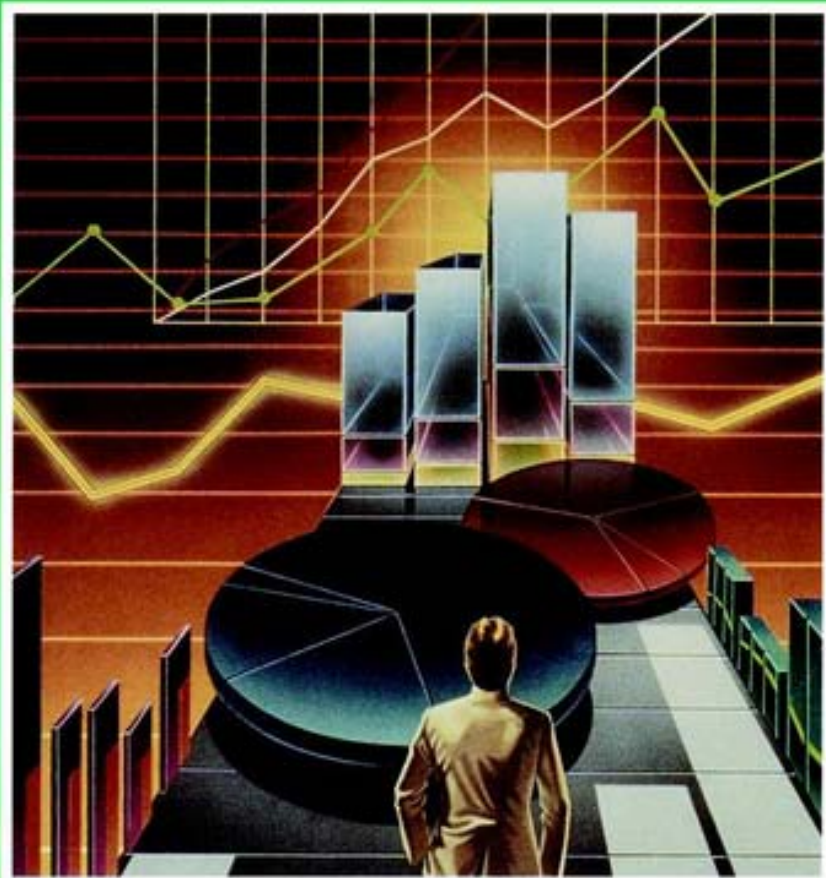


CLIFFS
QUICK
REVIEW®

STATISTICS



CliffsQuickReview™

Statistics

**By David H. Voelker, MA,
Peter Z. Orton, Ed M,
and Scott V. Adams**



Hungry Minds™

An International Data Group Company
New York, NY • Cleveland, OH • Indianapolis, IN

About the Authors

David Voelker is a respected educator and author from the San Francisco Bay area.

Peter Orton is Program Director of Global Curriculum Technology for IBM, designing leadership development Web sites, videos, and online simulators for IBMers globally. Orton lives in North Carolina with his wife and 10-year-old daughter.

Scott Adams is earning his PhD in physics at Vanderbilt University. His main interest is in molecular biophysics, especially electrophysiology.

Publisher's Acknowledgments

Editorial

Project Editor: Tracy Barr
Acquisitions Editor: Sherry Gomoll
Editorial Assistant: Michelle Hacker

Production

Indexer: TECHBOOKS Production Services
Proofreader: Christine Pingleton
Hungry Minds Indianapolis Production Services

CliffsQuickReview™ Statistics

Published by
Hungry Minds, Inc.
909 Third Avenue
New York, NY 10022
www.hungryminds.com
www.cliffsnotes.com

Copyright © 2001 Hungry Minds, Inc. All rights reserved. No part of this book, including interior design, cover design, and icons, may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording, or otherwise) without the prior written permission of the publisher.

Library of Congress Control Number: 2001024144

ISBN: 0-7645-6388-2

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

10/RU/QV/QR/IN

Distributed in the United States by Hungry Minds, Inc.

Distributed by CDG Books Canada Inc. for Canada; by Transworld Publishers Limited in the United Kingdom; by IDG Norge Books for Norway; by IDG Sweden Books for Sweden; by IDG Books Australia Publishing Corporation Pty. Ltd. for Australia and New Zealand; by TransQuest Publishers Pte Ltd. for Singapore, Malaysia, Thailand, Indonesia, and Hong Kong; by Gotop Information Inc. for Taiwan; by ICG Muse, Inc. for Japan; by Intersoft for South Africa; by Eyrolles for France; by International Thomson Publishing for Germany, Austria and Switzerland; by Distribuidora Cuspide for Argentina; by LR International for Brazil; by Galileo Libros for Chile; by Ediciones ZETA S.C.R. Ltda. for Peru; by WS Computer Publishing Corporation, Inc., for the Philippines; by Contemporanea de Ediciones for Venezuela; by Express Computer Distributors for the Caribbean and West Indies; by Micronesia Media Distributor, Inc. for Micronesia; by Chips Computadoras S.A. de C.V. for Mexico; by Editorial Norma de Panama S.A. for Panama; by American Bookshops for Finland.

For general information on Hungry Minds' products and services in the U.S., please call our Customer Care department at 800-762-2974.

For reseller information, including discounts and premium sales, please call our Reseller Customer Care department at 800-434-3422.

For information on where to purchase Hungry Minds' products outside the U.S., please contact our International Sales department at 317-572-3993 or fax 317-572-4002.

For consumer information on foreign language translations, please contact our Customer Care department at 800-434-3422, fax 317-572-4002, or e-mail rights@hungryminds.com.

For information on licensing foreign or domestic rights, please contact our Sub-Rights department at 212-884-5000.

For sales inquiries and special prices for bulk quantities, please contact our Order Services department at 800-434-3422 or write to Hungry Minds, Inc., Attn: Customer Care department, 10475 Crosspoint Boulevard, Indianapolis, IN 46256.

For information on using Hungry Minds' products and services in the classroom or for ordering examination copies, please contact our Educational Sales department at 800-434-2086 or fax 317-572-4005.

Please contact our Public Relations department at 212-884-5163 for press review copies or 212-884-5000 for author interviews and other publicity information or fax 212-884-5400.

For authorization to photocopy items for corporate, personal, or educational use, please contact Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, or fax 978-750-4470.

LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: THE PUBLISHER AND AUTHOR HAVE USED THEIR BEST EFFORTS IN PREPARING THIS BOOK. THE PUBLISHER AND AUTHOR MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS BOOK AND SPECIFICALLY DISCLAIM ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. THERE ARE NO WARRANTIES WHICH EXTEND BEYOND THE DESCRIPTIONS CONTAINED IN THIS PARAGRAPH. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES REPRESENTATIVES OR WRITTEN SALES MATERIALS. THE ACCURACY AND COMPLETENESS OF THE INFORMATION PROVIDED HEREIN AND THE OPINIONS STATED HEREIN ARE NOT GUARANTEED OR WARRANTED TO PRODUCE ANY PARTICULAR RESULTS, AND THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR EVERY INDIVIDUAL. NEITHER THE PUBLISHER NOR AUTHOR SHALL BE LIABLE FOR ANY LOSS OF PROFIT OR ANY OTHER COMMERCIAL DAMAGES, INCLUDING BUT NOT LIMITED TO SPECIAL, INCIDENTAL, CONSEQUENTIAL, OR OTHER DAMAGES.

Trademarks: Cliffs, CliffsNotes, the CliffsNotes logo, CliffsAP, CliffsComplete, CliffsTestPrep, CliffsQuickReview, CliffsNote-a-Day and all related logos and trade dress are registered trademarks or trademarks of Hungry Minds, Inc., in the United States and other countries. All other trademarks are property of their respective owners. Hungry Minds, Inc., is not associated with any product or vendor mentioned in this book.



Hungry Minds is a trademark of Hungry Minds, Inc.

Table of Contents

Introduction	1
Why You Need This Book	1
How to Use This Book	1
Visit Our Web Site	2
Chapter 1: Overview	3
Types of Statistics	3
Numerical statistics	3
Pictorial statistics	3
Method of Statistical Inference	4
Steps in the Process	4
Making Predictions	5
Comparing Results	6
Probability	7
Common Statistics Mistakes	7
Chapter 2: Graphic Displays	8
Bar Chart	9
Dot Plot	10
Pie Chart	10
Ogive	11
Frequency Histogram	12
Relative Frequency Histogram	13
Frequency Polygon	14
Frequency Distribution	15
Stem-and-Leaf	17
Box Plot (Box-and-Whiskers)	18
Scatter Plot	20
Chapter 3: Numerical Measures	23
Measures of Central Tendency	23
Mean	24
Median	24
Mode	25
Notation and formulae	25
Mean for grouped data	25
Median for grouped data	26
Symmetric distribution	28
Skewed curves	28
Measures of Variability	29
Range	30
Deviation and variance	30
Standard deviation	31
Notation	31

