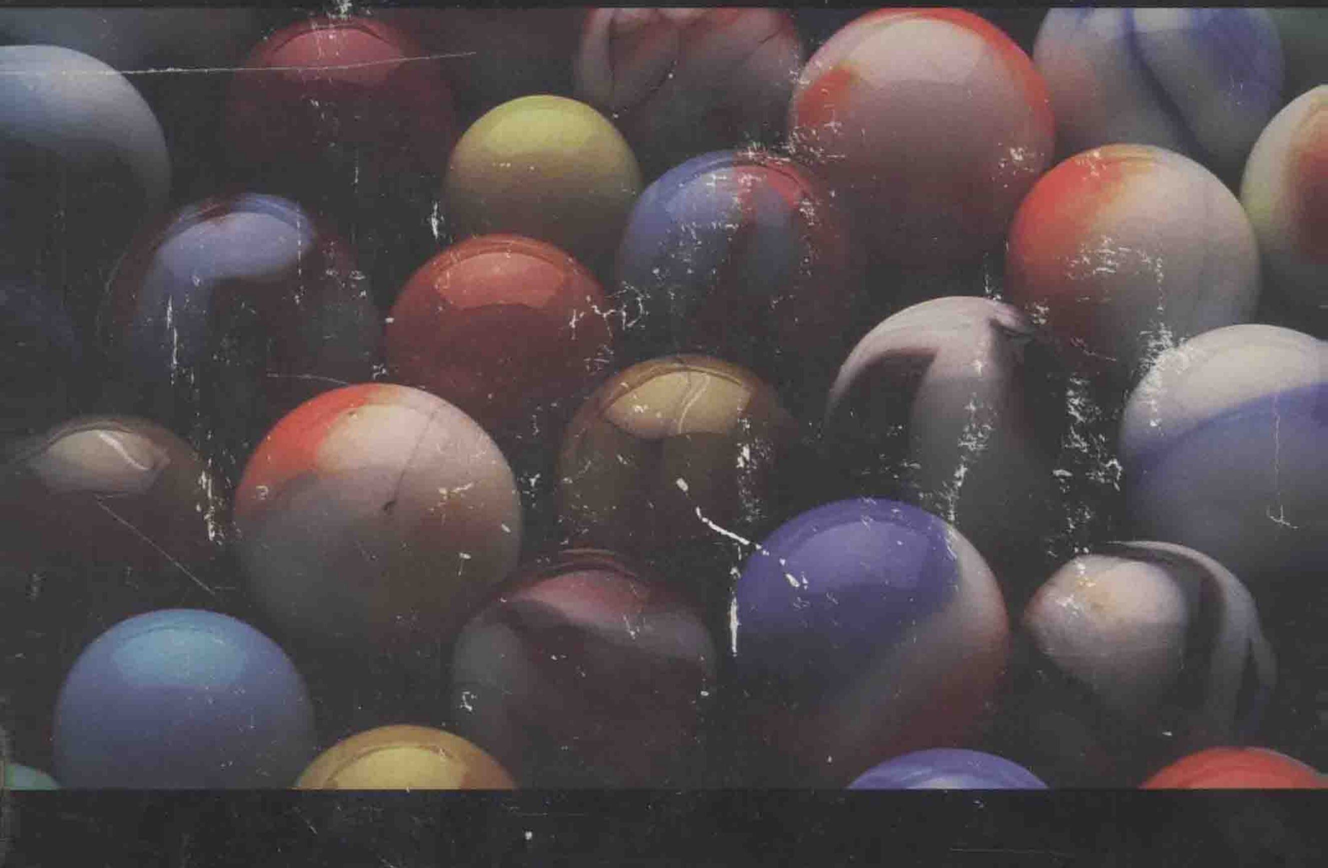
JUST THE ESSENTIALS

of

Elementary Statistics

SECOND EDITION



ROBERT JOHNSON

PATRICIA KUBY

JUST THE ESSENTIALS OF ELEMENTARY STATISTICS

SECOND EDITION

ROBERT JOHNSON
PATRICIA KUBY
Monroe Community College

Duxb
An Imprint of Brooks/Cole Publishing Company
I(T)P® An International Thomson Publishing Company

Pacific Grove, CA • Albany, NY • Belmont, CA • Boston • Cincinnati • Johannesburg • London • Madrid
Melbourne • Mexico City • New York • Scottsdale, AZ • Singapore • Tokyo • Toronto

Sponsoring Editor: *Carolyn Crockett*
Marketing Team: *Christine Davis/Michele Mootz*
Editorial Assistant: *Kimberly Raburn*
Production Editor: *Tessa A. McGlasson*
Production: *Greg Hubit Bookworks*
Permissions: *The Permissions Group*
Interior Designer: *Stuart Paterson*
Cover Designer: *Roy R. Neuhaus*
Cover Photo: *Roger Allyn Lee/ SuperStock, Inc.*
Typesetting: *Maryland Composition Corporation*
Printing and Binding: *World Color/ Taunton*

For more information, contact:

BROOKS/COLE PUBLISHING COMPANY
511 Forest Lodge Road
Pacific Grove, CA 93950
USA

International Thomson Publishing Europe
Berkshire House 168-173
High Holborn
London, WC1V7AA
England

Thomas Nelson Australia
102 Dodds Street
South Melbourne, 3205
Victoria, Australia

Nelson Canada
1120 Birchmount Road
Scarborough, Ontario
Canada M1K 5G4

International Thomson Editores
Seneca 53, Col. Polanco
México, D. F., México
C. P. 11560

