book value 6-3

box MATLAB command B-12

break-even point 2-3

C

calculus 5-1

call option - see option

Cartesian coordinate system 2-2, 4-3

Cartesian form A-40

Cauchy distribution 10-59

center of a circle 3-13

central moment 9-13

characteristic

function 9-18

of a logarithm 1-12

Chart Wizard Excel feature 11-5

Chebyshev's inequality 10-46

CHIDIST Excel function 10-43

CHIINV Excel function 10-43

chi-square

distribution 10-40

test 10-75

test for goodness of fit 10-76

CHITEST Excel function 10-77

circle 3-13

circular cone 3-18

circumference 3-13

clc MATLAB command B-2

clear MATLAB command B-2

coefficient 1-5

cofactor of a matrix - see matrix

column vector in MATLAB B-19

command

screen in MATLAB B-1

window in MATLAB B-1

commas in MATLAB B-8

comment line in MATLAB B-2

common

difference 1-23

logarithm A-32

prefixes A-36

ratio 1-24

comparison of alternate proposals 6-68

complementary error function 10-25

complex

conjugate in MATLAB B-4

fraction A-21

numbers B-3

compound interest - see interest

CONCATENATE Excel function 12-21

conditional probability 8-7

cone 3-18

CONFIDENCE Excel function 10-71

confidence

interval 10-36, 10-71

levels 10-71

limits 10-71

congruent triangles 3-7

constant value A-41

continuous random variable 8-2, 9-1

conv MATLAB function B-6

convertible bond - see bond

co-ordinates A-40

corporate bond - see bond

CORREL Excel function 11-13, 11-15

correlation coefficient 10-56

COUNT Excel function 12-3, 12-4, 12-9

coupon bond - see bond

COVAR Excel function 11-11

covariance 11-10, 11-11

Cramer's rule 2-16, 2-28, 11-3

CRITBINOM Excel function 10-62

cube 3-17

cubic

equation 1-15

root A-31

cumulative

distribution function 9-2

frequency distribution 9-2

frequency percentages 9-3

curve fitting 11-1

curved regression 11-7

cylinder 3-19

D

Data Analysis Toolpack Excel

Add-On 11-7, 12-6

data points in MATLAB B-14

deceleration 5-9

decibel B-13

decimal number A-1

deconv MATLAB function B-6

default color in MATLAB B-15

default line in MATLAB B-15

default marker in MATLAB B-15

default values in MATLAB B-12

definite integral 5-16

degree 4-1

degrees of freedom 10-36, 10-40

between groups 12-4

for error variation 12-12

for Factor A 12-19

- for Factor B 12-19
- for total variation 12-12
- within groups 12-5
- demo MATLAB command B-2
- dependent variable 5-1
- derivative 5-1
- descriptive geometry 3-1
- determinant of a matrix - see matrix
- determinant of order 2 - see matrix
- determinant of order 3 - see matrix
- diagonal elements of a matrix - see matrix
- diagonals in geometry 3-10
- diameter 3-13
- difference A-4
- difference quotient 5-1
- differentiable 5-4
- differential calculus 5-1
- differentials 5-15
- differentiation 5-1
- digital computers A-1
- directrix 3-15, 3-18
- discrete type random variable 8-2, 9-1
- discriminant of a quadratic equation 1-13
- display formats in MATLAB B-31
- dividend 6-39
- dot multiplication operator in MATLAB B-20
- dummy variable 9-10

E

- eccentricity 3-15,
- editor window in MATLAB B-1
- Editor/Debugger in MATLAB B-1
- effrr** MATLAB function 6-62
- element-by-element division in MATLAB B-21
 - division in MATLAB B-21
 - exponentiation in MATLAB B-21
 - multiplication in MATLAB B-20
- elements of the matrix - see matrix
- ellipse 3-14
- eps** MATLAB function B-22, B-27
- equality 1-2
- equation 1-1
- equilateral triangle 3-6
- equity stripping 6-5
- Erlang distribution C-16
- error 11-3, 12-20, 12-21
- error function 10-25
- Euclidean Geometry 3-1
- event 8-2, 9-1
- exercise price F-1, F-2, F-3
- exit** MATLAB command B-2
- EXP** Excel function 10-19
- EXP(GAMMALN(n))** Excel function C-5
 - expected
 - frequencies 10-76
 - value 11-11
- expiration date F-3
- EXPONDIST** Excel function 10-12
- exponent A-22, A-32
- exponential distribution 10-10
- Exponential Moving Average 1-18
- extrapolation 1-20

