

DOUGLAS C. MONTGOMERY | GEORGE C. RUNGER

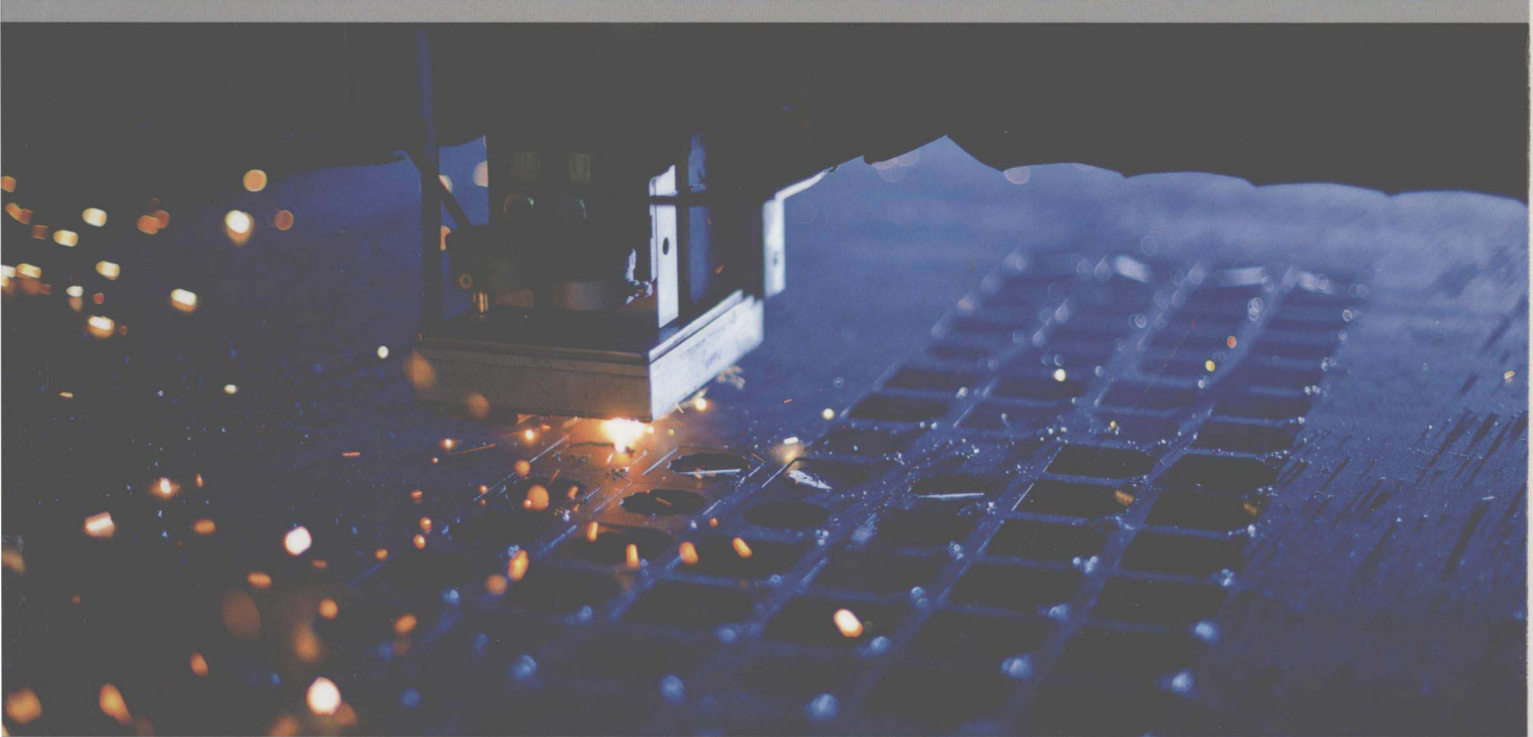
# Applied Statistics and Probability for Engineers

Sixth Edition

EXCLUSIVE TO THIS VERSION

- New exercise sets

- Coverage of SI units



INTERNATIONAL STUDENT VERSION

WILEY

# Applied Statistics and Probability for Engineers

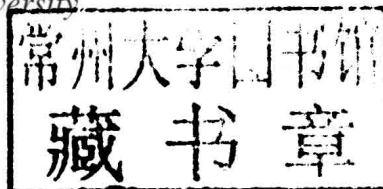
International Student Version

Douglas C. Montgomery

*Arizona State University*

George C. Runger

*Arizona State University*



WILEY

To:  
Meredith, Neil, Colin, and Cheryl  
Rebecca, Elisa, George, and Taylor

Copyright ©2014, 2011 John Wiley & Sons Singapore Pte. Ltd.

Cover image from ©Nikkolia/Shutterstock

Contributing Subject Matter Expert: Mengyun Wei, George Washington University

Founded in 1807, John Wiley & Sons, Inc. has been a valued source of knowledge and understanding for more than 200 years, helping people around the world meet their needs and fulfill their aspirations. Our company is built on a foundation of principles that include responsibility to the communities we serve and where we live and work. In 2008, we launched a Corporate Citizenship Initiative, a global effort to address the environmental, social, economic, and ethical challenges we face in our business. Among the issues we are addressing are carbon impact, paper specifications and procurement, ethical conduct within our business and among our vendors, and community and charitable support. For more information, please visit our website: [www.wiley.com/go/citizenship](http://www.wiley.com/go/citizenship).

All rights reserved. **This book is authorized for sale in Europe, Asia, Africa and the Middle East only and may not be exported outside of these territories.** Exportation from or importation of this book to another region without the Publisher's authorization is illegal and is a violation of the Publisher's rights. The Publisher may take legal action to enforce its rights. The Publisher may recover damages and costs, including but not limited to lost profits and attorney's fees, in the event legal action is required.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, website [www.copyright.com](http://www.copyright.com). Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, website <http://www.wiley.com/go/permissions>.