International Thomson Publishing GmbH
Königswinterer Strasse 418
53227 Bonn
Germany

International Thomson Publishing Asia
60 Albert Street
#15-01 Albert Complex
Singapore 189969

International Thomason Publishing, Japan
Hirakawacho Kyowa Building, 3F
2-2-1 Hirakawacho
Chiyoda-ku, Tokyo 102
Japan

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Johnson, Robert Russell
Just the essentials of elementary statistics. — 2nd ed. / Robert Johnson, Patricia Kuby.
p. cm.
Includes indexes.
ISBN 0-534-35779-2
1. Statistics. I. Kuby, Patricia. II. Title
QA276.12.J65 1999
519.5—dc21 98-22598

PREFACE

PURPOSE AND PREREQUISITES

This book was written for use as an introductory course for students who need a working knowledge of statistics but do not have a strong mathematical background. Statistics requires the use of many formulas and an occasional solution of a simple algebraic equation. Those students who have not completed intermediate algebra should complete at least one semester of college mathematics as a prerequisite before attempting this course.

OUR OBJECTIVES

The primary objective of *Just the Essentials, Second Edition,* is to present a truly concise introduction to statistics that can be covered in its entirety in a one-term course. We have aimed to write a very readable textbook that will promote learning, understanding, and motivation by presenting statistics in a context that relates to personal experiences. Simply, our goal is a brief, clear, and interesting statistics textbook.

Statistics is a practical discipline that evolves with the changing needs of our society. Today's student is the product of a particular cultural environment and is motivated differently from students of a few years ago. In this text we present statistics as a useful tool in learning about the world around us. While studying descriptive and inferential concepts, students will become aware of their real-world applications in such fields as the physical and social sciences, business, economics, and engineering.

IMPORTANT ONGOING FEATURES

This second edition continues to feature the following elements:

- A communication style that reflects current student culture;
- A strong computer flavor, with numerous annotated Minitab outputs and exercises that can be used as computer laboratory assignments;
- Many exercises that are intended to be solved with the aid of a computer, and the corresponding instrucitons, if Minitab is to be used;
- A Chapter 1 that introduces ideas of variability and data collection, as well as basic terms;

- Case Studies based on situations of interest and using real data;
- An early descriptive presentation of linear correlation and linear regression in Chapter 3.

THIS REVISION

We hope the users of this second edition will appreciate the following improvements:

- A presentation that is more approachable, clearer, and more visual throughout;
- The presentation is *more approachable* because it is often less technical, such as in the new discussion on hypothesis testing in Chapter 8;
- It is *clearer* because formulas have been expressed in "word algebra," paraphrasing algebraic symbolism in recognition of the frequent discomfort with notation. This is done any time a formula of any complexity is introduced, such as in Chapters 2 and 3 with the many descriptive statistics formulas (Formula 2-1) and with the inequality notation with class intervals (Table 2-4);
- It is *more visual* as well. There are new graphic displays that compare similar statistics (Figure 2-28), that organize procedures for finding percentiles (Figure 2-26), and that organize and compare the various statistical tests (Figures 9-1, 9-7, and 10-1). There are simulations that visually demonstrate—for example, level of confidence, alpha and *p*-value (Illustrations 8-3, 8-14, and 8-20);
- A rewritten Chapter 8 provides a more natural flow of ideas from sampling distribution, to estimation, then to hypothesis testing, using either the *p*-value or classical approach. It also contains more drill exercises, broken down into smaller steps;
- The exercises have been expanded and improved, from simple drills to critical thinking questions. The simplest exercises help in acquiring confidence in computations and single-step questions. Some appear in the margins of the text to be done as the text material is read. The end-of-section and end-of-chapter exercises have been expanded to include many real-life data situations and critical thinking questions all aimed at getting the student involved in real-world statistical applications;
- More real data are used and all significant data sets are included on a data disk;
- There is an increased number of exercises that demonstrate statistical theory through computer simulation. Several of these could be used as special assignments or as lab exercises. Karl Pearson tossed a coin 24,000 times, observing heads/tails. It must have taken days for him to have realized the power of the law of large numbers; Exercise 9.63 repeats this experiment in seconds—even repeats it several times in a few minutes. There are simulation exercises designed to help in understanding the importance of the assumptions to a statistical test (Exercise 9.41);
- Two specially designed tables determine *p*-values so that the initial experience is not overwhelming. Tables 5 and 7 in Appendix B allow *p*-values to be read directly for tests involving the *z*- or the *t*-statistic;
- The focus on interpreting computer output has been increased;

- The MINITAB, Version 11, instructions are presented with both session commands and menu commands being displayed;
- The Pareto diagram, a popular graph in industry, is presented in Chapter 2;
- New physical page layout makes use of the margins with the addition of margin exercises and margin "conversation."

To the Instructor: The Text as a Teaching Tool

One primary objective of this book is to cover only that which is truly essential in elementary statistics and can be taught in a one-term course. This is ultimately a subjective judgment, and it reflects our understanding of what most instructors teach in one term. We would, nevertheless, like to indicate the criteria that guided us in deciding on the coverage. They are as follows:

The first three chapters are introductory by nature: Chapter 1 is an introduction to the language of statistics; Chapter 2 covers the descriptive presentation of single-variable data; and Chapter 3 is the descriptive presentation of bivariate data. The bivariate material is presented at this point in the book because students often ask about the relationship between two sets of data (such as heights and weights) while studying Chapter 2.