Empirical rule: The normal curve	.31
Shortcut formulae	.32
Percentile	.33
Quartiles and interquartile range	.33
Measurement Scales	.34
Chapter 4: Probability	.37
Classic Theory	.37
Relative Frequency Theory	.38
Probability of Simple Events	.38
Independent Events	.39
Dependent Events	.39
Probability of Joint Occurrences	.40
Multiplication rule	.40
Addition rule	.41
Non-Mutually Exclusive Outcomes	.42
Double Counting	.42
Conditional Probability	.44
Probability Distributions	.44
Discrete vs. continuous variables	.45
Total of probabilities of discrete variables	.46
The Binomial	.46
Binomial experiments	.46
Binomial table	.47
Mean and standard deviation	.47
Chapter 5: Sampling	.50
Populations, Samples, Parameters, and Statistics	.50
Sampling Distributions	.52
Random and Systematic Error	.53
Central Limit Theorem	.54
Properties of the Normal Curve	.54
Normal Approximation to the Binomial	.58
Chapter 6: Principles of Testing	.62
Stating Hypotheses	.62
The Test Statistic	.63
One- and Two-tailed Tests	.66
Type I and II Errors	.68
Significance	.69
Point Estimates and Confidence Intervals	.71
Estimating a Difference Score	.72
Univariate Tests: An Overview	.73

Chapter 7: Univariate Inferential Tests	76
One-sample z -test	76
Confidence interval for population mean using z	78
Choosing a sample size	79
One-sample t -test	79
Confidence interval for population mean using t	82
Two-sample z -test for Comparing Two Means	83
Two-sample t -test for Comparing Two Means	84
Confidence interval for comparing two means	85
Pooled variance method	86
Paired Difference t -test	88
Test for a Single Population Proportion	89
Confidence interval for a single population proportion	90
Choosing a sample size	90
Test for Comparing Two Proportions	91
Confidence interval for comparing two proportions	93
Chapter 8: Bivariate Relationships	96
Correlation	96
Finding the significance of r	100
The coefficient of determination	102
Simple Linear Regression	102
Confidence interval for the slope	105
Confidence interval for prediction	107
Assumptions and cautions	108
Chi-Square (χ^2)	110
Appendix A: Common Mistakes	114
Appendix B: Tables	117
CQR Review	136
CQR Resource Center	142
Glossary	145
Index	151

INTRODUCTION

Statistics are a powerful tool for finding patterns in data and inferring important connections between events in the real world. We are accustomed to seeing statistics every day — in the news, for example. However, many people find statistics confusing and intimidating. Often statistics are considered to make an argument irrefutable. On the other hand, some people perceive statistics as a trick, used to support outrageous claims.

This book is intended to clarify some of the elements of statistical reasoning and analysis. The format is simple, and there are many examples to help translate abstract ideas into concrete applications.

Why You Need This Book

Can you answer yes to any of these questions?

- Do you need to review the fundamentals of statistics fast?
- Do you need a course supplement to statistics?
- Do you need a concise, comprehensive reference for statistics?
- Do you need to prepare for a statistics exam?

If so, then CliffsQuickReview *Statistics* is for you!

How to Use This Book

You're in charge here. You get to decide how to use this book. You can read it cover to cover or just look up the information you need right now. Here are some recommended ways to search for topics.

- Use the Pocket Guide to find essential information, such as basic statistics formulas and abbreviations.
- Flip through the book looking for subject areas at the top of each page.
- Get a glimpse of what you'll gain from a chapter by reading through the "Chapter Check-In" at the beginning of each chapter.

- Use the “Chapter Checkout” at the end of each chapter to gauge your grasp of the important information you need to know.
- Test your knowledge more completely in the CQR Review and look for additional sources of information in the CQR Resource Center.
- Use the glossary to find key terms fast. This book defines new terms and concepts where they first appear in the chapter. If a word is bold-faced, you can find a more complete definition in the book’s glossary.
- Or flip through the book until you find what you’re looking for—we organized this book to gradually build on key concepts.

Visit Our Web Site

A great resource, www.cliffsnotes.com, features review materials, valuable Internet links, quizzes, and more to enhance your learning. The site also features timely articles and tips, plus downloadable versions of many CliffsNotes books.

When you stop by our site, don’t hesitate to share your thoughts about this book or any Hungry Minds product. Just click the Talk to Us button. We welcome your feedback!

Chapter 1

OVERVIEW

Chapter Check-In

- Learning about the different kinds of statistics
- Understanding the steps in statistical reasoning
- Learning several uses for statistical reasoning

For many people, statistics means numbers—numerical facts, figures, or information. Reports of industry production, baseball batting averages, government deficits, and so forth, are often called statistics. To be precise, these numbers are **descriptive statistics** because they are numerical data that describe phenomena. Descriptive statistics are as simple as the number of children in each family along a city block or as complex as the annual report released from the U.S. Treasury Department.

Types of Statistics

In these pages, we are concerned with two ways of representing descriptive statistics: numerical and pictorial.

Numerical statistics

Numerical statistics are numbers, but clearly, some numbers are more meaningful than others. For example, if you are offered a purchase price of \$1 for an automobile on the condition that you also buy a second automobile, the price of the second automobile would be a major consideration (its price could be \$1,000,000 or only \$1,000), and thus, the average—or **mean**—of the two prices would be the important statistic. The different meaningful numerical statistics are discussed in Chapter 3.

Pictorial statistics

Taking numerical data and presenting it in pictures or graphs is what is known as **pictorial statistics**. Showing data in the form of a graphic can

make complex and confusing information appear more simple and straightforward. The more commonly used graphics in statistics are discussed in Chapter 2.

Method of Statistical Inference

Statistics is also a method, a way of working with numbers to answer puzzling questions about both human and non-human phenomena. Questions answerable by using the “method” of statistics are many and varied: Which of several techniques is best for teaching reading to third-graders? Will a new medicine be more effective than the old one? Can you expect it to rain tomorrow? What’s the probable outcome of the next presidential election? Which assembly-line process produces fewer faulty carburetors? How can a polling organization make an accurate prediction of a national election by questioning only a few thousand voters? And so on.

For our purposes, statistics is both a collection of numbers and/or pictures *and* a process: the art and science of making accurate guesses about outcomes involving numbers.

So, fundamentally, the goals of statistics are

- To describe sets of numbers
- To make accurate inferences about groups based upon incomplete information

Steps in the Process

Making accurate guesses requires groundwork. The statistician must do the following in order to make an educated, better-than-chance hunch:

1. **Gather data (numerical information).**
2. **Organize the data (sometimes in a pictorial).**
3. **Analyze the data (using tests of significance and so forth).**

This book shows you how to do precisely these procedures and finally how to use your analysis to draw an **inference**—an educated statistical guess—to solve a particular problem. While these steps may appear simple (and indeed some of them are), sound statistical method requires that you perform them in certain prescribed ways. Those ways are the heart and soul of statistics.

Making Predictions

Suppose that you decide to sell commemorative T-shirts at your town's centennial picnic. You know that you can make a tidy profit, but only if you can sell most of your supply of shirts because your supplier won't buy any of them back. How many shirts can you reasonably plan on selling?

Your first question, of course, is this: How many people will be attending the picnic? Suppose you know that 100,000 tickets have been sold to the event. How many T-shirts should you purchase from your supplier to sell on the day of the event? 10,000? 50,000? 100,000? 150,000? How many of each size—small, medium, large, extra large? And the important question: How many must you sell in order to make some profit for your time and effort and *not* be left with an inventory of thousands of unsold T-shirts?

Ideally you need to have an accurate idea—*before* you buy your inventory of T-shirts—just how many ticket holders will want to purchase centennial T-shirts and which sizes they will want. But, obviously, you have neither the time nor the resources to ask each of those 100,000 people if they plan to purchase a commemorative T-shirt. If, however, you could locate a small number of those ticket holders—for example, 100—and get an accurate count of how many of those 100 would purchase a T-shirt, you'd have a better idea of how many of the 100,000 ticket holders would be willing to buy one.

That is, of course, if the 100 ticket holders that you asked (called the **sample**) are not too different in their intentions to purchase a T-shirt from the total 100,000 ticket holders (called the **population**). If the sample is indeed representative (typical) of the population, you could expect about the same percentage of T-shirt sales (and sizes) for the population as for the sample, all things being equal. So if 50 of your sample of 100 people say they can't wait to plunk down \$10 for a centennial T-shirt, it would be reasonable to expect that you would sell about 50,000 T-shirts to your population of 100,000. (At just \$1 profit per T-shirt, that's \$50,000!)

But before you start shopping to buy a yacht with your profits, remember that this prediction of total T-shirt sales relies heavily upon the sample being representative (similar to the population), which may not necessarily be the case with your sample. You may have inadvertently selected a sample that has more expendable income or is a greater proportion of souvenir T-shirt enthusiasts or who knows what else. Are you reasonably certain that the intentions of the sample of 100 ticket holders reflect the

intentions of the 100,000 ticket holders? If not, you may quite possibly be stuck with tens of thousands of centennial T-shirts and no profit to splurge on a yacht.

You can see why choosing a random sample is a critical part of the process of statistics and why this book includes a chapter on sampling (Chapter 5).

Comparing Results

Making predictions is only one use of statistics. Suppose you have recently developed a new headache/pain remedy that you call *Ache-Away*. Should you produce *Ache-Away* in quantity and make it available to the public? That would depend, among other concerns, upon whether *Ache-Away* is more effective than the old remedy. How can you determine that?

One way might be to administer both remedies to two separate groups of people, collect data on the results, and then statistically analyze that data to determine if *Ache-Away* is more effective than the old remedy. And what if the results of this test showed *Ache-Away* to be more effective? How certain can you be that *this* particular test administration is indicative of *all* tests of these two remedies? Perhaps the group taking the old remedy (the control group) and the group taking *Ache-Away* (the treatment group) were so dissimilar that the results were due *not* to the pain remedies but to the differences between the groups.