ISBN: 978-1-118-74412-3

Printed in Asia

10 9 8 7 6 5 4 3 2 1

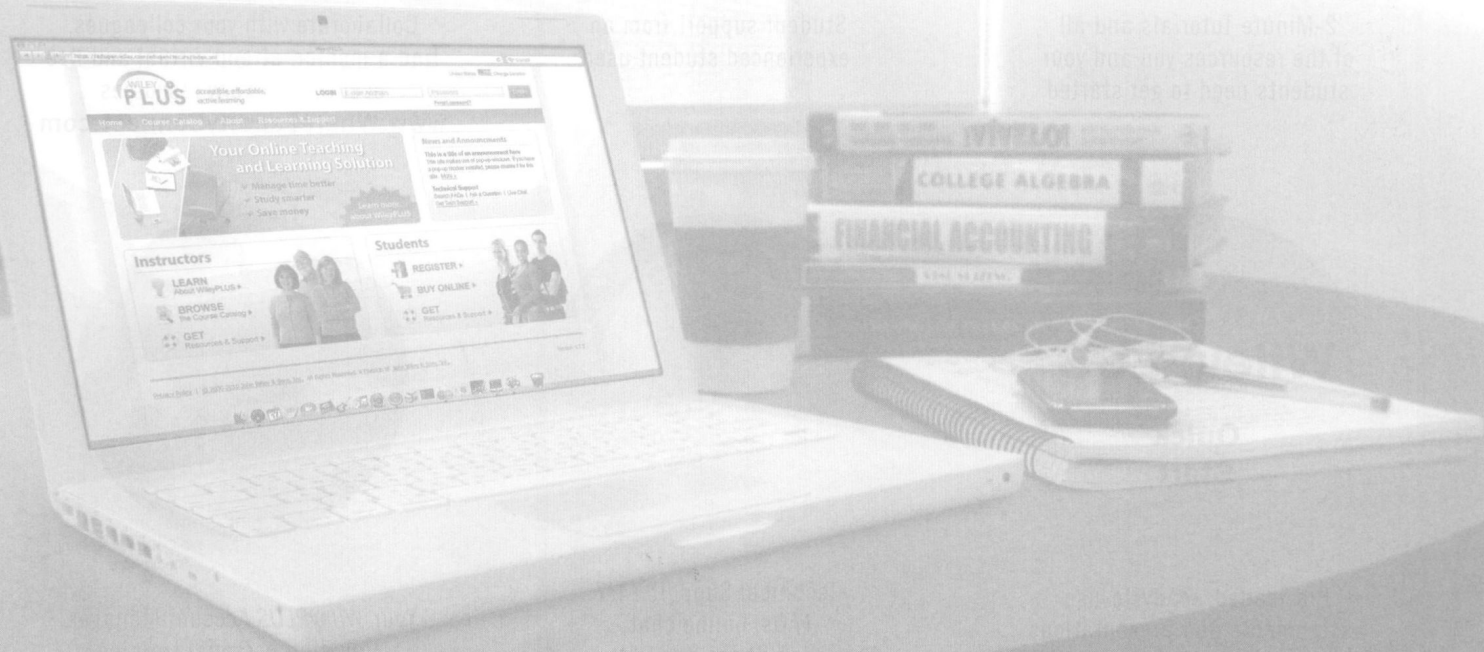
# WileyPLUS

**WileyPLUS is a research-based online environment for effective teaching and learning.**

*WileyPLUS* builds students' confidence because it takes the guesswork out of studying by providing students with a clear roadmap:

- what to do
- how to do it
- if they did it right

It offers interactive resources along with a complete digital textbook that help students learn more. With *WileyPLUS*, students take more initiative so you'll have greater impact on their achievement in the classroom and beyond.



For more information, visit [www.wileyplus.com](http://www.wileyplus.com)

Now available for



Blackboard



# WileyPLUS

**ALL THE HELP, RESOURCES, AND PERSONAL SUPPORT YOU AND YOUR STUDENTS NEED!**

**[www.wileyplus.com/resources](http://www.wileyplus.com/resources)**

**1<sup>st</sup>** DAY OF  
**CLASS**  
**... AND BEYOND!**

2-Minute Tutorials and all of the resources you and your students need to get started

**WileyPLUS**

**Student  
Partner  
Program**

Student support from an experienced student user

**Wiley Faculty Network**



Collaborate with your colleagues, find a mentor, attend virtual and live events, and view resources  
**[www.WhereFacultyConnect.com](http://www.WhereFacultyConnect.com)**

**WileyPLUS**

**Quick  
Start**

Pre-loaded, ready-to-use assignments and presentations created by subject matter experts



Technical Support 24/7  
FAQs, online chat, and phone support  
**[www.wileyplus.com/support](http://www.wileyplus.com/support)**



© Courtney Keating/  
iStockphoto

Your *WileyPLUS* Account Manager, providing personal training and support

# Preface

## INTENDED AUDIENCE

This is an introductory textbook for a first course in applied statistics and probability for undergraduate students in engineering and the physical or chemical sciences. These individuals play a significant role in designing and developing new products and manufacturing systems and processes, and they also improve existing systems. Statistical methods are an important tool in these activities because they provide the engineer with both descriptive and analytical methods for dealing with the variability in observed data. Although many of the methods we present are fundamental to statistical analysis in other disciplines, such as business and management, the life sciences, and the social sciences, we have elected to focus on an engineering-oriented audience. We believe that this approach will best serve students in engineering and the chemical/physical sciences and will allow them to concentrate on the many applications of statistics in these disciplines. We have worked hard to ensure that our examples and exercises are engineering- and science-based, and in almost all cases we have used examples of real data—either taken from a published source or based on our consulting experiences.

We believe that engineers in all disciplines should take at least one course in statistics. Unfortunately, because of other requirements, most engineers will only take one statistics course. This book can be used for a single course, although we have provided enough material for two courses in the hope that more students will see the important applications of statistics in their everyday work and elect a second course. We believe that this book will also serve as a useful reference.

We have retained the relatively modest mathematical level of the first five editions. We have found that engineering students who have completed one or two semesters of calculus and have some knowledge of matrix algebra should have no difficulty reading all of the text. It is our intent to give the reader an understanding of the methodology and how to apply it, not the mathematical theory. We have made many enhancements in this edition, including reorganizing and rewriting major portions of the book and adding a number of new exercises.

## ORGANIZATION OF THE BOOK

Perhaps the most common criticism of engineering statistics texts is that they are too long. Both instructors and students complain that it is impossible to cover all of the topics in the book in one or even two terms. For authors, this is a serious issue because there is great variety in both the content and level of these courses, and the decisions about what material to delete without limiting the value of the text are not easy. Decisions about which topics to include in this edition were made based on a survey of instructors.

Chapter 1 is an introduction to the field of statistics and how engineers use statistical methodology as part of the engineering problem-solving process. This chapter also introduces the reader to some engineering applications of statistics, including building empirical models, designing engineering experiments, and monitoring manufacturing processes. These topics are discussed in more depth in subsequent chapters.