In the chapters on probability (4 and 5), the concepts of permutations and combinations are deliberately avoided. Instead, this material is contained in Appendix A, "Basic Principles of Counting," so that it may be included as the instructor wishes. The binomial coefficient is introduced in connection with the binomial probability distribution in Chapter 5.

The instructor has several options in the selection of topics to be studied in a given course. We consider Chapters 1 through 9 to be the basic core of a course (some sections of Chapters 2, 4, and 6, and all of Chapter 3 may be omitted without affecting continuity). Following the completion of Chapter 9, any combination of sections from Chapters 3, 10, and 11 may be studied.

The *Getting Started* section provides a brief explanation of the components of each chapter and makes suggestions about ways to use this book more effectively. It is suggested reading for both the instructor and the student.

To the Student: The Text as a Learning Tool

Statistics is different from other courses:

1. It has its own extensive technical vocabulary.
2. It is highly cumulative, in that many of the concepts you will be learning at each step become the basis for other concepts learned throughout the rest of the course; therefore, failure to master each concept as presented can cause great difficulty later on.
3. It requires very precise measurements and calculations (a seemingly minor error will often be magnified and lead to wrong answers in some procedures).
4. While it is an academic subject, statistics is also very real and touches each of us frequently in everyday life. Plain talk and stress on common sense are the book's main characteristics as a learning tool. This approach should allow

you—provided you have the necessary basic mathematics skills—to work your way through the course with relative ease. Examples of this approach are: (a) Illustration 1-1 (p. 11), which is used to reemphasize the meaning of the eight basic definitions presented in Section 1.2, and (2) Chapter Objectives for Chapter 2 (p. 28), which use a familiar situation to motivate the topics of Chapter 2.

Our goal in writing this textbook is to motivate and involve you in the statistics that you are learning. Turn to page xvii and read *Getting Started* for a brief explanation of some of the ways you can use this book more effectively to succeed in learning statistics.

SUPPLEMENTS

Statsource is a multimedia supplement bound in the back of each book. It provides students with videos, a tutorial program, a lecture presentation that corresponds to the text, concept simulations using Java applets, and data sets—formatted for Minitab, Excel, StataQuest, JMP IN, and in ASCII—to solve the problems in the text.

The Statistical Tutor is a student manual that:

a. Contains the complete solutions to the odd-numbered exercises (the same exercises whose answers are in the back of the book).
b. Contains many helpful hints and suggestions to serve as a guide through the learning process. It includes many summaries and overviews.
c. Contains several review lessons to help refresh materials studied previously in other courses.

The **Instructor's Manual** is also intended to be uncommonly helpful. It contains:

a. Everything that is in *The Statistical Tutor.*
b. The complete solutions to all even-numbered exercises.
c. Many helpful teaching suggestions that an instructor might incorporate. The notes specifically intended for the instructor are set in a type different from that for the student material.

The **Test Bank** contains a combination of true-false, multiple-choice, short-answer, matching, and computational test questions for each chapter in the text.

The **MINITAB Student Supplement** is provided for those interested in teaching or learning the course interactively with the computer. This supplement is a text-specific introduction to the MINITAB statistical analysis system and is keyed to text discussion and examples.

ACKNOWLEDGMENTS

We owe a debt to many other books. Many of the ideas, principles, examples, and developments that appear in this text stem from thoughts provoked by these sources.

It is a pleasure to acknowledge the aid and the encouragement we have received throughout the development of this text from students and colleagues at Monroe Community College. A special thanks to Madeline A. Bradley, University of North Carolina, Greensboro; Donald A. Chambless, Auburn University at Montgomery; Charles

W. Haynes, University of Texas at Austin; Jiann-Shiun Huang, Columbia College; Debra L. Hydorn, Mary Washington College; Patricia P. O'Connor, University of North Carolina, Greensboro; Grayson S. Sallez, University of North Carolina, Greensboro; Furman H. Smith, Auburn University at Montgomery; Walker G. Weigel, University of North Carolina, Greensboro.

Thanks also to the many authors and publishers who so generously extended reproduction permissions for the news articles and tables used in the text. These acknowledgments are specified individually throughout the text.

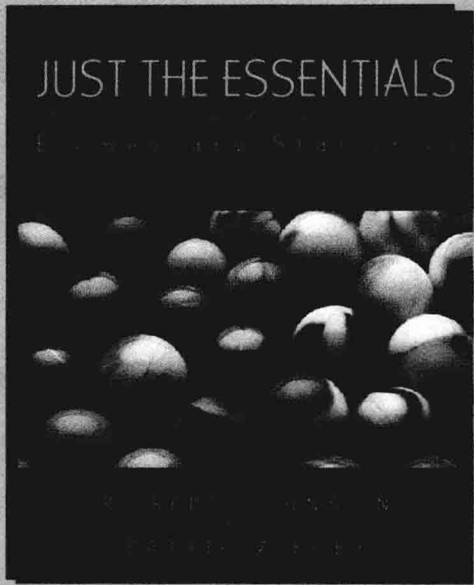
And last and certainly the most significant of all—thank you to our spouses, Barbara and Joe, for their assistance and their just "being there."

Robert Johnson
Patricia Kuby

Getting Started

Read this section to become familiar with the components of each chapter and their intended purpose.