F

- F critical value 12-7

- F distribution 10-43, 12-1
- F distribution table 12-13
- F ratio 12-1, 12-5, 12-12, 12-21
- F values 12-13
- face value 6-2
- Factor A 12-14
 - degrees of freedom 12-19
 - sum of squares 12-18
 - variance 12-19
- Factor B 12-14
 - degrees of freedom 12-19
 - sum of squares 12-19
 - variance 12-20
- factorial 10-1
- FDIST** Excel function 10-45, 12-6
- figure window in MATLAB B-13
- financial ratios G-1
- FINV** Excel function 10-45, 12-6
- first derivative 5-1
- flipping 6-5
- fmin** MATLAB function B-27
- focus (foci) 3-14
- format MATLAB command B-31
- fplot** MATLAB command B-27
- fractal geometry 3-1
- fractional number A-11
- F-ratio 12-1
- frequency response B-12
- frustum of cone 3-18
- frustum of pyramid 3-18
- F-test 10-74
- FTEST** Excel function 10-75
- function 5-1
- function file in MATLAB B-26
- function-product relations 4-11
- fundamental theorem of
 - integral calculus 5-17
- fundamental trigonometric identities 4-11
- FV** Excel function 6-48
- fvfix** MATLAB function 6-64
- fvvar** MATLAB function 6-65
- fzero** MATLAB function B-26, B-28

G

- gamma distribution 10-36, C-15
- gamma function C-1
- gamma** MATLAB function C-3, C-10
- GAMMADIST** Excel function C-16
- GAMMALN** Excel function C-5
- Gaussian distribution 10-12
- Gaussian Elimination Method 2-17
- generalized factorial function C-1
- generatrix 3-18
- geometric
 - distribution 10-59
 - progression 1-24
 - series 1-24
- geometry 3-1
- golden
 - proportion 1-28
 - rectangle 1-28
- grand
 - mean 12-9
 - sum 12-9
- graphical solution 2-3
- grid** MATLAB command B-12

- gtext** MATLAB command B-13

H

- harmonic
 - progression 1-26
 - series 1-26
- help matlab\i ofun MATLAB command B-2
- help plot MATLAB command B-15
- heteroscedastic 10-74
- homoscedastic 10-74
- hundreds position A-25
- hyperbola 3-15
- hypergeometric distribution 10-53
- HYPGEOMDIST** Excel function 10-55
- hypotenuse 3-3

I

- identity matrix - see matrix
- iir** Excel function 6-54
- imag(z)** MATLAB function B-23
- improper integral C-1
- increments between points in MATLAB B-14
- indefinite integral 5-16
- indenture 6-2
- independent variable 5-1
- indeterminate form 5-4
- infinite sequence 1-23
- inner numbers A-21
- instantaneous
 - acceleration 5-3
 - velocity 5-2
- integer number A-10
- integral calculus 5-1
- integration 5-15
- interaction 12-15
 - degrees of freedom 12-20
 - plot 12-22
 - sum of squares 12-20
 - variance 12-20
- interest 6-6
 - compound 6-8
 - simple 6-6
- interior angle - see angle
- International System of Units A-35
- inverse of a matrix - see matrix
- inverse trigonometric functions 4-13
- IPMT** Excel function 6-58
- irr** MATLAB function 6-61
- isosceles triangle 3-5
- ISPMT** Excel function 6-60

J

- joint
 - occurrence 8-7
 - probability 8-7
 - probability density function 9-12
 - probability distribution function 9-11
- junk bond - see bond