Chapters 2, 3, 4, and 5 cover the basic concepts of probability, discrete and continuous random variables, probability distributions, expected values, joint probability distributions, and independence. We have given a reasonably complete treatment of these topics but have avoided many of the mathematical or more theoretical details.

Chapter 6 begins the treatment of statistical methods with random sampling; data summary and description techniques, including stem-and-leaf plots, histograms, box plots, and probability plotting; and several types of time series plots. Chapter 7 discusses sampling distributions, the central limit theorem, and point estimation of parameters. This chapter also introduces some of the important properties of estimators, the method of maximum likelihood, the method of moments, and Bayesian estimation.

Chapter 8 discusses interval estimation for a single sample. Topics included are confidence intervals for means, variances or standard deviations, proportions, prediction intervals, and tolerance intervals. Chapter 9 discusses hypothesis tests for a single sample. Chapter 10 presents tests and confidence intervals for two samples. This material has been extensively rewritten and reorganized. There is detailed information and examples of methods for determining appropriate sample sizes. We want the student to become familiar with how these techniques are used to solve real-world engineering problems and to get some understanding of the concepts behind them. We give a logical, heuristic development of the procedures rather than a formal, mathematical one. We have also included some material on nonparametric methods in these chapters.

Chapters 11 and 12 present simple and multiple linear regression including model adequacy checking and regression model diagnostics and an introduction to logistic regression. We use matrix algebra throughout the multiple regression material (Chapter 12) because it is the only easy way to understand the concepts presented. Scalar arithmetic presentations of multiple regression are awkward at best, and we have found that undergraduate engineers are exposed to enough matrix algebra to understand the presentation of this material.

Chapters 13 and 14 deal with single- and multifactor experiments, respectively. The notions of randomization, blocking, factorial designs, interactions, graphical data analysis, and fractional factorials are emphasized. Chapter 15 introduces statistical quality control, emphasizing the control chart and the fundamentals of statistical process control.

## WHAT'S NEW IN THIS EDITION

We received much feedback from users of the fifth edition of the book, and in response we have made substantial changes in this new edition.

- Because computer intensive methods are so important in the modern use of statistics, we have added material on the bootstrap and its use in constructing confidence intervals.
- We have increased the emphasis on the use of  $P$ -value in hypothesis testing. Many sections of several chapters were rewritten to reflect this.
- Many sections of the book have been edited and rewritten to improve the explanations and try to make the concepts easier to understand.
- The introductory chapter on hypothesis testing now includes coverage of equivalence testing, a technique widely used in the biopharmaceutical industry, but which has widespread applications in other areas.
- Combining  $P$ -values when performing multiple tests is included.
- Decision theory is briefly introduced in Chapter 15.
- We have added brief comments at the end of examples to emphasize the practical interpretations of the results.
- Many new examples and homework exercises have been added.

## FEATURED IN THIS BOOK

### Definitions, Key Concepts, and Equations

Throughout the text, definitions and key concepts and equations are highlighted by a box to emphasize their importance.

### Learning Objectives

Learning objectives at the start of each chapter guide the students in what they are expected to take away from this chapter and serve as a study reference.

#### Learning Objectives

After careful study of this chapter, you should be able to do the following:

1. Determine probabilities from probability density functions
2. Determine probabilities from cumulative distribution functions and cumulative distribution functions from probability density functions, and the reverse
3. Calculate means and variances for continuous random variables
4. Understand the assumptions for some common continuous probability distributions
5. Select an appropriate continuous probability distribution to calculate probabilities in specific applications
6. Calculate probabilities, determine means and variances for some common continuous probability distributions
7. Standardize normal random variables
8. Use the table for the cumulative distribution function of a standard normal distribution to calculate probabilities
9. Approximate probabilities for some binomial and Poisson distributions

### 4-3 Cumulative Distribution Functions

An alternative method to describe the distribution of a discrete random variable can also be used for continuous random variables.

Cumulative Distribution Function

The cumulative distribution function of a continuous random variable  $X$  is

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u) du \quad (4-3)$$

for  $-\infty < x < \infty$ .

The cumulative distribution function is defined for all real numbers. The following example illustrates the definition.

**Example 4-3** Electric Current For the copper current measurement in Example 4-1, the cumulative distribution function of the random variable  $X$  consists of three expressions. If  $x < 4.9$ ,  $f(x) = 0$ .

Therefore,

$$F(x) = 0, \quad \text{for } x < 4.9$$

and

$$F(x) = \int_{4.9}^x f(u) du = 5x - 24.5, \quad \text{for } 4.9 \leq x < 5.1$$

Finally,

$$F(x) = \int_{-\infty}^x f(u) du = 1, \quad \text{for } 5.1 \leq x$$

Therefore,

$$F(x) = \begin{cases} 0 & x < 4.9 \\ 5x - 24.5 & 4.9 \leq x < 5.1 \\ 1 & 5.1 \leq x \end{cases}$$

The plot of  $F(x)$  is shown in Fig. 4-6.

Notice that in the definition of  $F(x)$ , any  $<$  can be changed to  $\leq$  and vice versa. That is, in Example 4-3  $F(x)$  can be defined as either  $5x - 24.5$  or 0 at the end-point  $x = 4.9$ , and  $F(x)$  can be defined as either  $5x - 24.5$  or 1 at the end-point  $x = 5.1$ . In other words,  $F(x)$  is a continuous function. For a discrete random variable,  $F(x)$  is not a continuous function. Sometimes a continuous random variable is defined as one that has a continuous cumulative distribution function.

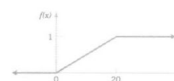


FIGURE 4-6 Cumulative distribution function for Example 4-3.

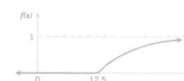


FIGURE 4-7 Cumulative distribution function for Example 4-4.

### Seven-Step Procedure for Hypothesis Testing

The text introduces a sequence of seven steps in applying hypothesis-testing methodology and explicitly exhibits this procedure in examples.

#### 9-1.6 GENERAL PROCEDURE FOR HYPOTHESIS TESTS

This chapter develops hypothesis-testing procedures for many practical problems. Use of the following sequence of steps in applying hypothesis-testing methodology is recommended.