Your Guide to Getting the Most Out of Just the Essentials, Second Edition



First realize that statistics is more than the mathematics of formulas and data. Statistics includes the processes of problem solving, statistical thinking, data collection, obtaining numerical and graphical results, and the follow-up questioning of those results. Sometimes statistics requires the use of mathematics, and sometimes it does not. Everybody (and everything) is an individual, and is uniquely different from all the others. However, when it comes to a single trait, many values of a single variable taken from many individuals will generally form a pattern. Statistical methodologies are used to describe and help explain these patterns. Statistics uses mathematical techniques to quantify the ideas being investigated and to reduce the information to a numeric format, in which it can be treated graphically or algebraically. When concepts become quantified, they become an application of mathematics, not *the* mathematics.

As important as it is to be able to take a set of data and calculate statistical values (mean, correlation coefficient, line of best fit, etc.) or to draw a graphic display (histograms, scatter diagrams, etc.), it is far more important that you understand the circumstances being investigated, that you understand the variables involved, that you understand why you are investigating the problem, and that you must learn to question the data and the statistical results. Your life experience and understanding of real-life situations is the foundation for understanding statistics. Don't lose sight of the fact that statistics is about describing the world around us. You will see many statistical examples from business, the physical and social sciences, and many other fields and professions as you study from this textbook.

To get the most out of this book, become familiar with its many learning features. On the following pages are examples of several features this book contains and suggestions on how to make the best use of them. Take a moment to look them over and, then, please use them. Active involvement in learning about statistics is the single most important factor in determining success and satisfaction. Now open up your mind and let your imagination and your natural curiosity go to work.

GETTING STARTED

A person cannot just get in their car and start driving and expect to arrive at the correct destination. They must know where they are going and what route to take before they start. Studying statistics is much the same. Every chapter in this textbook opens with two important study tools.

CHAPTER OUTLINES appear at the beginning of each chapter to give a schematic overview of what is to be presented. These outlines are annotated to give a first impression of a key concept from each section of the chapter.

9

INFERENCES INVOLVING ONE POPULATION

CHAPTER OUTLINE

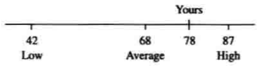
9.1 Inferences About Mean μ (σ unknown)
When the standard deviation of the population is unknown, the Student's t-distribution is used to make inferences.

9.2 Inferences About the Binomial Probability of Success
The observed sample proportion p' is approximately normally distributed under certain conditions.

CHAPTER OBJECTIVES prepare the way for new material by describing why the material is important and how it relates to previously studied topics. Use Chapter Objectives to understand the motivation for learning this material.

● CHAPTER OBJECTIVES

Imagine that you took an exam at the last meeting of your favorite class. Today your instructor returns your exam paper and it has a grade of 78. If you are like most students in this situation, as soon as you see your grade you want to know how your grade compares to those of the rest of the class and you immediately ask: "What was the average exam grade?" Your instructor replies, "The class average was 68." Since 78 is 10 points above the average, you ask: "How close to the top is my grade?" Your instructor replies that the grades ranged from 42 to 87 points. The accompanying figure summarizes the information we have so far.



A third question that is sometimes asked is "How are the grades distributed?" Your instructor replies that half the class had grades between 65 and 75. With this information you conclude that your grade is fairly good.

The preceding illustration demonstrates the basic process of statistics and how students make use of statistics on a regular basis: There is a set of data (set of exam scores), the data is described with a few descriptive statistics (average grade and high and low grades), and based on this information students are able to draw conclusions about their relative success.

A large part of this chapter is devoted to learning how to present and describe sets of data. There are four basic types of descriptive statistics: (1) measures of central tendency, (2) measures of dispersion (spread), (3) measures of position, and (4) types of distribution. The measures of central tendency are used to describe or locate the "middle" of the data values. The measures of dispersion measure the spread or variability within the data values. When very large sets of data are involved, the measures of position become useful. For example, when college board exams are taken, thousands of scores result. Your concern is the position of your score with respect to all the others. Averages, ranges, and so on, are not enough. A measure of position will tell you that your score is better than a certain percentage of the other scores. For example, your score is better than 85% of all the scores. The fourth concept, the type of distribution, describes the data by telling you whether the values are evenly distributed or clustered (bunched) around a certain value. Another very important part of this chapter will explain how to interpret the findings so that we know what the data is telling us about the sampled population.

In this chapter, we deal with single-variable data, which is data collected for one numerical variable at a time. In Chapter 3 we will learn to work with two variables at a time and learn about their relationship.

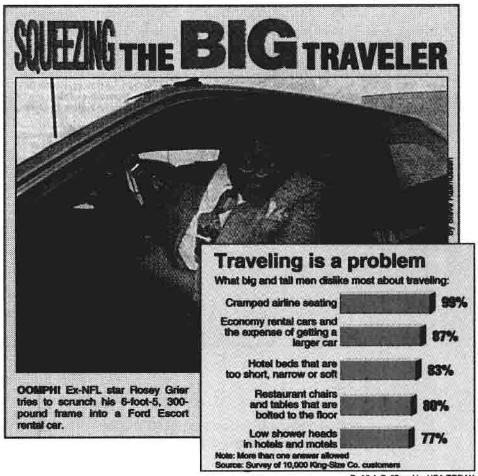
MAKING STATISTICS COME ALIVE

Case Study 1-1

Measuring Physical Discomfort

Sitting in an uncomfortable seat for long periods of time is no fun. Slipping into a seat on a jet airliner when you are larger than average can be outright painful.