K

- Kelvin's Law 6-71

L

L' Hôpital's rule C-2
Lambda index E-1
law of large numbers 10-46
laws of
 cosines 4-12
 exponents 1-5
 logarithms 1-10
 sines 4-12
 tangents 4-12
least significant digit A-26
least squares
 line 11-2, 11-17
 method 11-2, 11-3
 curve 11-2, 11-17
 parabola 11-2, 11-7
level of significance 10-67, 10-71
limit 5-4
lims = MATLAB function B-27
linear
 correlation coefficient 11-13
 extrapolation 1-20
 factor B-9
 graph A-41
 interpolation 1-20
 regression 11-2
linspace MATLAB command B-14
ln (natural logarithm) B-12
locus 3-14
log (common logarithm) B-12
log(x) MATLAB function B-13
log10(x) MATLAB function B-13
log2(x) MATLAB function B-13
loglog MATLAB command B-12
lower limit of integration 5-18
lower-tail test 10-65

M

$m \times n$ order matrix - see matrix
main diagonal of a matrix - see matrix
major axis 3-14
mantissa 1-12
Markov chain D-2
MATLAB 10-21
MATLAB demos B-2
MATLAB's Editor/Debugger B-1
matrix (matrices)
 adjoint of 2-19
 cofactor of 2-12
 conformable for addition 2-7
 conformable for multiplication 2-8
 conformable for subtraction 2-7
 defined 2-7
 determinant of 2-10
 determinant of **order 2** 2-10
 determinant of **order 3** 2-11
 diagonal elements of 2-7
 elements of 2-7
 identity 2-10
 inverse of 2-20
 $m \times n$ order 2-7, 2-26
 main diagonal of 2-7
 left division in MATLAB 2-25
 multiplication B-18, B-19
 non-singular 2-19
 singular 2-19

 square 2-7
 stochastic D-1
 transition D-2
 zero 2-7
Maxwell distribution 10-61
mean 1-15
mean squares 12-12
mean-square value 9-13
measure of central tendency 1-15
median 1-15, 3-2
mesh(x,y,z) MATLAB command B-17
meshgrid(x,y) MATLAB command B-17
method of least squares 11-2, 11-3
metric system A-35
m-file in MATLAB B-1, B-26
minor axis 3-14
minor of a determinant 2-12
minuend A-4
minutes 4-1
MINVERSE Excel function 2-23
MIRR Excel function 6-56
mixed number A-11
MMULT Excel function 2-23
mode 1-15
moment about the mean 9-13
monotonic function 9-4
most significant digit A-25
moving average 1-18
multinomial distribution 10-52
municipal bond - see bond
mutually exclusive events 8-3

N

NaN in MATLAB B-26, B-27
negative binomial distribution 10-59
negative number A-1
NEGBINOMDIST Excel function 10-60
non-Euclidean geometry 3-1
non-parametric statistical method 11-13
non-singular matrix - see matrix
normal distribution 10-12
NORMSDIST Excel function 10-69, E-2
NORMSINV Excel function 10-70
NPER Excel function 6-51
NPV Excel function 6-52
null hypothesis 10-65
number of levels 12-14

O

obtuse angle - see angle
one-tail
 left test 10-65
 right test 10-65
 test 10-39, 10-65
one-way ANOVA - see ANOVA
option
 call F-1
 put F-2
order of precedence 1-19
ordinary annuity 6-5
ordinate 2-2, A-40
origin date 6-24
outer numbers A-21
overdetermined system 11-3