It's possible that the results of this test are far off the results that you would get if you tried the test several more times. You certainly do not want to foist a questionable drug upon an unsuspecting public based upon untypical test results. How certain can you be that you can put your faith in the results of your tests? You can see that the problems of comparing headache remedy results can produce headaches of their own.

This type of statistical procedure is discussed in Chapters 6 and 7.

Probability

One of the most familiar uses of statistics is to determine the chance of some occurrence. For instance, what are the chances that it will rain tomorrow or that the Boston Red Sox will win a World Series? These kinds of probabilities, although interesting, are not the variety under discussion here. Rather, we are examining the probability in statistics that deals with classic theory and frequency theory—events that can be repeated over and over again, independently, and under the same conditions.

Coin tossing and card drawing are two such examples. A *fair* coin (one that is not weighted or *fixed*) has an equal chance of landing heads as landing tails. A typical deck of cards has 52 different cards—13 of each suit (hearts, clubs, diamonds, and spades)—and each card or suit has an equal chance of being drawn. This kind of event forms the basis of your understanding of probability and enables you to find solutions to everyday problems that seem far removed from coin tossing or card drawing.

Probability is discussed in Chapter 4.

Common Statistics Mistakes

Whether you are new to the study of statistics or have been using statistical analysis for years, there is a likelihood that you will make (or already have made) errors in your application of principles explained in this book. In fact, some particular errors are committed so frequently in statistical analysis that we feel obliged to call them to your attention in the hope that by, being prepared for them, you will avoid their pitfalls. Appendix A lists these common mistakes.

Chapter Checkout

1. True or False: Statistics are only about things in the past, so you can't use them to make predictions.
2. True or False: An inference is an educated guess, made after analyzing data.
3. True or False: A population is a randomly chosen portion of a larger sample.
4. True or False: Statistical analysis will allow you to be absolutely sure of your conclusions.

Answers: 1. False 2. True 3. False 4. False

Chapter 2

GRAPHIC DISPLAYS

Chapter Check-In

- ❑ Displaying data in simple graphical formats such as bar charts, pie charts, dot plots, and ogives
- ❑ Constructing frequency histograms and distributions from data and learning about the basic properties of these displays
- ❑ Using stem-and-leaf plots and box plots to display numerical measures of data
- ❑ Understanding scatter plots

Pie charts, bar charts, and histograms are graphic displays of numerical data—simple collections of numbers made more understandable by use of visuals. This section discusses the more common graphic displays used in statistical analysis.

As an example, consider the yearly expenditures of a college undergraduate. After collecting her **data** (expense records) for the past year, she finds the expenditures shown in Table 2-1.

Table 2-1 Yearly Expenses of a College Undergraduate

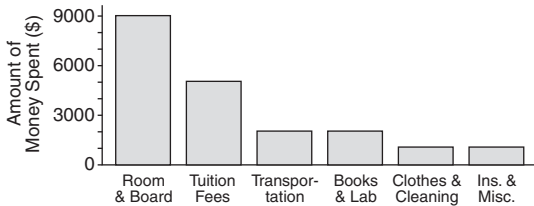
<i>Item</i>	<i>Amount</i>
Tuition fees	\$5,000
Room and board	9,000
Books and lab	2,000
Clothes/cleaning	1,000
Transportation	2,000
Insurance and miscellaneous	1,000

These figures, although presented in categories, do not allow for easy analysis. The reader must expend extra effort in order to compare amounts spent or relate individual proportions to the total. For ease of analysis, these data can be presented pictorially.

Bar Chart

One way to pictorially display the numbers shown in Table 2-1 is with a **bar chart** (see Figure 2-1).

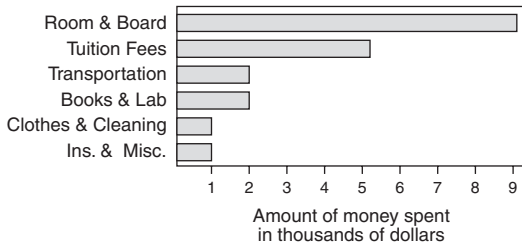
Figure 2-1 Vertical bar chart presentation of the expenditures of a college undergraduate for the past year.



Comparing the size of the bars, you can quickly see that room and board expenses are nearly double tuition fees, and tuition fees are more than double books and lab or transportation expenses.

A bar chart may also be placed on its side with the bars going horizontally, as shown in Figure 2-2.

Figure 2-2 Horizontal bar chart presentation of the expenditures of a college undergraduate for the past year.



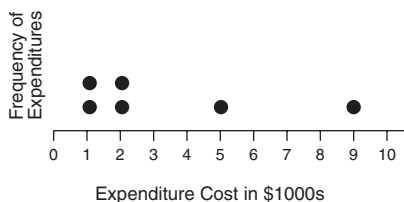
In each bar chart, vertical or horizontal, the amounts are ordered from highest to lowest or lowest to highest, making the chart clearer and more easily understood. Space is left between each of the bars in order to define the categories as being different.

The bottom line in Figure 2-1 and the left-hand side in Figure 2-2 indicate 0. Although typical, this presentation need not always be used. Finally, while the lengths of the bars may be different, their *thicknesses* are the same.

Dot Plot

Dot plots are similar to bar graphs. Typically used for a small set of values, a dot plot uses a *dot* for each unit of measurement; thus, the preceding undergraduate expense data represented in a dot plot could look like that shown in Figure 2-3.

Figure 2-3 Dot plot of the expenditures of a college undergraduate for the past year.

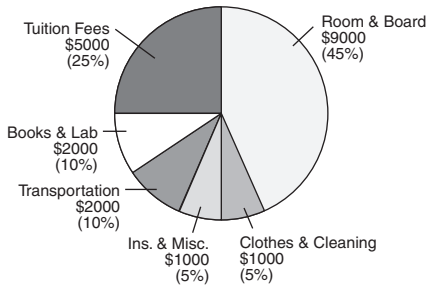


Pie Chart

The previous charts have a limitation: It's difficult to see what portion of the total each item comprises. If knowing about a "part of the whole" is an important consideration, then a **pie chart** is a better choice for showing the same data. A pie chart may also display each category's percentage of the total. Using the same data from the undergraduate expenditures, we get the pie chart shown in Figure 2-4.

The parts of the circle (or *pie*) match in size each category's percentage of the total. The parts of the pie chart are ordered from highest to lowest for easier interpretation of the data. Pie charts work best with only a few categories; too many categories will make a pie chart confusing.

Figure 2-4 Pie chart presentation of the expenditures of a college undergraduate for the past year.

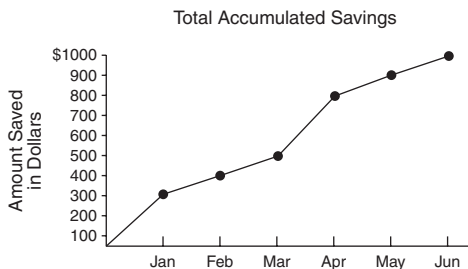


Ogive

Data may be expressed using a single line. An **ogive** (a cumulative line graph) is best used when you want to display the total at any given time. The relative slopes from point to point will indicate greater or lesser increases; for example, a steeper slope means a greater increase than a more gradual slope. An ogive, however, is not the ideal graphic for showing comparisons between categories because it simply combines the values in each category and thus indicates an *accumulation*, a growing or lessening total. If you simply want to keep track of a total and your individual values are periodically combined, an ogive is an appropriate display.

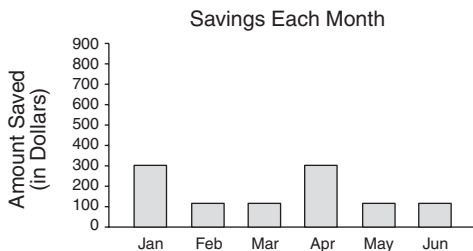
For example, if you saved \$300 in both January and April and \$100 in each of February, March, May, and June, an ogive would look like Figure 2-5.

Figure 2-5 Ogive of accumulated savings for one year.



An ogive displays a *running total*. Although each individual month's savings could be expressed in a bar chart (as shown in Figure 2-6), you could not easily see the amount of total growth or loss, as you can in an ogive.

Figure 2-6 Vertical bar chart of accumulated savings for one year.



The choice of graphic display, therefore, depends on what information is important for your purposes: percentages (parts of the whole), running total, comparisons of categories, and so forth.

Frequency Histogram

One of the more commonly used pictorials in statistics is the frequency histogram, which in some ways is similar to a bar chart and tells how many items are in each numerical category. For example, suppose that after a garage sale, you want to determine which items were the most popular: the high-priced items, the low-priced items, and so forth. Let's say you sold a total of 32 items for the following prices: \$1, \$2, \$2, \$2, \$5, \$5, \$5, \$5, \$7, \$8, \$10, \$10, \$10, \$10, \$11, \$15, \$15, \$15, \$19, \$20, \$21, \$21, \$25, \$25, \$29, \$29, \$29, \$30, \$30, \$30, \$35, \$35.

The items sold *ranged* in price from \$1 to \$35. First, divide this **range** of \$1 to \$35 into a number of categories, called **class intervals**. Typically, no fewer than 5 or more than 20 class intervals work best for a frequency histogram.