1. **Parameter of interest:** From the problem context, identify the parameter of interest.
2. **Null hypothesis,  $H_0$ :** State the null hypothesis,  $H_0$ .
3. **Alternative hypothesis,  $H_1$ :** Specify an appropriate alternative hypothesis,  $H_1$ .
4. **Test statistic:** Determine an appropriate test statistic.
5. **Reject  $H_0$  if:** State the rejection criteria for the null hypothesis.
6. **Computations:** Compute any necessary sample quantities, substitute these into the equation for the test statistic, and compute that value.
7. **Draw conclusions:** Decide whether or not  $H_0$  should be rejected and report that in the problem context.



## Figures

Numerous figures throughout the text illustrate statistical concepts in multiple formats.

TABLE • 11-1 Oxygen and Hydrocarbon Levels

Observation Number	Hydrocarbon Level $x$ (%)	Purity $y$ (%)
1	0.99	90.01
2	1.02	89.05
3	1.15	91.43
4	1.29	93.74
5	1.46	96.73
6	1.36	94.45
7	0.87	87.59
8	1.23	91.77
9	1.55	99.42
10	1.40	93.65
11	1.19	93.54
12	1.15	92.52
13	0.99	90.56
14	1.02	89.54
15	1.15	89.85
16	1.29	90.39
17	1.46	93.25
18	1.36	93.41
19	0.87	94.98
20	1.23	87.33

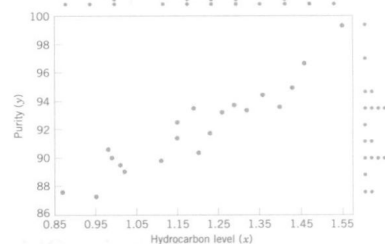
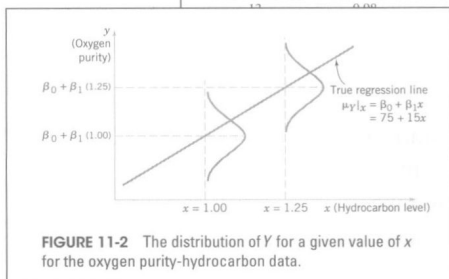


FIGURE 11-1 Scatter diagram of oxygen purity versus hydrocarbon level from Table 11-1.

## Computer Output

Examples throughout the book use computer output to illustrate the role of modern statistical software.

Stem-and-leaf of Strength

$N = 80$	Leaf	Unit = 1.0
1	7	6
2	8	7
3	9	7
5	10	15
8	11	058
11	12	013
17	13	133455
25	14	12356899
37	15	001344678888
(10)	16	0003357789
33	17	0112445668
23	18	0011346
16	19	034699
10	20	0178
6	21	8
5	22	189
2	23	7
1	24	5

FIGURE 6-6 A typical computer-generated stem-and-leaf diagram.

## Example Problems

A set of example problems provides the student with detailed solutions and comments for interesting, real-world situations. Brief practical interpretations have been added in this edition.

### Example 10-1

#### Paint Drying Time

A product developer is interested in reducing the drying time of a primer paint. Two formulations of the paint are tested; formulation 1 is the standard chemistry, and formulation 2 has a new drying ingredient that should reduce the drying time. From experience, it is known that the standard deviation of drying time is 8 minutes, and this inherent variability should be unaffected by the addition of the new ingredient. Ten specimens are painted with formulation 1, and another 10 specimens are painted with formulation 2; the 20 specimens are painted in random order. The two sample average drying times are  $\bar{x}_1 = 121$  minutes and  $\bar{x}_2 = 112$  minutes, respectively. What conclusions can the product developer draw about the effectiveness of the new ingredient, using  $\alpha = 0.05$ ?

We apply the seven-step procedure to this problem as follows:

- 1. Parameter of interest:** The quantity of interest is the difference in mean drying times,  $\mu_1 - \mu_2$ , and  $\Delta_0 = 0$ .
- 2. Non hypothesis:**  $H_0: \mu_1 - \mu_2 = 0$ , or  $H_0: \mu_1 = \mu_2$ .
- 3. Alternative hypothesis:**  $H_1: \mu_1 > \mu_2$ . We want to reject  $H_0$  if the new ingredient reduces mean drying time.
- 4. Test statistic:** The test statistic is

$$z_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \text{where } \sigma_1^2 = \sigma_2^2 = (8)^2 = 64 \text{ and } n_1 = n_2 = 10.$$

- 5. Reject  $H_0$  if:** Reject  $H_0: \mu_1 = \mu_2$  if the  $P$ -value is less than 0.05.
- 6. Computations:** Because  $\bar{x}_1 = 121$  minutes and  $\bar{x}_2 = 112$  minutes, the test statistic is

$$z_0 = \frac{121 - 112}{\sqrt{\frac{(8)^2}{10} + \frac{(8)^2}{10}}} = 2.52$$

- 7. Conclusion:** Because  $z_0 = 2.52$ , the  $P$ -value is  $P = 1 - \Phi(2.52) = 0.0059$ , so we reject  $H_0$  at the  $\alpha = 0.05$  level.

**Practical Interpretation:** We conclude that adding the new ingredient to the paint significantly reduces the drying time. This is a strong conclusion.

## Exercises

Each chapter has an extensive collection of exercises, including **review exercises** that emphasize the material in each section, **supplemental exercises** that cover the scope of chapter topics and require the student to make a decision about the approach they will use to solve the problem, and **mind-expanding exercises** that often require the student to extend the text material somewhat or to apply it in a novel situation. Answers are provided to most odd-numbered exercises in Appendix C in the text, and the *WileyPLUS* online learning environment includes for students complete detailed solutions to selected exercises.

For more information on the exercises, visit the *WileyPLUS* website at [www.wiley.com/college/montgomery](http://www.wiley.com/college/montgomery).

## Important Terms and Concepts

At the end of each chapter is a list of important terms and concepts for an easy self-check and study tool.

## Exercises

### 6-1 Numerical Summaries of Data

6-1. Will the sample mean always correspond to one of the observations in the sample?  
6-2. Will exactly half of the observations in a sample fall below the mean?

6-7. Eight measurements were made on the inside diameter of forged piston rings used in an automobile engine. The data (in millimeters) are 74.002, 74.003, 74.015, 74.000, 74.005, 74.002, 74.005, and 74.005. Calculate the sample mean and

## Supplemental Exercises

5-101. Show that the following function satisfies the properties of a joint probability mass function:

$x$	$y$	$f(x, y)$
0	0	1/4
0	1	1/8
1	0	1/8
1	1	1/4

(e) Conditional mean of the number of people who suffer severe side effects given that 19 suffer minor side effects

5-103. The backoff torque required to remove bolts in a steel plate is rated as high, moderate, or low. Historically, the probability of a high, moderate, or low rating is 0.6, 0.3, or 0.1, respectively. Suppose that 20 bolts are evaluated and that the torque ratings are independent.