P

packing 6-5
paired test 10-73
par value 6-2
parabola 3-15
parabolic curve 11-1
parallel lines 3-11
parallelepiped 3-16
parallelogram 3-11
Pascal distribution 10-59
Pearson correlation
 coefficient 11-9, 11-13
PEARSON Excel function 11-13
percentile 10-32
PERCENTILE Excel function 10-35
perimeter 3-2
perpendicular 3-2
perpetual bond - see bond
perpetuity 6-1
phasor A-41
Plot Area in Excel 11-5
plot MATLAB command B-10, B-12, B-15
plot3 MATLAB command B-15
PMT Excel function 6-49
Poisson distribution 10-47
POISSON Excel function 10-51
polar form A-40
polar plot in MATLAB B-24
polar(theta,r) MATLAB function B-24
poly MATLAB function B-4
polyder MATLAB function B-6, B-7
polygon 3-1
polynomial 5-6
 construction from known roots
 using MATLAB B-4
polyval MATLAB function B-5
population 8-2, 10-36
 correlation coefficient 11-12
positive number A-1
posteriori probabilities 8-11
power A-32
PPMT Excel function 6-59
precedence 1-1
present value of an ordinary annuity 1-6
price volatility F-3
principal angle 4-2
prism 3-17
probability 8-1
 differential 9-9
 distribution 9-2
 function 9-2
 integral 10-25
 vector D-1
promissory note 6-4
proportion 1-27
put option - see option
PV Excel function 6-46
P-value 12-8
pvfxx MATLAB function 6-62
pvvar MATLAB function 6-63
pyramid 3-17
Pythagorean theorem 3-4

Q

q function 10-30
quadrant 4-3

quadratic
 curve 11-1
 equation 1-13
 factor B-9
 quadrilateral 3-10
quit MATLAB command 2

R

radian 4-1
 radius 3-13
 random
 experiment 8-2
 variable (rv) 8-2, 9-1
 variation 12-11
 rank correlation coefficient 11-13
RANK Excel function 11-13
 rare event 10-47
RATE Excel function 6-50
 rational polynomial B-8
 Rayleigh distribution 10-57
 real numbers A-2
real(z) MATLAB function B-23
 reciprocal of a number A-11
 rectangle 3-11, 3-12
 rectangular
 coordinate system 4-3
 form A-40
 parallelepiped 3-16
 regression 11-2
 Regression Analysis Excel feature 11-7
 regular stochastic matrix D-5
 relative frequency 8-2
 research hypothesis 10-65
 result 8-2
 rhombus 3-11
 right
 angle 4-1
 circular cylinder 3-19
 triangle 3-3
 rise (in a slope) 2-2
 root(s) 1-13, B-3
roots MATLAB function B-3
ROUND Excel function 10-19
round(n) MATLAB function B-24
 row vector in MATLAB B-3, B-19
 royal flush 8-7
 run (in a slope) 2-2
 rv (random variable) 9-1

S

sample 1-16, 10-36
 point 9-1
 space 9-1
 sampled mean 10-63
 sampling with replacement 10-53
 script file B-26
 second 4-1
 second order differential equation 5-20
 semicolon in MATLAB 8
semilogx MATLAB command B-12
semilogy MATLAB command 12
 set of sample points 9-1
 similar triangles 3-10
 simple
 differential equation 5-15
 harmonic motion 5-20
 interest - see interest
 sinewave A-42
 sinewave graph A-42
 Single Factor Analysis Tool in Excel 12-6
 single sample point 9-1
 singular matrix - see matrix
 sinking fund 6-6
 slope 1-20
 slope 5-1
 slope defined 2-2
 solution by Cramer's rule 2-5
 solution by matrix inversion 2-5
 sphere 3-19
SQRT Excel function 10-19
 square 3-11, 3-12
 square matrix - see matrix
 square root A-30
 standard
 deviation 11-11
 prefixes A-36
STANDARDIZE Excel function 10-15
 standardized form of the normal
 distribution 10-13
 state space D-2
 statistical
 independence 8-9
 regularities 8-2
 STDEV Excel function 10-16
 step function 9-8
 Stirling's asymptotic series C-9
 stochastic
 matrix - see matrix
 process D-1
 strike price F-1, F-2
 string in MATLAB B-16
 Student's t-distribution 10-36
subplot MATLAB command B-18
 substitution method of solution 2-5
 subtrahend A-4
SUM Excel function 12-9
 sum
 identity 1-24
 of squares for error 12-20
 of squares for Factor A 12-18
 of squares for Factor B 12-19