Choose the first class interval to include your lowest (smallest value) data and make sure that no *overlap* exists so that one piece of data does not fall into two class intervals. For example, you would not have your first class interval be \$1-\$5 and your second class interval \$5-\$10 because the four items that sold for \$5 would belong in both the first and the second intervals. Instead, use \$1-\$5 for the first interval and \$6-\$10 for the second. Class intervals are mutually exclusive.

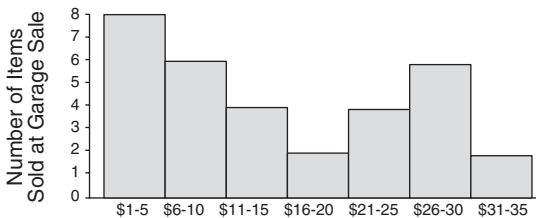
First, make a table of how your data is distributed. The number of observations that falls into each class interval is called the **class frequency**. See Table 2-2.

Table 2-2 Distribution of Items Sold at Garage Sale

<i>Class</i>	<i>Interval</i>	<i>Frequency</i>
1	\$1-\$5	8
2	\$6-\$10	6
3	\$11-\$15	4
4	\$16-\$20	2
5	\$21-\$25	4
6	\$26-\$30	6
7	\$31-\$35	2

Note that each class interval has the same width. That is, \$1-\$5 has a width of five dollars, inclusive; \$6-\$10 has a width of five dollars, inclusive; \$11-\$15 has a width of five dollars, inclusive, and so forth. From the data, a frequency histogram would look like that shown in Figure 2-7:

Figure 2-7 Frequency histogram of items sold at a garage sale.



Unlike in a bar chart, the class intervals are drawn immediately adjacent to each other.

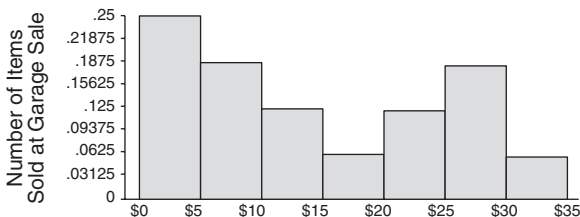
Relative Frequency Histogram

A **relative frequency histogram** uses the same information as a frequency histogram but compares each class interval to the total number of items. For example, the first interval (\$1-\$5) contains 8 out of the total of 32 items, so the relative frequency of the first class interval is $8/32$. See Table 2-3.

Table 2-3 Distribution of Items Sold at Garage Sale, Including Relative Frequencies

<i>Class</i>	<i>Interval</i>	<i>Frequency</i>	<i>Relative Frequency</i>
1	\$1-\$5	8	.25
2	\$6-\$10	6	.1875
3	\$11-\$15	4	.125
4	\$16-\$20	2	.0625
5	\$21-\$25	4	.125
6	\$26-\$30	6	.1875
7	\$31-\$35	2	.0625

The only difference between a frequency histogram and a relative frequency histogram is that the vertical axis uses relative or proportional frequency instead of simple frequency (see Figure 2-8).

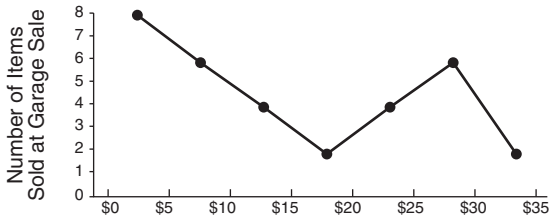
Figure 2-8 Relative frequency histogram of items sold at a garage sale.

Frequency Polygon

Relative frequencies of class intervals can also be shown in a **frequency polygon**. In this chart, the frequency of each class is indicated by points or dots drawn at the midpoints of each class interval. Those points are then connected by straight lines.

Comparing the frequency polygon (shown in Figure 2-9) to the frequency histogram (refer to Figure 2-7), you will see that the major difference is that points replace the bars.

Figure 2-9 Frequency polygon display of items sold at a garage sale.



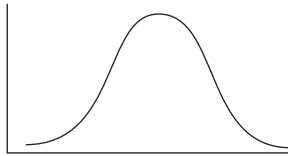
Whether to use bar charts or histograms depends on the data. For example, you may have **qualitative** data—numerical information about categories that vary significantly in kind. For instance, gender (male or female), types of automobile owned (sedan, sports car, pickup truck, van, and so forth), and religious affiliations (Christian, Jewish, Moslem, and so forth) are all qualitative data. On the other hand, **quantitative** data can be measured in amounts: age in years, annual salaries, inches of rainfall. Typically, qualitative data are better displayed in bar charts, quantitative data in histograms. As you will see, histograms play an extremely important role in statistics.

Frequency Distribution

Frequency distributions are like frequency polygons (refer to Figure 2-9); however, instead of straight lines, a frequency distribution uses a smooth curve to connect the points and, similar to a graph, is plotted on two axes: The horizontal axis from left to right (or x axis) indicates the different possible **values** of some **variable** (a phenomenon where observations vary from trial to trial). The vertical axis from bottom to top (or y axis) measures frequency or how many times a particular value occurs.

For example, in Figure 2-10, the x axis might indicate annual income (the values would be in thousands of dollars); the y axis might indicate frequency (millions of people or percentage of working population). Notice that in Figure 2-10 the highest percentage of the working population would thus have an annual income in the middle of the dollar values. The lowest percentages would be at the extremes of the values: nearly 0 and extremely high.

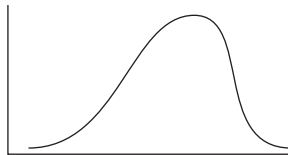
Figure 2-10 A symmetric bell curve.



Notice that this frequency curve displays perfect **symmetry**; that is, one half (the left side) is the mirror image of the other half (the right side). A **bell-shaped** or **mound-shaped curve** is also *normal*, giving it special properties.

The negatively skewed curve, shown in Figure 2-11, is *skewed to the left*. Its greatest frequency occurs at a value near the right of the graph.

Figure 2-11 Negatively skewed bell curve.



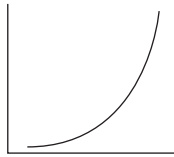
The positively skewed curve (see Figure 2-12) is *skewed to the right*. Its greatest frequency occurs at a value near the left of the graph. This distribution is probably a more accurate representation of the annual income of working Americans than is Figure 2-10.

Figure 2-12 Positively skewed bell curve.



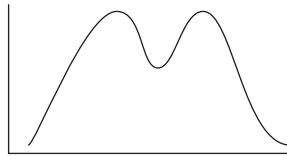
Figure 2-13 shows a **J-shaped** curve.

Figure 2-13 J-shaped curve.



Unlike Figure 2-10, a **bimodal** curve (shown in Figure 2-14) has two high points.

Figure 2-14 A bimodal curve has two maximum peaks.



Stem-and-Leaf

Another useful pictorial is the **stem-and-leaf**. It is similar to a histogram in that it shows the range of data, where the data are concentrated, if there are any **outliers** (occasional extremely high or extremely low scores), and the general shape of the distribution.

For example, look at the following data—test scores of 17 high school students: 69, 75, 77, 79, 82, 84, 87, 88, 89, 89, 89, 90, 91, 93, 96, 100, 100. The stem-and-leaf takes the first digit of each score as the **stem** and uses the remaining digits as the **leaf**.

As an example, for the score of 69, the 6 is the stem and the 9 is the leaf; for the next three grades (75, 77, and 79), 7 is the stem, and 5, 7, and 9 are the leaves.

Note, too, that along the extreme left side of the chart is a vertical column that keeps a running count or total. (Some stem-and-leaves do not include this running total.) Having a running total enables the reader to locate quickly the median. Median is discussed further in Chapter 3.

The completed stem-and-leaf for the high school students' test scores looks like that shown in Table 2-4.

Table 2-4 Stem-and-Leaf Display of Students' Test Scores

<i>Running Count</i>	<i>Stem</i>	<i>Leaves</i>
1	6	9
4	7	5 7 9
11	8	2 4 7 8 9 9 9
15	9	0 1 3 6
17	10	0 0

Notice that, like a histogram, each stem determines a class interval and, also like a histogram, the class intervals are all equal. (In this case, each interval width is from a possible low of 0 to a possible high of 9.) All 17 scores are displayed in the stem-and-leaf so that you can see not only the frequencies and the shape of the distribution but also the actual value of every score.

Box Plot (Box-and-Whiskers)

Box plots, sometimes called **box-and-whiskers**, take the stem-and-leaf one step further. A box plot will display a number of values of a distribution of numbers:

- The median value
- The lower quartile (Q_1)
- The upper quartile (Q_3)
- The interquartile range or *IQR* (distance between the lower and upper quartiles)
- The symmetry of the distribution
- The highest and lowest values

Use the set of values in Table 2-5 to examine each of the preceding items.

Table 2-5 Verbal SAT Scores of 20 High School Students.

280, 340, 440, 490, 520, 540, 560, 560, 580, 580,
600, 610, 630, 650, 660, 680, 710, 730, 740, 740

The median (the middle value in a set that has been ordered lowest to highest) is the value above which half of the remaining values fall and below which the other half of the remaining values fall. Because there are an even number of scores in our example (20), the median score is the average of the two middle scores (10th and 11th)—580 and 600—or 590.