## Mind-expanding exercises

5-126. Show that if  $X_1, X_2, \dots, X_p$  are independent, continuous random variables,  $P(X_1 \in A_1, X_2 \in A_2, \dots, X_p \in A_p) = P(X_1 \in A_1)P(X_2 \in A_2) \dots P(X_p \in A_p)$  for any regions  $A_1, A_2, \dots, A_p$  in the range of  $X_1, X_2, \dots, X_p$  respectively.

5-127. Show that if  $X_1, X_2, \dots, X_p$  are independent random variables and  $Y = c_1X_1 + c_2X_2 + \dots + c_pX_p$ ,  $V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \dots + c_p^2V(X_p)$ .

You may assume that the random variables are continuous.

5-128. Suppose that the joint probability function of the continuous random variables  $X$  and  $Y$  is constant on the rectangle  $0 < x < a$ ,  $0 < y < b$ . Show that  $X$  and  $Y$  are independent.

5-129. Suppose that the range of the continuous variables  $X$  and  $Y$  is  $0 < x < a$  and  $0 < y < b$ . Also suppose that the joint probability density function  $f_{XY}(x, y) = g(x)h(y)$ , where  $g(x)$  is a function only of  $x$ , and  $h(y)$  is a function only of  $y$ . Show that  $X$  and  $Y$  are independent.

5-130. This exercise extends the hypergeometric distribution to multiple variables. Consider a population with  $N$  items of  $k$  different types. Assume that there are  $N_1$  items of type 1,  $N_2$  items of type 2, ...,  $N_k$  items of type  $k$  so that  $N_1 + N_2 + \dots + N_k = N$ . Suppose that a random sample of size  $n$  is

selected, without replacement, from the population. Let  $X_1, X_2, \dots, X_k$  denote the number of items of each type in the sample so that  $X_1 + X_2 + \dots + X_k = n$ . Show that for feasible values of  $n, x_1, x_2, \dots, x_k, N_1, N_2, \dots, N_k$ , the probability is

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \frac{\binom{N_1}{x_1} \binom{N_2}{x_2} \dots \binom{N_k}{x_k}}{\binom{N}{n}}$$

5-131. Use the properties of moment generating functions to show that a sum of  $p$  independent normal random variables with means  $\mu_i$  and variances  $\sigma_i^2$  for  $i=1, 2, \dots, p$  has a normal distribution.

5-132. Show that by expanding  $e^X$  in a power series and taking expectations term by term we may write the moment-generating function as

$$M_X(t) = E(e^{tX}) = 1 + \mu'_1 t + \mu'_2 \frac{t^2}{2!} + \dots + \mu'_r \frac{t^r}{r!} + \dots$$

Thus, the coefficient of  $t^r/r!$  in this expansion is  $\mu'_r$ , the  $r$ th origin moment.

Write the power series expansion for  $M_X(t)$  for a gamma random variable and determine  $\mu'_1$  and  $\mu'_2$  using this approach.

## Important Terms and Concepts

Bivariate distribution  
Bivariate normal distribution  
Conditional mean  
Conditional probability density function  
Conditional probability mass function  
Conditional variance

Contour plots  
Correlation  
Covariance  
Error propagation  
General functions of random variables  
Independence

Joint probability density function  
Joint probability distribution  
Joint probability mass function  
Linear functions of random variables

Marginal probability distribution  
Moment-generating functions  
Multinomial distribution  
Reproductive property of the normal distribution

## STUDENT RESOURCES

- **Data Sets** Data sets for all examples and exercises in the text. Visit the student section of the book Web site at [www.wiley.com/college/montgomery](http://www.wiley.com/college/montgomery) to access these materials.

## INSTRUCTOR RESOURCES

The following resources are available only to instructors who adopt the text:

- **Solutions Manual** All solutions to the exercises in the text.
- **Data Sets** Data sets for all examples and exercises in the text.
- **Image Gallery of Text Figures**
- **PowerPoint Lecture Slides**
- **Section on Logistic Regression**

These instructor-only resources are password-protected. Visit the instructor section of the book Web site at [www.wiley.com/college/montgomery](http://www.wiley.com/college/montgomery) to register for a password to access these materials.

## WileyPLUS

This online teaching and learning environment integrates the **entire digital textbook** with the most effective instructor and student resources to fit every learning style.

With *WileyPLUS*:

- Students achieve concept mastery in a rich, structured environment that's available 24/7.
- Instructors personalize and manage their course more effectively with assessment, assignments, grade tracking, and more.

*WileyPLUS* can complement your current textbook or replace the printed text altogether.

## For Students

### Personalize the learning experience

Different learning styles, different levels of proficiency, different levels of preparation—each of your students is unique. *WileyPLUS* empowers them to take advantage of their individual strengths:

- Students receive timely access to resources that address their demonstrated needs, and get immediate feedback and remediation when needed.
- Integrated, multi-media resources—including audio and visual exhibits, demonstration problems, and much more—provide multiple study-paths to fit each student's learning preferences and encourage more active learning.
- *WileyPLUS* includes many opportunities for self-assessment linked to the relevant portions of the text. Students can take control of their own learning and practice until they master the material.

## For Instructors

### Personalize the teaching experience

*WileyPLUS* empowers you with the tools and resources you need to make your teaching even more effective:

- You can customize your classroom presentation with a wealth of resources and functionality from PowerPoint slides to a database of rich visuals. You can even add your own materials to your *WileyPLUS* course.
- With *WileyPLUS* you can identify those students who are falling behind and intervene accordingly, without having to wait for them to come to office hours.
- *WileyPLUS* simplifies and automates such tasks as student performance assessment, making assignments, scoring student work, keeping grades, and more.

## COURSE SYLLABUS SUGGESTIONS

This is a very flexible textbook because instructors' ideas about what should be in a first course on statistics for engineers vary widely, as do the abilities of different groups of students. Therefore, we hesitate to give too much advice, but will explain how we use the book.

We believe that a first course in statistics for engineers should be primarily an applied statistics course, not a probability course. In our one-semester course we cover all of Chapter 1 (in one or two lectures); overview the material on probability, putting most of the emphasis on the normal distribution (six to eight lectures); discuss most of Chapters 6 through 10 on confidence intervals

and tests (twelve to fourteen lectures); introduce regression models in Chapter 11 (four lectures); give an introduction to the design of experiments from Chapters 13 and 14 (six lectures); and present the basic concepts of statistical process control, including the Shewhart control chart from Chapter 15 (four lectures). This leaves about three to four periods for exams and review. Let us emphasize that the purpose of this course is to introduce engineers to how statistics can be used to solve real-world engineering problems, not to weed out the less mathematically gifted students. This course is not the “baby math-stat” course that is all too often given to engineers.