By Del Jones, *USA Today*, 5-27-94



OOMPH! Ex-NFL star Rosey Grier tries to scrunch his 6-foot-5, 300-pound frame into a Ford Escort rental car.

By Nick Galifianakis, USA TODAY

EXERCISE 1.1

Refer to "Traveling is a problem."

a. Who was surveyed?
b. How many were surveyed?
c. Explain the meaning of "Cramped airline seating: 99%."
d. Why are so many percentages reported?

The world is small for big and tall business travelers. Ask Rosey Grier, a 6-foot-5, 300-pound former NFL defensive lineman, who never met a running back he couldn't tackle. But he can't win against airline seats built for 5-foot-9, 170-pounders.

Grier's not alone. 13 million other men are at least 6-foot-2 or 225 pounds.

Airline seats have never been roomy. "They were originally designed to fit a very well-known athlete: (jockey) Willie Shoemaker," jokes Ed Perkins, editor of Consumer Reports Travel Letter, which measures the distance between seats ev-

Case Studies are designed to demonstrate how statistics work in the everyday real world. Throughout the text, case studies can be found that incorporate statistical concepts as these concepts are presented. Margin exercises also accompany these case studies and provide an excellent way to try out your new knowledge in a real-world setting.

FIGURE 2-26

Finding P_k Procedure

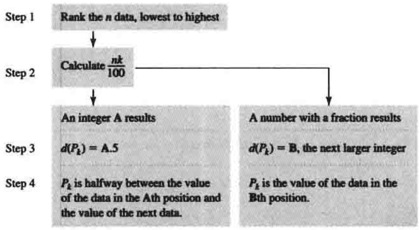
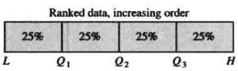


FIGURE 2-23

Quartiles



The procedure for determining the value of the quartiles is the same as that for percentiles and is shown in the following description of percentiles.

Graphical displays can be found in the form of *charts, graphs,* and *tables*. Graphics are very important in statistics. They are the pictures that either demonstrate the theory or condense vast amounts of data into an easy-to-understand format.

UNDERSTANDING STATISTICS

All the important **definitions** and **key terms** are highlighted. The term also appears in the margin to assist in locating referenced information quickly.

Ungrouped freq. dist.

Grouped freq. dist.

The frequency f is the number of times the value x occurs in the sample. Table 2-3 is an **ungrouped frequency distribution**—"ungrouped" because each value of x in the distribution stands alone. When a large set of data has many different x-values instead of a few repeated values, as in the previous example, we can group the values into a set of classes and construct a **grouped frequency distribution.** The stem-and-leaf display in Figure 2-5 (p. 34) shows, in picture form, a grouped frequency distribution. Each stem represents a class. The number of leaves on each stem is the same as the frequency for that same class. The data represented in Figure 2-5 are listed as a frequency distribution in Table 2-4.

Completely **worked-out examples** present the step-by-step solution process.

▼ ILLUSTRATION 2-14

Find the standard scores for (a) 92 and (b) 72 with respect to a sample of exam grades that have a mean score of 75.9 and a standard deviation of 11.1.

SOLUTION

a. $x = 92, \bar{x} = 75.9; s = 11.1$. Thus, $z = \frac{x - \bar{x}}{s} = \frac{92 - 75.9}{11.1} = \frac{16.1}{11.1} = \mathbf{1.45}$.

b. $x = 72, \bar{x} = 75.9; s = 11.1$. Thus, $z = \frac{x - \bar{x}}{s} = \frac{72 - 75.9}{11.1} = \frac{-3.9}{11.1} = \mathbf{-0.35}$

▲ This means that the score 92 is approximately one-and-one-half standard deviations above the mean, while the score 72 is approximately one-third of a standard deviation below the mean.

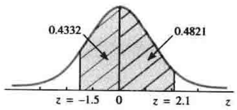
Some exercises are located in the book's margin; these **margin exercises** have been placed there for an initial practice on the adjacent concept. Solving these exercises before reading further is an excellent way to begin homework assignments.

▼ ILLUSTRATION 6-6

EXERCISE 6.6

Find the area between $z = -1.83$ and $z = 1.23$, $P(-1.83 < z < 1.23)$.

The area between $z = -1.5$ and $z = 2.1$, $P(-1.5 < z < 2.1)$, is found by adding the two areas together. Both probabilities are read directly from Table 3.



Therefore, we obtain

$P(-1.5 < z < 2.1) = P(-1.5 < z < 0) + P(0 < z < 2.1) = 0.4332 + 0.4821 = \mathbf{0.9153}$

SECTION 9.1 Inferences About Mean μ (σ unknown) 391

9.29 State the null hypothesis, H_0, and the alternative hypothesis, H_a, that would be used to test each of the following claims.
   a. The mean weight of honey bees is at least 11 g.
   b. The mean age of patients at Memorial Hospital is no more than 54 years.
   c. The mean amount of salt in granola snack bars is different from 75 mg.