T

tables of integrals 5-16
TDIST Excel function 10-39
 t-distribution 10-36
 tens position A-25
 terminal date 6-24
 test statistic 10-66
text MATLAB command B-14
 thousands position 25
title MATLAB command B-12
 total variation - see variation
 transition diagram D-4
 transition matrix - see matrix
 transversal line 3-9
 trapezoid 3-11, 3-13
 treasury
 bill 6-2
 bond 6-1
 note 6-2
 treatment 12-14
 Trendline Excel feature 11-8
 triangle 3-1
 trigonometric
 identities 4-2
 reduction formulas 4-8
 trigonometry 4-1
 trimodal 1-17
 t-test 10-73
TTEST Excel function 10-74
 two-dimensional normal distribution 10-56
 two-factor with replication 12-8
 Two-Factor with Replication Analysis
 Tool in Excel 12-6
 two-factor without replication 12-8
 Two-Factor without Replication Analysis
 Tool in Excel 12-6
 two-factor without replication ANOVA -
 see ANOVA
 two-tail test 10-39, 10-66
 two-way ANOVA - see ANOVA
 Type
 I error 10-67
 II error 10-67

U

undetermined system 11-3
 uniform distribution 10-6
 units of measurements A-25
 upper limit of integration 5-18
 upper-tail test 10-65

V

variance 11-11
 between groups 12-4
 for Factor B 12-20
 for interaction 12-20
 variation
 between columns 12-11, 12-17
 between rows 12-9
 due to error 12-11
 due to interaction 12-15
 total 12-11
 vector A-41
 velocity 5-2
 vertex 4-1
 volatility F-3

W

Wallis's formulas C-15
 Weibull distribution 10-60
 weighted moving average 1-18
 within groups
 degrees of freedom 12-5
 sum of squares 12-4
 variance 12-5
 variation 12-1

X

x axis A-40
xlabel MATLAB command B-12

Y

y axis A-40
 y-intercept 2-2

ylabel MATLAB command B-12

Z

zero coupon bond - see bond

zero matrix - see matrix

z-test 10-72

ZTEST Excel function 10-73

Mathematics

for Business, Science, and Technology

with MATLAB® and Excel® Computations

Third Edition

Students and working professionals will find Mathematics for Business, Science, and Technology with MATLAB® and Excel® Computations to be a concise and easy-to-learn text. It provides complete, clear, and detailed explanations of the principal mathematical concepts including probability and statistics, and these are illustrated with numerous practical examples.

This text includes the following chapters and appendices:

- Numbers and Arithmetic Operations • Elementary Algebra • Intermediate Algebra
- Fundamentals of Geometry • Fundamentals of Plane Trigonometry • Fundamentals of Calculus • Introduction to Probability and Statistics • Random Variables • Common Probability Distributions and Tests • Curve Fitting, Regression, and Correlation • Analysis of Variance (ANOVA) • Introduction to MATLAB • The Gamma and Beta Functions and their Distributions
- Introduction to Markov Chains • The Lambda Index • The Black-Scholes Stock Options Model
- The Allman Bankruptcy Z-Score

Each chapter contains numerous practical applications supplemented with detailed instructions for using MATLAB and Excel to obtain quick answers.

Steven T. Karris is the president and founder of Orchard Publications, has undergraduate and graduate degrees in electrical engineering, and is a registered professional engineer in California and Florida. He has more than 35 years of professional engineering experience and more than 30 years of teaching experience as an adjunct professor, most recently at UC Berkeley, California.



Orchard Publications

Visit us on the Internet
www.orchardpublications.com or email us:
info@orchardpublications.com

ISBN-13: 978-1-934404-01-0

ISBN-10: 1-934404-01-2



ISBN-13: 978-1-934404-01-0

ISBN-10: 1-934404-01-2

\$69.95 U.S.A.



EX

LIBRIS

Eugene A.

Katkovsky