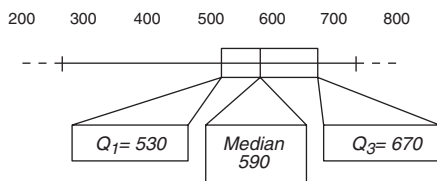
The **lower quartile** (Q_1 or 25th percentile) is the median of the bottom half. The bottom half of this set consists of the first 10 numbers (ordered from low to high): 280, 340, 440, 490, 520, 540, 560, 560, 580, 580. The median of those 10 is the average of the 5th and 6th scores—520 and 540—or 530. The lower quartile score is 530.

The **upper quartile** (Q_3 or 75th percentile) is the median score of the top half. The top half of this set consists of the last 10 numbers: 600, 610, 630, 650, 660, 680, 710, 730, 740, 740. The median of these 10 is again the average of the 5th and 6th scores—in this case, 660 and 680—or 670. So 670 is the upper quartile score for this set of 20 numbers.

A **box plot** can now be constructed as follows: The left side of the box indicates the lower quartile; the right side of the box indicates the upper quartile; and the line inside the box indicates the median. A horizontal line is then drawn from the lowest value of the distribution through the box to the highest value of the distribution. (This horizontal line is the “whiskers.”)

Using the Verbal SAT scores in Table 2-5, a box plot would look like Figure 2-15.

Figure 2-15 A box plot of SAT scores displays median and quartiles.



Without reading the actual values, you can see by looking at the box plot in Figure 2-15 that the scores range from a low of 280 to a high of 740; that the lower quartile (Q_1) is at 530; the median is at 590; and the upper

quartile (Q_3) is at 670. Because the median is slightly nearer the lower quartile than the upper quartile and the interquartile range is situated far to the right of the range of values, the distribution departs from symmetry.

Scatter Plot

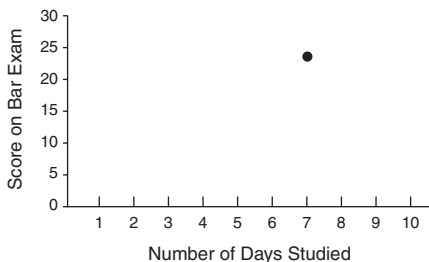
Sometimes you want to display information about the relationship involving two different phenomena. For example, suppose you collected data about the number of days that law school candidates studied for a state bar examination and their resulting scores on the exam. The data from eight candidates is shown in Table 2-6.

Table 2-6 Law School Candidates' Prep Times and Test Scores

Candidate #	1	2	3	4	5	6	7	8
Days studied	7	9	5	1	8	4	3	6
Score earned	23	25	14	5	22	15	11	17

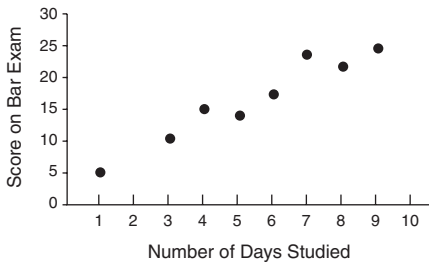
One dot would then be plotted for each examinee, giving a total of only 8 dots, yet displaying 16 pieces of numerical information. For example, candidate #1 studied for 7 days and received a score of 23. Candidate #1's dot would be plotted at a vertical of 23 and a horizontal of 7 (see Figure 2-16).

Figure 2-16 A representative point of data on a scatter plot.



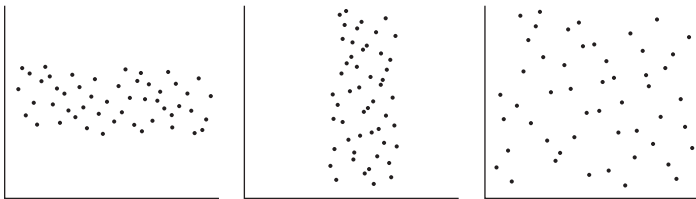
A completed scatter plot would look like Figure 2-17.

Figure 2-17 A scatter plot displaying the relationship between preparation time and test score.



There is a strong **positive relationship** between the number of days studied and the score on the bar exam; that is, the data displayed indicates that an increase in days studied for the exam correlates with an increase in score achieved. A **negative relationship** would be indicated if the dots suggest a line going down from left to right, meaning that as one variable increases, the other decreases. And no relationship would be indicated if the scatter plot dots suggest a completely horizontal line, a completely vertical line, or no line at all (see Figure 2-18).

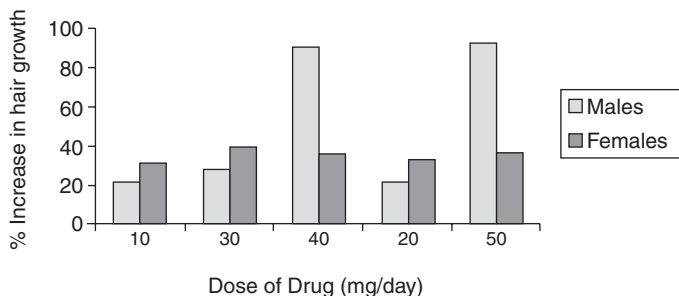
Figure 2-18 Scatter plots which display no relationship between the variables plotted.



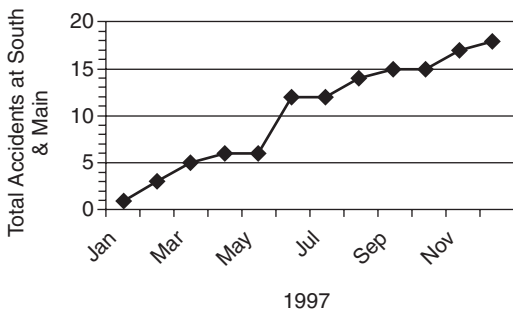
These relationships are discussed in more detail in Chapter 8, which discusses bivariate relationships.

Chapter Checkout

1. A car salesman takes inventory and finds that he has a total of 125 cars to sell. Of these, 97 are the 2001 model, 11 are the 2000 model, 12 are the 1999 model, and 5 are the 1998 model. Which two types of charts are most appropriate to display the data? Construct one of the plots.
2. Given the bar graph shown here, take the information and construct a scatter plot.



3. Using this ogive, answer the following questions:



- a. How many total accidents have occurred at the intersection of South and Main Streets in the last year?
- b. How many occurred between June and August?
- c. What interval saw the greatest increase in the rate of accidents?

Answers: 1. Pie or bar 2. Provide your own answer. 3a. 18 b. 2 c. May-June

Chapter 3

NUMERICAL MEASURES

Chapter Check-In

- ❑ Calculating and understanding the central tendencies of data
- ❑ Using measures of variability to describe the spread in data
- ❑ Discussing the principle types of measurement scales

This section discusses how to graph numerical information—to transform it into a picture—to get a better sense of the data. Another way to get a sense of data is to use numerical measures, certain numbers that give special insight into your values. Two types of numerical measures are important in statistics: **measures of central tendency** and **measures of variation**. Each of these individual measures can provide information about the entire set of data.

Measures of Central Tendency

Measures of central tendency are numbers that tend to cluster around the “middle” of a set of values. Three such middle numbers are the **mean**, the **median**, and the **mode**.

For example, suppose your earnings for the past week were the values shown in Table 3-1.

Table 3-1 Earnings for the Past Week

<i>Day</i>	<i>Amount</i>
Monday	\$350
Tuesday	\$150

(continued)

Table 3-1 (continued)

<i>Day</i>	<i>Amount</i>
Wednesday	\$100
Thursday	\$350
Friday	\$50

Mean

You could express your daily earnings from Table 3-1 in a number of ways. One way is to use the average, or **mean**, of each day. The arithmetic mean is the sum of the measures in the set divided by the number of measures in the set. Totaling all the measures and dividing by the number of measures, you get $\$1,000 \div 5 = \200 .

Median

Another measure of central tendency is the **median**, which is defined as the middle value when the numbers are arranged in increasing or decreasing order. When you order the daily earnings shown in Table 3-1, you get \$50, \$100, \$150, \$350, \$350. The middle value is \$150, and therefore, \$150 is the median.

If there is an even number of items in a set, the median is the average of the two middle values. For example, if we had four values—4, 10, 12, 26—the median would be the average of the two middle values, 10 and 12; thus, 11 is the median in that case. The median may sometimes be a better indicator of central tendency than the mean, especially when there are **outliers**, or extreme values.

Example 1: For example, given the four annual salaries of a corporation shown in Table 3-2, determine the mean and the median.

Table 3-2 Four Annual Salaries

<i>Position</i>	<i>Salary</i>
CEO	\$1,000,000
Manager	\$50,000
Administrative	\$30,000
Clerical	\$20,000

The mean of these four salaries is \$275,000. The median is the average of the middle two salaries, or \$40,000. In this instance, the median appears to be a better indicator of central tendency because the CEO's salary is an extreme outlier causing the mean to lie far from the other three salaries.

Mode

Another indicator of central tendency is the **mode**, or the value that occurs most often in a set of numbers. In the set of weekly earnings in Table 3-1, the mode would be \$350 because it appears twice and the other values appear only once.

Notation and formulae

The mean of a sample is typically denoted as \bar{x} (read as x bar). The mean of a population is typically denoted as μ (read as mew.) The sum (or total) of measures is typically denoted with a Σ . The formula for a sample mean is

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

Mean for grouped data

Occasionally you may have data that do not consist of actual values but rather **grouped measures**. For example, you may know that, in a certain working population, 32 percent earn between \$25,000 and \$29,999, 40 percent earn between \$30,000 and \$34,999, 27 percent earn between \$35,000 and \$39,999, and the remaining 1 percent earn between \$80,000 and \$85,000. This type of information is similar to that presented in a frequency table. (Refer to Chapter 2 for information about frequency tables and graphic displays.) Although you do not have precise individual measures, you can, nevertheless, compute measures for **grouped data**, data presented in a frequency table.