9.30 Determine the *p*-value for the following hypothesis tests involving the Student's *t*-distribution with 10 degrees of freedom.
   a. H_0: $\mu = 15.5$, H_a: $\mu < 15.5$, $t^* = -2.01$ b. H_0: $\mu = 15.5$, H_a: $\mu > 15.5$, $t^* = 2.01$
   c. H_0: $\mu = 15.5$, H_a: $\mu \neq 15.5$, $t^* = 2.01$ d. H_0: $\mu = 15.5$, H_a: $\mu \neq 15.5$, $t^* = -2.01$

9.31 Determine the test criteria that would be used to test the null hypotheses below.
   a. H_0: $\mu = 10$ ($\alpha = 0.05, n = 15$) H_a: $\mu \neq 10$ b. H_0: $\mu = 37.2$ ($\alpha = 0.01, n = 25$) H_a: $\mu > 37.2$
   c. H_0: $\mu = -20.5$ ($\alpha = 0.05, n = 18$) H_a: $\mu < -20.5$ d. H_0: $\mu = 32.0$ ($\alpha = 0.01, n = 42$) H_a: $\mu > 32.0$

9.32 Compare the *p*-value and classical approaches to hypothesis testing by comparing the *p*-value and decision of the *p*-value approach to the critical values and decision of the classical approach for each of the following situations. Use $\alpha = 0.05$.
   a. H_0: $\mu = 128$, H_a: $\mu \neq 128$, $n = 15$, $t^* = 1.60$
   b. H_0: $\mu = 18$, H_a: $\mu > 18$, $n = 25$, $t^* = 2.16$
   c. H_0: $\mu = 38$, H_a: $\mu < 38$, $n = 45$, $t^* = -1.73$
   d. Compare the results of the two techniques for each case.

9.33 A student group maintains that the average student must travel for at least 25 minutes in order to reach college each day. The college admissions office obtained a random sample of 22 one-way travel times from students. The sample had a mean of 19.4 min and a standard deviation of 9.6 min. Does the admissions office have sufficient evidence to reject the students' claim? Use $\alpha = 0.01$.
   a. Solve using the *p*-value approach.
   b. Solve using the classical approach.

9.34 Homes in a nearby college town have a mean value of \$88,950. It is assumed that homes in the vicinity of the college have a higher value. To test this theory, a random sample of 12 homes is chosen from the college area. Their mean valuation is \$92,460 and the standard deviation is \$5,200. Complete a hypothesis test using $\alpha = 0.05$. Assume prices are normally distributed.
   a. Solve using the *p*-value approach.
   b. Solve using the classical approach.

End-of-section exercises give the opportunity to practice concepts that are presented in the section. The best way to learn statistics is to practice, and once a concept is thought to be understood, practice it some more for good measure. This text contains hundreds of exercises with many applications. The *icons* next to some of the exercises indicate the type of application—business, sports, criminal justice, and so on.

Figure skaters practice their skating for many hours in order to make the Olympics. Homework time is the statistics students' practice time. Effort and persistence are the hallmark of every gold medalist and all successful statistics students.

CHAPTER EXERCISES

9.83 One of the objectives of a large medical study was to estimate the mean physician fee for cataract removal. For 25 randomly selected cases the mean fee was found to be \$1550 with a standard deviation of \$125. Set a 99% confidence interval on μ, the mean fee for all physicians. Assume fees are normally distributed.

9.84 A natural-gas utility is considering a contract for purchasing tires for its fleet of service trucks. The decision will be based on expected mileage. For a sample of 100 tires tested, the mean mileage was 36,000 and the standard deviation was 2000 miles. Estimate the mean mileage that the utility should expect from these tires using a 96% confidence interval.

9.85 According to an article in *Changing Times* (March 1991), financial assets in the United States fell by \$6700 per household between the end of 1989 and the end of 1990. Suppose that in order to check this statement, a survey of 500 households is conducted and that the mean decrease in financial assets was \$6300 with a standard deviation equal to \$3250. Calculate the value of the test statistic and the *p*-value for the research hypothesis $\mu < \$6700$.

9.86 It has been suggested that abnormal male children tend to occur more in children born to older-than-average parents. Case histories of 20 abnormal males were obtained, and the ages of the 20 mothers were

31	21	29	28	34	45	21	41	27	31
43	21	39	38	32	28	37	28	16	39

Chapter Exercises are based on all the techniques studied in the chapter. They offer the opportunity to integrate conceptual and computational skills as well as to identify the appropriate procedure needed to produce desired results.

Chapter Exercises 409

IN RETROSPECT

We have been studying inferences, both confidence intervals and hypothesis tests, for the two basic population parameters (mean μ and proportion p) of a single population. Most inferences about a single population are concerned with one of these two parameters. Figure 9-7 on page 410 presents a visual organization of the techniques studied throughout Chapters 8 and 9 along with the key questions that you must ask yourself as you are deciding which test statistic and which formula to use.

In this chapter we also used the maximum error of estimate, formula (9-7), to determine the size of sample required to make estimates about the population proportion with the desired accuracy. Case Study 9-2 presents a graph showing the "margin of error" (maximum error of estimate) for sample sizes up to 1500. These are for 95% confidence interval estimates as you found out in Exercise 9.50. By combining a point estimate with its corresponding maximum error of estimate (sampling error), we can construct an interval estimate based on the information reported. Most polls and surveys use the 95% confidence level, even though they do not report the 95%.

In the next chapter we will discuss inferences about two populations whose respective means and proportions are to be compared.

An **In Retrospect** section summarizes the concepts learned in the chapter, pointing out the relationships and the interrelationships to material covered previously.

Doing Statistics on the Computer

MINITAB is a statistics computer software package that is easy to use and lets the reader focus on learning concepts while it handles the calculations and graphing of data. In the real world, most statistical calculations are now done by computers.

MINITAB (Release 11) commands to construct a Pareto diagram with the categories listed in C1 and their corresponding counts or frequencies listed in C2.