If a second semester is available, it is possible to cover the entire book, including much of the supplemental material, if appropriate for the audience. It would also be possible to assign and work many of the homework problems in class to reinforce the understanding of the concepts. Obviously, multiple regression and more design of experiments would be major topics in a second course.

## USING THE COMPUTER

In practice, engineers use computers to apply statistical methods to solve problems. Therefore, we strongly recommend that the computer be integrated into the class. Throughout the book we have presented typical example of the output that can be obtained with modern statistical software. In teaching, we have used a variety of software packages, including Minitab, Statgraphics, JMP, and Statistica. We did not clutter up the book with operational details of these different packages because how the instructor integrates the software into the class is ultimately more important than which package is used. All text data are available in electronic form on the textbook Web site. In some chapters, there are problems that we feel should be worked using computer software. We have marked these problems with a special icon in the margin.



In our own classrooms, we use the computer in almost every lecture and demonstrate how the technique is implemented in software as soon as it is discussed in the lecture. Student versions of many statistical software packages are available, and students can either purchase their own copy or use the products available through the institution. We have found that this greatly improves the pace of the course and student understanding of the material.

Users should be aware that final answers may differ slightly due to different numerical precision and rounding protocols among softwares.

## ACKNOWLEDGMENTS

We would like to express our grateful appreciation to the many organizations and individuals who have contributed to this book. Many instructors who used the previous editions provided excellent suggestions that we have tried to incorporate in this revision.

We would like to thank the following who assisted in contributing to and/or reviewing material for the *WileyPLUS* course:

Michael DeVasher, *Rose-Hulman Institute of Technology*

Craig Downing, *Rose-Hulman Institute of Technology*

Julie Fortune, *University of Alabama in Huntsville*

Rubin Wei, *Texas A&M University*

We would also like to thank the following for their assistance in checking the accuracy and completeness of the exercises and the solutions to exercises.

Dr. Abdelaziz Berrado

Dr. Connie Borror

Aysegul Demirtas

Kerem Demirtas

Patrick Egbunonu

Sindhura Gangu

James C. Ford

Dr. Alejandro Heredia-Langner

Dr. Jing Hu

Dr. Busaba Laungrungrong

Dr. Fang Li

Dr. Nuttha Lurponglikana

Sarah Streett

Yolande Tra

Dr. Lora Zimmer

We are also indebted to Dr. Smiley Cheng for permission to adapt many of the statistical tables from his excellent book (with Dr. James Fu), *Statistical Tables for Classroom and Exam Room*. Wiley, Prentice Hall, the Institute of Mathematical Statistics, and the editors of *Biometrics* allowed us to use copyrighted material, for which we are grateful.

**Douglas C. Montgomery**

**George C. Runger**

Chapter Opener Photos: © babyblueut/iStockphoto, Erik Isakson/Getty Images, Thinkstock Images/Comstock Images/Getty Images, © amriphoto/iStockphoto, © Ralph125/iStockphoto, © Caro/Alamy, © EasyBuy4u/iStockphoto, © jkullander/iStockphoto, © Alain Nogues/Sygma/Corbis, © HKPNC/iStockphoto, © EdStock/iStockphoto, © skynesher/iStockphoto, © Vasko/iStockphoto, © RedHelga/iStockphoto, © doc-stock/©Corbis, Courtesy of Mike Johnson, [www.redbead.com](http://www.redbead.com)



# Contents

## Chapter 1 The Role of Statistics in Engineering 1

- 1-1 The Engineering Method and Statistical Thinking 2
- 1-2 Collecting Engineering Data 4
  - 1-2.1 Basic Principles 4
  - 1-2.2 Retrospective Study 5
  - 1-2.3 Observational Study 5
  - 1-2.4 Designed Experiments 6
  - 1-2.5 Observing Processes Over Time 8
- 1-3 Mechanistic and Empirical Models 11
- 1-4 Probability and Probability Models 12

## Chapter 2 Probability 15

- 2-1 Sample Spaces and Events 16
  - 2-1.1 Random Experiments 16
  - 2-1.2 Sample Spaces 17
  - 2-1.3 Events 20
  - 2-1.4 Counting Techniques 22
- 2-2 Interpretations and Axioms of Probability 26
- 2-3 Addition Rules 29
- 2-4 Conditional Probability 32
- 2-5 Multiplication and Total Probability Rules 35
- 2-6 Independence 38
- 2-7 Bayes' Theorem 41
- 2-8 Random Variables 43

## Chapter 3 Discrete Random Variables and Probability Distributions 65

- 3-1 Discrete Random Variables 66
- 3-2 Probability Distributions and Probability Mass Functions 67
- 3-3 Cumulative Distribution Functions 68
- 3-4 Mean and Variance of a Discrete Random Variable 71
- 3-5 Discrete Uniform Distribution 74
- 3-6 Binomial Distribution 76
- 3-7 Geometric and Negative Binomial Distributions 80
  - 3-7.1 Geometric Distribution 80
- 3-8 Hypergeometric Distribution 85
- 3-9 Poisson Distribution 88

## Chapter 4 Continuous Random Variables and Probability Distributions 107

- 4-1 Continuous Random Variables 108

- 4-2 Probability Distributions and Probability Density Functions 108
- 4-3 Cumulative Distribution Functions 111
- 4-4 Mean and Variance of a Continuous Random Variable 112
- 4-5 Continuous Uniform Distribution 114
- 4-6 Normal Distribution 116
- 4-7 Normal Approximation to the Binomial and Poisson Distributions 123
- 4-8 Exponential Distribution 126
- 4-9 Erlang and Gamma Distributions 130
- 4-10 Weibull Distribution 133
- 4-11 Lognormal Distribution 134
- 4-12 Beta Distribution 136