The formula for a sample mean for grouped data is

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

where x denotes the midpoint of the interval; fx denotes the sum of the measurements of the interval.

For example, if 8 is the midpoint of a class interval and there are 10 measurements in the interval, $fx = 10(8) = 80$, the sum of the 10 measurements in the interval.

$\sum fx$ denotes the sum of all the measurements in all class intervals. Dividing that sum by the number of measurements yields the sample mean for grouped data.

For example, consider the information shown in Table 3-3.

Table 3-3 Distribution of the Prices of Items Sold at a Garage Sale

<i>Class Interval</i>	<i>Frequency (f)</i>	<i>Midpoint (x)</i>	<i>fx</i>
\$1-\$5	8	3	24
\$6-\$10	6	8	48
\$11-\$15	4	13	52
\$16-\$20	2	18	36
\$21-\$25	4	23	92
\$26-\$30	6	28	168
\$31-\$35	2	33	66
	$n = 32$		$\sum fx = 486$

Substituting into the formula:

$$\bar{x} = \frac{\sum fx}{n} = \frac{486}{32} = 15.19$$

Therefore, the average price of items sold was about \$15.19.

Median for grouped data

The median for grouped data may not necessarily be computed precisely because the actual values of the measurements may not be known. In that case, you can find the particular interval that contains the median and then approximate the median.

Using Table 3-3, you can see that there is a total of 32 measures. The median is between the 16th and 17th measure, and therefore the median falls within the \$11-\$15 interval. The formula for the best approximation of the median for grouped data is

$$\text{median} = L + \frac{w}{f_{\text{med}}} (.5n - \sum f_b)$$

where L = lower class limit of the interval that contains the median

n = total number of measurements

w = class width

f_{med} = frequency of the class containing the median

$\sum f_b$ = sum of the frequencies for all classes before the median class

Consider the information in Table 3-4.

Table 3-4 Distribution of Prices of Items Sold at a Garage Sale

Class Boundaries	Frequency (f)
\$995-\$5.995	8
\$5.995-\$10.995	6
\$10.995-\$15.995	4
\$15.995-\$20.995	2
\$20.995-\$25.995	4
\$25.995-\$30.995	6
\$30.995-\$35.995	2
	$n = 32$

As we already know, the median is located in class interval \$11-\$15. So $L = 11$, $n = 32$, $w = 4.99$, $f_{\text{med}} = 4$, $\sum f_b = 14$.

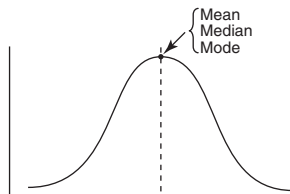
Substituting into the formula:

$$\begin{aligned}
 \text{median} &= L + \frac{H}{f_{\text{modal}}} (.5n - \sum f_c) \\
 &= 10.995 + \frac{4.99}{4} (.5(32) - 14) \\
 &= 10.995 + \frac{4.99}{4} (16 - 14) \\
 &= 10.995 + \frac{4.99}{4} (2) \\
 &= 10.995 + 2.495 = 13.49
 \end{aligned}$$

Symmetric distribution

In a distribution displaying perfect symmetry, the mean, the median, and the mode are all at the same point, as shown in Figure 3-1.

Figure 3-1 For a symmetric distribution, mean, median, and mode are equal.



Skewed curves

As you have seen, an outlier can significantly alter the mean of a series of numbers, whereas the median will remain at the center of the series. In such a case, the resulting curve drawn from the values will appear to be **skewed**, tailing off rapidly to the left or right. In the case of negatively skewed or positively skewed curves, the median remains in the center of these three measures.

Figure 3-2 shows a negatively skewed curve.

Figure 3-2 A negatively skewed distribution, mean < median < mode.

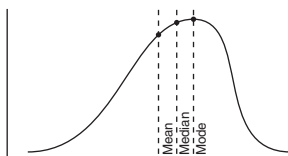
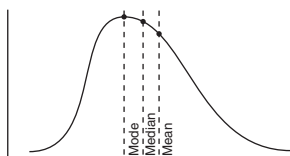


Figure 3-3 shows a positively skewed curve.

Figure 3-3 A positively skewed distribution, mode < median < mean.



Measures of Variability

Measures of central tendency locate only the center of a distribution of measures. Other measures are often needed to describe data.

For example, consider the two sets of numbers presented in Table 3-5.

Table 3-5 Earnings of Two Employees

<i>Earnings of Employee A/Day</i>	<i>Earnings of Employee B/Day</i>
\$200	\$200
\$210	\$20
\$190	\$400
\$201	\$0
\$199	\$390
\$195	\$10
\$205	\$200
\$200	\$380

The mean, the median, and the mode of each employee's daily earnings all equal \$200. Yet, there is significant difference between the two sets of numbers. For example, the daily earnings of Employee A are much more consistent than those of Employee B, which show great variation. This example illustrates the need for **measures of variation** or spread.

Range

The most elementary measure of variation is **range**. Range is defined as the difference between the largest and smallest values. Note that the range for Employee A is $\$210 - \$190 = \$20$; the range for Employee B is $\$400 - \$0 = \$400$.

Deviation and variance

The deviation is defined as the distance of the measurements away from the mean. In Table 3-5, Employee A's earnings have considerably less deviation than do Employee B's. The **variance** is defined as the sum of the squared deviations of n measurements from their mean divided by $(n-1)$.

So, from the table of employee earnings (Table 3-5), the mean for Employee A is \$200, and the deviations from the mean are as follows:

$$0, +10, -10, +1, -1, -5, +5, 0$$

The squared deviations from the mean are, therefore, the following:

$$0, 100, 100, 1, 1, 25, 25, 0$$

The sum of these squared deviations from the mean equals 252. Dividing by $(n-1)$, or $8-1$, yields $252/7$, which equals 36. The variance = 36.

For Employee B, the mean is also \$200, and the deviations from the mean are as follows:

$$0, -180, +200, -200, +190, -190, 0, +180$$

The squared deviations are, therefore, the following:

$$0; 32,400; 40,000; 40,000; 36,100; 36,100; 0; 32,400$$

The sum of these squared deviations equals 217,000. Dividing by $(n - 1)$ yields $217,000/7$, which equals 31,000.

Although they earned the same totals, there is significant difference in variance between the daily earnings of the two employees.

Standard deviation

The **standard deviation** is defined as the positive square root of the variance; thus, the standard deviation of Employee A's daily earnings is the positive square root of 36, or 6. The standard deviation of Employee B's daily earnings is the positive square root of 31,000, or about 176.

Notation

s^2 denotes the variance of a sample.

σ^2 denotes the variance of a population.

s denotes the standard deviation of a sample.

σ denotes the standard deviation of a population.

Empirical rule: The normal curve

The practical significance of the standard deviation is that with mound-shaped (bell-shaped) distributions, the following rules apply:

- The interval from one standard deviation below the mean to one standard deviation above the mean contains approximately 68 percent of the measurements.
- The interval from two standard deviations below the mean to two standard deviations above the mean contains approximately 95 percent of the measurements.
- The interval from three standard deviations below the mean to three standard deviations above the mean contains approximately all of the measurements.

These mound-shaped curves are usually called **normal distributions** or **normal curves** (see Figures 3-4 through 3-6).

Figure 3-4 The interval $\pm \sigma$ from the mean contains 68% of the measurements.

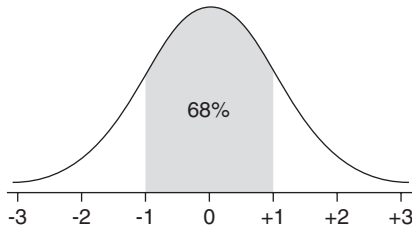


Figure 3-5 The interval $\pm 2\sigma$ from the mean contains 95% of the measurements.

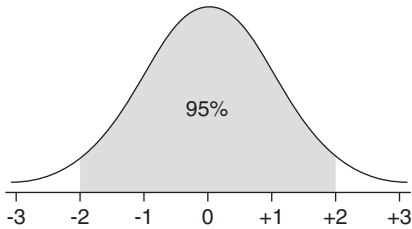
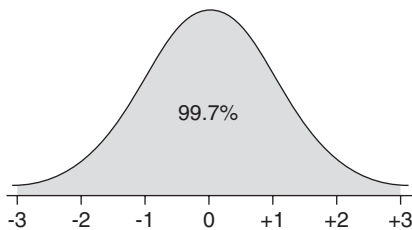


Figure 3-6 The interval $\pm 3\sigma$ from the mean contains 99.7% of the measurements.



Shortcut formulae

A shortcut method of calculating variance and standard deviation requires two quantities: sum of the values and sum of the squares of the values.