Session commands	*Menu commands*
Enter:	Choose: Stat > SPC > Pareto Chart
%PARETO C1;	Select: Chart defect table
COUNts C2;	Enter: Label in C1
TITLe 'your title'.	Frequency in C2
	Your Title

MINITAB commands are introduced in the text, and using them as much as possible is encouraged. Several exercises list the Minitab needed to complete that exercise; however Minitab is not the only tool you can use. The exercise may be completed by using equivalent commands with another software or without the aid of a computer.

2.12 The USA Snapshot® "How to say I love you" reports the results of a David Michaelson & Associates survey for Ethel M Chocolates, on the best way to show affection.

If you use MINITAB, use %PARETO

Best way to show affection	Give gift	Hold hands	Hugging/kissing	Smiling	Other
Percent who said	10%	10%	51%	20%	9%

Draw a Pareto diagram picturing this information.

2.13 The January 10, 1991, USA Snapshot® "What's in U.S. landfills" reports the percentages of each type of waste in our landfills: food—4%, glass—2%, metal—14%, paper—38%, plastic—18%, yard waste—11%, other—13%.
a. Construct a Pareto diagram displaying this information.
b. Because of the size of the "other" category, the Pareto diagram may not be the best graph to use. Explain why, and describe what additional information is needed to make the Pareto diagram more appropriate.

2.14 The final-inspection defect report for assembly-line A12 is reported on a Pareto diagram.

Some exercises are also designed to encourage taking advantage of the "friendly" power of the computer. The "Helpful Hint" on page 74 tells you how to incorporate Cut and Paste techniques to make the work even easier.

406 CHAPTER 9 INFERENCES INVOLVING ONE POPULATION

9.63 Karl Pearson once tossed a coin 24,000 times and recorded 12,012 heads.
a. Calculate the point estimate for $p = P(\text{head})$ based on Pearson's results.
b. Determine the standard error of proportion.
c. Determine the 95% confidence interval estimate for $p = P(\text{head})$.
d. It must have taken Mr. Pearson many hours to toss a coin 24,000 times. You can simulate 24,000 coin tosses using the MINITAB commands listed below.

(*Note:* A Bernoulli experiment is like a "single" trial binomial experiment. That is, one toss of a coin is one Bernoulli experiment with $p = 0.5$; and 24,000 tosses of a coin either is a binomial experiment with $n = 24{,}000$ or is 24,000 Bernoulli experiments. Code: 0 = tail, 1 = head. The sum of the 1's will be the number of heads in the 24,000 tosses.)
e. How do your simulated results compare to Pearson's?
f. Use these six commands and generate another set of 24,000 coin tosses. Compare these results to those above. Explain what you can conclude from these results.

```
RANDOM 24000 C1;
  BERNoulli 0.5.
SUM C1 K1
LET K2 = K1/24000
PRINt K2
```

DOING STATISTICS ON YOUR OWN

WORKING WITH YOUR OWN DATA

The central limit theorem is very important to the development of the rest of this course. Its proof, which requires the use of calculus, is beyond the intended level of this course. However, the truth of the CLT can be demonstrated both theoretically and by experimentation. The following series of questions will help to verify the central limit theorem both ways:

A | THE POPULATION

Consider the theoretical population that contains the three numbers 0, 3, and 6 in equal proportions.

1. a. Construct the theoretical probability distribution for the drawing of a single number, with replacement, from this population.
   b. Draw a histogram of this probability distribution.
   c. Calculate the mean μ and the standard deviation σ for this population.

A **Working with Your Own Data** section is included at the end of each part, encouraging further exploration. These sections provide a personalized learning experience beginning with the collection of data and continuing through the application techniques that have been studied. Experience shows that more concepts are retained by applying methods just learned in regular homework assignments to data that are familiar and understandable.

STUDYING AND PREPARING FOR EXAMS

VOCABULARY LIST

Be able to define each term. Pay special attention to the key terms, which are printed in red. In addition, describe in your own words, and give an example of, each term. Your examples should not be the ones given in class or in the textbook.

The bracketed numbers indicate the chapter in which the term first appeared, but you should define the terms again to show increased understanding of their meaning. Page numbers indicate the first appearance of the term in Chapter 5.

binomial coefficient (p. 229)
binomial experiment (p. 227)
binomial probability function (p. 229)
binomial random variable (p. 227)
constant function (p. 214)
continuous random variable (p. 211)
discrete random variable (p. 211)
experiment [1, 4] (p. 227)
failure (p. 227)
independent trials (p. 227)
mean of probability distribution (p. 220)
mutually exclusive events [4] (p. 213)
population parameter [1] (p. 220)
probability distribution (p. 213)
probability function (p. 214)
probability histogram (p. 216)
random variable (p. 210)
sample statistic [1] (p. 219)
standard deviation of probability distribution (p. 220)
success (p. 227)
trial (p. 227)
variance of probability distribution (p. 220)

The first step in preparing for an exam should always be: Complete all the reading and exercises assigned.

A **Vocabulary list** near the end of every chapter is an aid in deciding how much of the material is truly understood. Consider this informal self-test: try defining all of the vocabulary list to a friend. This will determine if more practice is needed.