## Chapter 5 Joint Probability Distributions 153

- 5-1 Two or More Random Variables 154
  - 5-1.1 Joint Probability Distributions 154
  - 5-1.2 Marginal Probability Distributions 157
  - 5-1.3 Conditional Probability Distributions 159
  - 5-1.4 Independence 163
  - 5-1.5 More Than Two Random Variables 165
- 5-2 Covariance and Correlation 168
- 5-3 Common Joint Distributions 173
  - 5-3.1 Multinomial Probability Distribution 173
  - 5-3.2 Bivariate Normal Distribution 174
- 5-4 Linear Functions of Random Variables 177
- 5-5 General Functions of Random Variables 179
- 5-6 Moment-Generating Functions 181

## Chapter 6 Descriptive Statistics 197

- 6-1 Numerical Summaries of Data 198
- 6-2 Stem-and-Leaf Diagrams 203
- 6-3 Frequency Distributions and Histograms 206
- 6-4 Box Plots 209
- 6-5 Time Sequence Plots 211
- 6-6 Scatter Diagrams 212
- 6-7 Probability Plots 215

## Chapter 7 Point Estimation of Parameters and Sampling Distributions 235

- 7-1 Point Estimation 236

- 7-2 Sampling Distributions and the Central Limit Theorem 237
- 7-3 General Concepts of Point Estimation 244
  - 7-3.1 Unbiased Estimators 244
  - 7-3.2 Variance of a Point Estimator 245
  - 7-3.3 Standard Error: Reporting a Point Estimate 246
  - 7-3.4 Bootstrap Standard Error 247
  - 7-3.5 Mean Squared Error of an Estimator 248
- 7-4 Methods of Point Estimation 249
  - 7-4.1 Method of Moments 249
  - 7-4.2 Method of Maximum Likelihood 251
  - 7-4.3 Bayesian Estimation of Parameters 257

## Chapter 8 Statistical Intervals for a Single Sample 267

- 8-1 Confidence Interval on the Mean of a Normal Distribution, Variance Known 269
  - 8-1.1 Development of the Confidence Interval and Its Basic Properties 269
  - 8-1.2 Choice of Sample Size 272
  - 8-1.3 One-Sided Confidence Bounds 273
  - 8-1.4 General Method to Derive a Confidence Interval 273
  - 8-1.5 Large-Sample Confidence Interval for  $\mu$  275
- 8-2 Confidence Interval on the Mean of a Normal Distribution, Variance Unknown 277
  - 8-2.1  $t$  Distribution 277
  - 8-2.2  $t$  Confidence Interval on  $\mu$  278
- 8-3 Confidence Interval on the Variance and Standard Deviation of a Normal Distribution 280
- 8-4 Large-Sample Confidence Interval for a Population Proportion 282
- 8-5 Guidelines for Constructing Confidence Intervals 286
- 8-6 Bootstrap Confidence Interval 287
- 8-7 Tolerance and Prediction Intervals 288
  - 8-7.1 Prediction Interval for a Future Observation 288
  - 8-7.2 Tolerance Interval for a Normal Distribution 289

## Chapter 9 Tests of Hypotheses for a Single Sample 301

- 9-1 Hypothesis Testing 302
  - 9-1.1 statistical hypotheses 302
  - 9-1.2 Tests of Statistical Hypotheses 304

- 9-1.3 One-Sided and Two-Sided Hypotheses 309
- 9-1.4  $P$ -Values in Hypothesis Tests 310
- 9-1.5 Connection Between Hypothesis Tests and Confidence Intervals 312
- 9-1.6 General Procedure for Hypothesis Tests 314

### 9-2 Tests on the Mean of a Normal Distribution, Variance Known 316

- 9-2.1 Hypothesis Tests on the Mean 316
- 9-2.2 Type II Error and Choice of Sample Size 319
- 9-2.3 Large-Sample Test 323

### 9-3 Tests on the Mean of a Normal Distribution, Variance Unknown 323

- 9-3.1 Hypothesis Tests on the Mean 323
- 9-3.2 Type II Error and Choice of Sample Size 328

### 9-4 Tests on the Variance and Standard Deviation of a Normal Distribution 330

- 9-4.1 Hypothesis Tests on the Variance 330
- 9-4.2 Type II Error and Choice of Sample Size 332

### 9-5 Tests on a Population Proportion 332

- 9-5.1 Large-Sample Tests on a Proportion 333
- 9-5.2 Type II Error and Choice of Sample Size 335

### 9-6 Summary Table of Inference Procedures for a Single Sample 337

### 9-7 Testing for Goodness of Fit 337

### 9-8 Contingency Table Tests 340

### 9-9 Nonparametric Procedures 342

- 9-9.1 The Sign Test 343
- 9-9.2 The Wilcoxon Signed-Rank Test 347
- 9-9.3 Comparison to the  $t$ -Test 349

### 9-10 Equivalence Testing 349

### 9-11 Combining $P$ -Values 351

## Chapter 10 Statistical Inference for Two Samples 369

### 10-1 Inference on the Difference in Means of Two Normal Distributions, Variances Known 370

- 10-1.1 Hypothesis Tests on the Difference in Means, Variances Known 372
- 10-1.2 Type II Error and Choice of Sample Size 373
- 10-1.3 Confidence Interval on the Difference in Means, Variances Known 375

### 10-2 Inference on the Difference in Means of Two Normal Distributions, Variances Unknown 377

- 10-2.1 Hypotheses Tests on the Difference in Means, Variances Unknown 377
- 10-2.2 Type II Error and Choice of Sample Size 384
- 10-2.3 Confidence Interval on the Difference in Means, Variances Unknown 384
- 10-3 A Nonparametric Test for the Difference in Two Means 386
  - 10-3.1 Description of the Wilcoxon Rank-Sum Test 386
  - 10-3.2 Large-Sample Approximation 388
  - 10-3.3 Comparison to the  $t$ -Test 388
- 10-4 Paired  $t$ -Test 389
- 10-5 Inference on the Variances of Two Normal Distributions 393
  - 10-5.1  $F$  Distribution 393
  - 10-5.2 Hypothesis Tests on the Ratio of Two Variances 395
  - 10-5.3 Type II Error and Choice of Sample Size 397
  - 10-5.4 Confidence Interval on the Ratio of Two Variances 397
- 10-6 Inference on Two Population Proportions 398
  - 10-6.1 Large-Sample Tests on the Difference in Population Proportions 399
  - 10-6.2 Type II Error and Choice of Sample Size 400
  - 10-6.3 Confidence Interval on the Difference in Population Proportions 402
- 10-7 Summary Table and Road Map for Inference Procedures for Two Samples 403