Σx = sum of the measures

Σx^2 = sum of the squares of the measures

For example, using these six measures: 3, 9, 1, 2, 5, and 4:

$$\begin{aligned}\Sigma x &= 3 + 9 + 1 + 2 + 5 + 4 = 24 \\ \Sigma x^2 &= 3^2 + 9^2 + 1^2 + 2^2 + 5^2 + 4^2 \\ &= 9 + 81 + 1 + 4 + 25 + 16 = 136\end{aligned}$$

The quantities are then substituted into the shortcut formula to find $\Sigma(x - \bar{x})^2$.

$$\begin{aligned}\Sigma(x - \bar{x})^2 &= \Sigma x^2 - \frac{(\Sigma x)^2}{n} \\ &= 136 - \frac{(24)^2}{6} \\ &= 136 - \frac{576}{6} = 40\end{aligned}$$

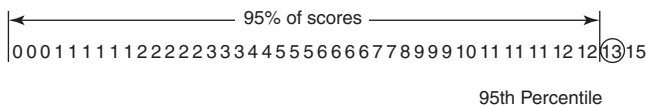
The variance and standard deviation are now found as before:

$$\begin{aligned}s^2 &= \frac{\Sigma(x - \bar{x})^2}{n - 1} = \frac{40}{5} = 8 \\ s &= \sqrt{s^2} = \sqrt{8} = 2.828\end{aligned}$$

Percentile

The **Nth percentile** is defined as the value such that N percent of the values lie below it. So, for example, a score of 5 percent from the top score would be the 95th percentile because it is above 95 percent of the other scores (see Figure 3-7).

Figure 3-7 95% of the test scores are less than the value of the 95th percentile.



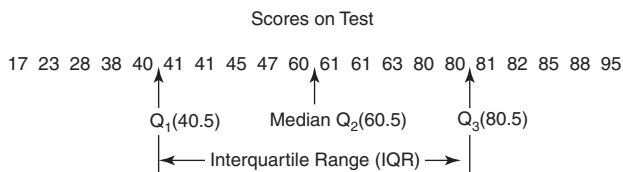
Quartiles and interquartile range

The **lower quartile** (Q_1) is defined as the 25th percentile; thus, 75 percent of the measures are above the lower quartile. The **middle quartile** (Q_2) is defined as the 50th percentile, which is, in fact, the median of all

the measures. The **upper quartile** (Q_3) is defined as the 75th percentile; thus, only 25 percent of the measures are above the upper quartile.

The **interquartile range** (*IQR*) is the set of measures lying between the upper and lower quartiles, or the middle 50 percent of the measures (see Figure 3-8).

Figure 3-8 An illustration of the quartiles and the interquartile range.



Measurement Scales

Different measurement scales allow for different levels of exactness, depending upon the characteristics of the variables being measured. The four types of scales available in statistical analysis are

- **Nominal:** A scale that measures data by name only. For example, religious affiliation (measured as Jewish, Christian, Buddhist, and so forth), political affiliation (measured as Democratic, Republican, Libertarian, and so forth), or style of automobile (measured as sedan, sports car, station wagon, van, and so forth).
- **Ordinal:** Measures by rank order only. Other than rough order, no precise measurement is possible. For example, medical condition (measured as satisfactory, fair, poor, guarded, serious, and critical); social-economic status (measured as lower class, lower-middle class, middle class, upper-middle class, upper class); or military officer rank (measured as lieutenant, captain, major, lieutenant colonel, colonel, general). Such rankings are not absolute but rather “relative” to each other: Major is higher than captain, but we cannot measure the exact difference in numerical terms. Is the difference between major and captain equal to the difference between colonel and general? You cannot say.
- **Interval:** Measures by using equal intervals. Here you can compare differences between pairs of values. The Fahrenheit temperature scale, measured in degrees, is an interval scale, as is the centigrade scale. The

temperature difference between 50 and 60 degrees centigrade (10 degrees) equals the temperature difference between 80 and 90 degrees centigrade (10 degrees). Note that the 0 in each of these scales is arbitrarily placed, which makes the interval scale different from ratio.

- **Ratio:** Similar to an interval scale, a ratio scale includes a 0 measurement that signifies the point at which the characteristic being measured vanishes (absolute 0). For example, income (measured in dollars, with 0 equal to no income at all), years of formal education, items sold, and so forth, are all ratio scales.

Chapter Checkout

Q&A

1. A doodad machine is capable of making up to 200 doodads per hour.
 - a. For four consecutive hours, the doodad machine makes 120, 118, 124, and 112 doodads. Find the average rate of doodad production.
 - b. During the fifth hour, the machine breaks after making only 21 doodads. The remainder of the hour is spent repairing it. Normally, the machine must average 115 doodads per hour, over the course of an 8 hour day, in order to meet quota. What must the average rate be for the next 3 hours to meet this quota?
2. A meteorologist measures the average rate of rainfall in order to test the saying, "When it rains, it pours." Her data, showing the rainfall rate over the course of a rainy morning, are as follows:

<i>Time of Day</i>	<i>Rainfall Rate (inches/min)</i>
10:00 - 10:15 a.m.	0.011
10:15 - 10:30 a.m.	0.007
10:30 - 10:45 a.m.	0.001
10:45 - 11:00 a.m.	0.010
11:00 - 11:15 a.m.	0.003
11:15 - 11:30 a.m.	0.045
11:30 - 11:45 a.m.	0.002
11:45 - 12:00 p.m.	0.006

- a. Find the median and mean.
 - b. Calculate the variance and standard deviation.
3. Can a distribution have the mean equal to the median, but the mode not equal to either mean or median? If so, draw it.

Answers: 1a. 118.5 b. 141.67 2a. 0.006, 0.0106 b. 0.0002, 0.014 3. Yes (a symmetric but bimodal distribution is a good example).

Chapter 4

PROBABILITY

Chapter Check-In

- ❑ Applying classic probability theory to simple events
- ❑ Learning how to analyze combinations of simple events based on these rules
- ❑ Learning about probability distributions, especially the binomial distribution

Probability theory plays a central role in statistics. After all, statistical analysis is applied to a collection of data in order to discover something about the underlying events. These events may be connected to one another—for example, mutually exclusive—but the individual choices involved are assumed to be random. Alternatively, we may sample a population at random and make inferences about the population as a whole from the sample by using statistical analysis. Therefore, a solid understanding of probability theory—the study of random events—is necessary to understand how the statistical analysis works and also to correctly interpret the results.

You have an intuition about probability. As you will see, in some cases, probability theory seems obvious. But be careful, for on occasion a seemingly obvious answer will turn out to be wrong—because sometimes your intuition about probability will fail. Even in seemingly simple cases, it is best to follow the rules of probability as described in this chapter rather than rely on your hunches.

Classic Theory

The **classic theory of probability** underlies much of probability in statistics. Briefly, this theory states that the chance of a particular outcome

occurring is determined by the ratio of the number of favorable outcomes (or “successes”) to the total number of outcomes. Expressed as a formula,

$$P(A) = \frac{\text{number of favorable outcomes}}{\text{total number of possible outcomes}}$$

For example, the probability of randomly drawing an ace from a well-shuffled deck of cards is equal to the ratio $4/52$. Four is the number of favorable outcomes (the number of aces in the deck), and 52 is the number of total outcomes (the number of cards in the deck). The probability of randomly selecting an ace in one draw from a deck of cards is therefore $4/52$, or .077. In statistical analysis, probability is usually expressed as a decimal and ranges from a low of 0 (no chance) to a high of 1.0 (certainty).

The classic theory assumes that all outcomes have equal likelihood of occurring. In the example just cited, each card must have an equal chance of being chosen—no card is larger than any other or in any way more likely to be chosen than any other card.

The classic theory pertains only to outcomes that are **mutually exclusive** (or **disjoint**), which means that those outcomes may not occur at the same time. For example, one coin flip can result in a head or a tail, but one coin flip cannot result in a head and a tail. So the outcome of a head and the outcome of a tail are said to be “mutually exclusive” in one coin flip, as is the outcome of an ace and a king as the outcome of one card being drawn.

Relative Frequency Theory

The **relative frequency theory of probability** holds that if an experiment is repeated an extremely large number of times and a particular outcome occurs a percentage of the time, then that particular percentage is close to the probability of that outcome.

For example, if a machine produces 10,000 widgets one at a time, and 1,000 of those widgets are faulty, the probability of that machine producing a faulty widget is approximately 1,000 out of 10,000, or .10.

Probability of Simple Events

Example 1: What is the probability of simultaneously flipping three coins—a penny, a nickel, and a dime—and having all three land heads?

Using the classic theory, determine the ratio of number of favorable outcomes to the number of total outcomes. Table 4-1 lists all possible outcomes.

Table 4-1 Possible Outcomes of Penny, Nickel, and Dime Flipping

<i>Outcome #</i>	<i>Penny</i>	<i>Nickel</i>	<i>Dime</i>
1	H	H	H
2	H	H	T
3	H	T	H
4	H	T	T
5	T	H	H
6	T	H	T
7	T	T	H
8	T	T	T

There are 8 different outcomes, only one of which is favorable (outcome #1: all three coins landing heads); therefore, the probability of three coins landing heads is $1/8$, or .125.

What is the probability of exactly two of the three coins landing heads? Again, there are the 8 total outcomes, but in this case only 3 favorable outcomes (outcomes #2, #3, and #5); thus, the probability of exactly two of three coins landing heads is $3/8$ or .375.

Independent Events

Each of the three coins being flipped in the preceding example is what is known as an **independent event**. Independent events are defined as outcomes that are not affected by other outcomes. In other words, the flip of the penny does not affect the flip of the nickel, and vice versa.

Dependent Events

Dependent events, on the other hand, are outcomes that are affected by other outcomes. Consider the following example.

Example 2: What is the probability of randomly drawing an ace from a deck of cards and then drawing an ace again from the same deck of cards, without returning the first drawn card back to the deck?