The **Chapter Practice Test** provides a formal self-evaluation of the mastery of the material before being tested in class. Correct responses are in the back of the text. The most effective way to use a practice test is to: (1) complete the material assigned for the pending test; (2) use the vocabulary list as an informal self-test; (3) then on the day before the exam, take the practice test, *under test conditions;* (4) correct it using the answers in the back; and (5) restudy the concepts related to the items missed. Do not use the practice tests as homework; use them to test yourself and they will help eliminate those "silly errors." WARNING: Use of the Practice Test too early in the study process defeats its role in the learning process.

Chapter Practice Test

(1–5) Answer "True" if the statement is always true. If the statement is not always true, replace the words shown in bold with the words that make the statement always true.

4.1 The probability of an event is a **whole number**.

4.2 The sample points of a sample space are **equally likely** events.

4.3 The probabilities of complementary events always **are equal**.

4.4 If two events are mutually exclusive, they are also **independent**.

4.5 If the sets of sample points belonging to two different events do not intersect, the events are **independent**.

4.6 A computer is programmed to generate the eight single-digit integers 1, 2, 3, 4, 5, 6, 7, and 8 with equal frequency. Consider the experiment—"the next integer generated." Define:

Event A = "Odd number" = {1, 3, 5, 7}

Event B = "Number more than 4" = {5, 6, 7, 8}

Event C = "1 or 2" = {1, 2}

Find:

a. $P(A)$	b. $P(B)$	c. $P(C)$	d. $P(\overline{C})$
e. $P(A \text{ and } B)$	f. $P(A \text{ or } B)$	g. $P(B \text{ and } C)$	h. $P(B \text{ or } C)$
i. $P(A \text{ and } C)$	j. $P(A \text{ or } C)$	k. $P(A \mid B)$	l. $P(B \mid C)$
m. $P(A \mid C)$			

n. Are events A and B mutually exclusive? Explain.
o. Are events B and C mutually exclusive? Explain.
p. Are events A and C mutually exclusive? Explain.
q. Are events A and B independent? Explain.
r. Are events B and C independent? Explain.
s. Are events A and C independent? Explain.

A Recommended Study Procedure

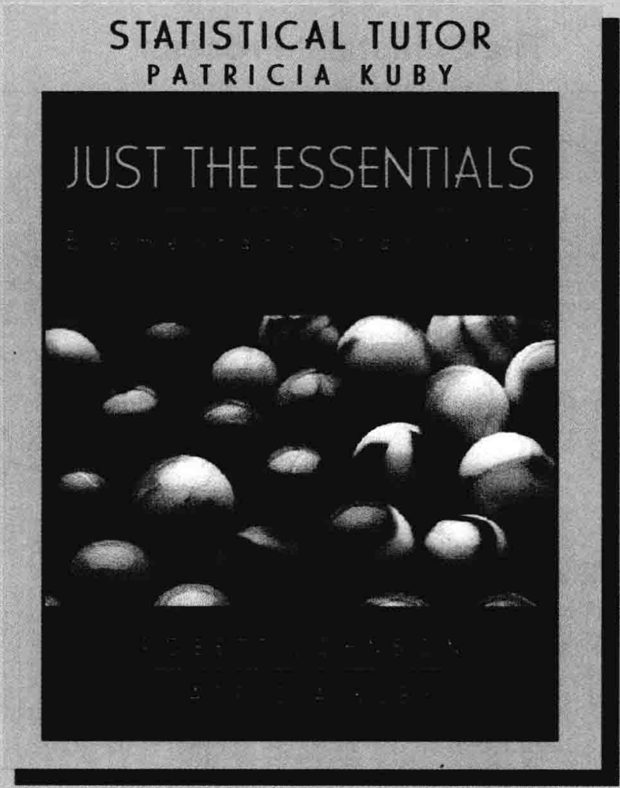
Based on the assumption that efficient and effective use of study time is a high priority, the following suggestions are made: Ask your instructor for a schedule indicating what material (pages in text) will be covered in each class and then spend 10 minutes before class reading that day's material. Do not take notes or highlight anything. This is a "warm-up" process, just "skim" read as a preview of the class to come. When you get to class you will be off to a "rolling start" since you have already seen the key words and concepts. When these concepts are discussed in class you will be hearing about them for the second time, and will therefore "hear" more in class. This will in turn reduce the study time required to sucessfully learn the material. (The before-class reading is just a skim read; it is not necessary to study the material at that time.)

As soon as possible after class, read the material again, thoroughly this time, making notes, highlighting the important points, and completing the assigned exercises. Do the margin exercises as you read.

Form a study group with two or three classmates. Friends are not always the best study partners; find study partners that have the same goals for this class and have the same determination to reach those goals as you. Establish a regular once-a-week meeting time and study together. Perhaps add an extra study session the week before a major exam. (True story: At midsemester a few semesters ago, three students all had grades of D in one of my statistics classes. None of them found this acceptable; they did not know each other and were all determined to do something about their grade. At my suggestion, they formed a study group and met twice a week for the rest of the semester. They also met with me a few times for extra help. Their final grades were B, B, and C+. The study group approach worked well for them and it can work for you.) In fact, the further you go in education, the more important study partners become.

GETTING ADDITIONAL ASSISTANCE

One of the best forms of assistance is to see your instructor outside of class, preferably during office hours that are set up for that purpose.



The Statistical Tutor is a student manual that:

a. Contains the complete solutions to all margin exercises and the odd-numbered exercises (the same exercises whose answers are in the back of the book).
b. Contains many overviews, helpful hints and suggestions to serve as a guide through the learning process.
c. Contains several review lessons to help refresh materials studied previously in other courses.