## Chapter 11 Simple Linear Regression and Correlation 419

- 11-1 Empirical Models 420
- 11-2 Simple Linear Regression 423
- 11-3 Properties of the Least Squares Estimators 427
- 11-4 Hypothesis Tests in Simple Linear Regression 428
  - 11-4.1 Use of  $t$ -Tests 428
  - 11-4.2 Analysis of Variance Approach to Test Significance of Regression 430
- 11-5 Confidence Intervals 432
  - 11-5.1 Confidence Intervals on the Slope and Intercept 432
  - 11-5.2 Confidence Interval on the Mean Response 432
- 11-6 Prediction of New Observations 434
- 11-7 Adequacy of the Regression Model 436

- 11-7.1 Residual Analysis 436
- 11-7.2 Coefficient of Determination ( $R^2$ ) 438
- 11-8 Correlation 439
- 11-9 Regression on Transformed Variables 444
- 11-10 Logistic Regression 446

## Chapter 12 Multiple Linear Regression 467

- 12-1 Multiple Linear Regression Model 468
  - 12-1.1 Introduction 468
  - 12-1.2 Least Squares Estimation of the Parameters 471
  - 12-1.3 Matrix Approach to Multiple Linear Regression 473
  - 12-1.4 Properties of the Least Squares Estimators 478
- 12-2 Hypothesis Tests In Multiple Linear Regression 479
  - 12-2.1 Test for Significance of Regression 479
  - 12-2.2 Tests on Individual Regression Coefficients and Subsets of Coefficients 482
- 12-3 Confidence Intervals In Multiple Linear Regression 486
  - 12-3.1 Confidence Intervals on Individual Regression Coefficients 486
  - 12-3.2 Confidence Interval on the Mean Response 487
- 12-4 Prediction of New Observations 488
- 12-5 Model Adequacy Checking 489
  - 12-5.1 Residual Analysis 489
  - 12-5.2 Influential Observations 492
- 12-6 Aspects of Multiple Regression Modeling 494
  - 12-6.1 Polynomial Regression Models 494
  - 12-6.2 Categorical Regressors and Indicator Variables 497
  - 12-6.3 Selection of Variables and Model Building 499
  - 12-6.4 Multicollinearity 506

## Chapter 13 Design and Analysis of Single-Factor Experiments: The Analysis of Variance 529

- 13-1 Designing Engineering Experiments 530
- 13-2 Completely Randomized Single-Factor Experiment 531
  - 13-2.1 Example: Tensile Strength 531
  - 13-2.2 Analysis of Variance 532
  - 13-2.3 Multiple Comparisons Following the ANOVA 539

- 13-2.4 Residual Analysis and Model Checking 541
- 13-2.5 Determining Sample Size 542
- 13-3 The Random-Effects Model 544
  - 13-3.1 Fixed Versus Random Factors 544
  - 13-3.2 ANOVA and Variance Components 545
- 13-4 Randomized Complete Block Design 547
  - 13-4.1 Design and Statistical Analysis 547
  - 13-4.2 Multiple Comparisons 552
  - 13-4.3 Residual Analysis and Model Checking 553

## Chapter 14 Design of Experiments with Several Factors 565

- 14-1 Introduction 566
- 14-2 Factorial Experiments 568
- 14-3 Two-Factor Factorial Experiments 572
  - 14-3.1 Statistical Analysis of the Fixed-Effects Model 572
  - 14-3.2 Model Adequacy Checking 577
  - 14-3.3 One Observation per Cell 578
- 14-4 General Factorial Experiments 579
- 14-5  $2^k$  Factorial Designs 581
  - 14-5.1  $2^2$  Design 581
  - 14-5.2  $2^k$  Design for  $k \geq 3$  Factors 587
  - 14-5.3 Single Replicate of the  $2^k$  Design 594
  - 14-5.4 Addition of Center Points to a  $2^k$  Design 598
- 14-6 Blocking and Confounding in the  $2^k$  Design 600
- 14-7 Fractional Replication of the  $2^k$  Design 605
  - 14-7.1 One-Half Fraction of the  $2^k$  Design 605
  - 14-7.2 Smaller Fractions: The  $2^{k-p}$  Fractional Factorial 611
- 14-8 Response Surface Methods and Designs 618

## Chapter 15 Statistical Quality Control 647

- 15-1 Quality Improvement and Statistics 648
  - 15-1.1 Statistical Quality Control 649
  - 15-1.2 Statistical Process Control 650
- 15-2 Introduction to Control Charts 650
  - 15-2.1 Basic Principles 650
  - 15-2.2 Design of a Control Chart 654
  - 15-2.3 Rational Subgroups 655
  - 15-2.4 Analysis of Patterns on Control Charts 656
- 15-3  $\bar{X}$  and  $R$  or  $S$  Control Charts 658
- 15-4 Control Charts for Individual Measurements 665

- 15-5 Process Capability 668
- 15-6 Attribute Control Charts 672
  - 15-6.1  $P$  Chart (Control Chart for Proportions) 672
  - 15-6.2  $U$  Chart (Control Chart for Defects per Unit) 675
- 15-7 Control Chart Performance 677
- 15-8 Time-Weighted Charts 678
  - 15-8.1 Cumulative Sum Control Chart 678
  - 15-8.2 Exponentially Weighted Moving-Average Control Chart 684
- 15-9 Other SPC Problem-Solving Tools 688
- 15-10 Decision Theory 689
  - 15-10.1 Decision Models 689
  - 15-10.2 Decision Criteria 691
- 15-11 Implementing SPC 693

## Appendix A. Statistical Tables and Charts 713

- Table I Summary of Common Probability Distributions 714
- Table II Cumulative Binomial Probabilities  $P(X \leq x)$  715
- Table III Cumulative Standard Normal Distribution 718
- Table IV Percentage Points  $\chi^2_{\alpha, v}$  of the Chi-Squared Distribution 720
- Table V Percentage Points  $t_{\alpha, v}$  of the  $t$  Distribution 721
- Table VI Percentage Points  $f_{\alpha, v_1, v_2}$  of the  $F$  Distribution 722
- Chart VII Operating Characteristic Curves 727
- Table VIII Critical Values for the Sign Test 736
- Table IX Critical Values for the Wilcoxon Signed-Rank Test 736
- Table X Critical Values for the Wilcoxon Rank-Sum Test 737
- Table XI Factors for Constructing Variables Control Charts 738
- Table XII Factors for Tolerance Intervals 739

## Appendix B: Bibliography 741

Appendix C: Answers to Selected Exercises (this section is available online)

## Glossary 745

## Index 761

## Index of Applications in Examples and Exercises