For the first draw, the probability of a favorable outcome is $4/52$, as explained earlier; however, once that first card has been drawn, the total number of outcomes is no longer 52, but now 51 because a card has been removed from the deck. And if that first card drawn resulted in a favorable outcome (an ace), there would now be only 3 aces in the deck. Or the number of favorable outcomes would remain at 4 if that first card drawn were not an ace. So the second draw is a dependent event because its probability changes depending upon what happens on the first draw.

If, however, you replace that drawn card back into the deck and shuffle well again before the second draw, then the probability for a favorable outcome for each draw will now be equal ($4/52$), and these events will be independent.

Probability of Joint Occurrences

Another way to compute the probability of all three flipped coins landing heads is as a series of three different events: first flip the penny, then flip the nickel, then flip the dime. Will the probability of landing three heads still be .125?

Multiplication rule

To compute the probability of two or more independent events all occurring—**joint occurrence**—multiply their probabilities.

For example, the probability of the penny landing heads is $1/2$, or .5; the probability of the nickel next landing heads is $1/2$, or .5; and the probability of the dime landing heads is $1/2$, or .5; thus, note that

$$.5 \times .5 \times .5 = .125$$

which is what you determined with the classic theory by assessing the ratio of number of favorable outcomes to number of total outcomes. The notation for joint occurrence is $P(AB) = P(A) \times P(B)$ and reads: The probability of A and B both happening is equal to the probability of A times the probability of B.

Using the **multiplication rule**, you can also determine the probability of drawing two aces in a row from a deck of cards. The only way to draw two aces in a row from a deck of cards is for both draws to be favorable. For the first draw, the probability of a favorable outcome is $4/52$. But because

the first draw is favorable, only 3 aces are left among 51 cards. So the probability of a favorable outcome on the second draw is $3/51$. For both events to happen, you simply multiply those two probabilities together:

$$\frac{4}{52} \times \frac{3}{51} = \frac{12}{2652} = .0045$$

Note that these probabilities are not independent. If, however, you had decided to return the initial card drawn back to the deck before the second draw, then the probability of drawing an ace on each draw is $4/52$, as these events are now independent. Drawing an ace twice in a row, with the odds being $4/52$ both times, gives the following:

$$\frac{4}{52} \times \frac{4}{52} = \frac{16}{2704} = .0059$$

In either case, you use the multiplication rule because you are computing probability for favorable outcomes in all events.

Addition rule

Given mutually exclusive events, finding the probability of *at least one* of them occurring is accomplished by adding their probabilities.

For example, what is the probability of one coin flip resulting in at least one head or at least one tail?

The probability of one coin flip landing heads is $.5$, and the probability of one coin flip landing tails is $.5$. Are these two outcomes mutually exclusive in one coin flip? Yes, they are. You cannot have a coin land both heads and tails in one coin flip; therefore, you can determine the probability of at least one head or one tail resulting from one flip by adding the two probabilities:

$$.5 + .5 = 1.0 \text{ (or certainty)}$$

Example 3: What is the probability of at least one spade or one club being randomly chosen in one draw from a deck of cards?

The probability of drawing a spade in one draw is $13/52$; the probability of drawing a club in one draw is $13/52$. These two outcomes are mutually exclusive in one draw because you cannot draw both a spade and a club in one draw; therefore, you can use the **addition rule** to determine the probability of drawing at least one spade or one club in one draw:

$$\frac{13}{52} + \frac{13}{52} = \frac{26}{52} = .50$$

Non-Mutually Exclusive Outcomes

For the addition rule to apply, the events must be mutually exclusive. Now consider the following example.

Example 4: What is the probability of the outcome of at least one head in two coin flips?

Should you add the two probabilities as in the preceding examples? In the first example, you added the probability of getting a head and the probability of getting a tail because those two events were mutually exclusive in one flip. In the second example, the probability of getting a spade was added to the probability of getting a club because those two outcomes were mutually exclusive in one draw. Now when you have two flips, should you add the probability of getting a head on the first flip to the probability of getting a head on the second flip? Are these two events mutually exclusive?

Of course, they are not mutually exclusive. You can get an outcome of a head on one flip and a head on the second flip. So, because they are not mutually exclusive, you cannot use the addition rule. If you did use the addition rule, you would get

$$\frac{1}{2} + \frac{1}{2} = \frac{2}{2} = 1.0$$

or certainty, which is absurd. There is no certainty of getting at least one head on two flips. (Try it several times, and see that there is a possibility of getting two tails and no heads.)

Double Counting

By using the addition rule in a situation that is not mutually exclusive, you are in fact **double counting**. One way of realizing this double counting is to use the classic theory of probability: List all the different outcomes when flipping a coin twice and assess the ratio of favorable outcomes to total outcomes (see Table 4-2).

Table 4-2 All Possible Outcomes of Flipping the Same Coin Twice

<i>First Flip</i>	<i>with</i>	<i>Second Flip</i>
head	+	head
head	+	tail

First Flip	with	Second Flip
tail	+	head
tail	+	tail

There are 4 total outcomes. Three of the outcomes have at least one head; therefore, the probability of throwing at least one head in two flips is $3/4$ or $.75$, not 1.0 . But if you had used the addition rule, you would have added the two heads from the first flip to the two heads from the second flip and gotten 4 heads in 4 flips, $4/4 = 1.0$! But the two heads in that first pair constitute only one outcome, and so by counting both heads for that outcome, you are double counting because this is the joint-occurrence outcome that is not mutually exclusive.

To use the addition rule in a non-mutually exclusive situation, you must subtract any events that double count. In this case:

$$\frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4} = .75$$

The notation, therefore, for at least one favorable occurrence in two events is

$$P(A + B) = P(A) + P(B) - P(AB)$$

This rule is read: The probability of at least one of the events A or B equals the probability of A plus the probability of B minus the probability of their joint occurrence. (Note that if they are mutually exclusive, then $P(AB)$ —the joint occurrence—equals 0, and you simply add the two probabilities.)

Example 5: What is the probability of drawing either a spade or an ace from a deck of cards?

The probability of drawing a spade is $13/52$; the probability of drawing an ace is $4/52$. But the probability of their joint occurrence (an ace of spades) is $1/52$. Thus,

$$\begin{aligned} P(A + B) &= P(A) + P(B) - P(AB) \\ &= \frac{13}{52} + \frac{4}{52} - \frac{1}{52} \\ &= \frac{16}{52} = \frac{4}{13} = .3077 \end{aligned}$$

Conditional Probability

Sometimes you have more information than simply total outcomes and favorable outcomes and, hence, are able to make more informed judgments regarding probabilities. For example, suppose you know the following information: In a particular village, there are 60 women and 40 men. Twenty of those women are 70 years of age or older; 5 of the men are 70 years of age or older. See Table 4-3.

Table 4-3 Distribution of People in a Particular Village

	70+ years	69 or less	Totals
Women	20	40	60
Men	5	35	40
Totals	25	75	100

What is the probability that a person selected at random in that town will be a woman? Because women constitute 60 percent of the total population, the probability is .60.

What is the probability that a person 70+ years of age selected at random will be a woman? This question is different because the probability of A (being a woman) given B (the person in question is 70+ years of age) is now conditional upon B (being 70+ years of age). Because women number 20 out of the 25 people in the 70+ years-old group, the probability of this latter question is $20/25$, or .80.

Conditional probability is found using this formula:

$$P(A|B) = \frac{AB}{B} = \frac{P(AB)}{P(B)}$$

which is read: The probability of A given B equals the proportion of the total of A and B to the total of B . The vertical bar in the expression $A|B$ is read *given that* or *given*.

Probability Distributions

A **probability distribution** is a pictorial display of the probability— $P(x)$ —for any value of x . Consider the number of possible outcomes of two coins

being flipped (see Table 4-4). Table 4-5 shows the probability distribution of the results of flipping two coins. Figure 4-1 displays this information graphically.

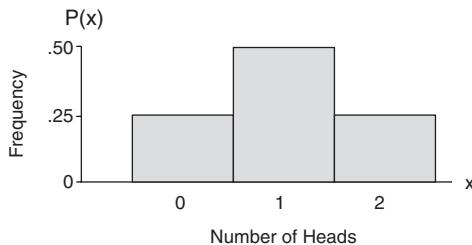
Table 4-4 Possible Outcomes of Two Flipped Coins

H + H	=	2 heads
H + T	=	1 head
T + H	=	1 head
T + T	=	0 heads

Table 4-5 Probability Distribution: Number of Heads

x	$P(x)$
0	1/4 or .25
1	1/2 or .50
2	1/4 or .25

Figure 4-1 Probability distribution of the results of flipping two coins.



Discrete vs. continuous variables

The number of heads resulting from coin flips can be counted only in integers (whole numbers). The number of aces drawn from a deck can be counted only in integers. These “countable” numbers are known as **discrete variables**: 0, 1, 2, 3, 4, and so forth. Nothing in between two variables is possible. For example, 2.6 is not possible.

Qualities such as height, weight, temperature, distance, and so forth, however, can be measured using fractions or decimals as well: 34.27, 102.26, and so forth. These are known as **continuous variables**.

Total of probabilities of discrete variables

As you have noted, the probability of a discrete variable lies somewhere between 0 and 1, inclusive. For example, the probability of tossing one head in two coin flips is, as earlier, .50. The probability of tossing two heads in two coin flips is .25, and the probability of tossing no heads in two coin flips is .25. The sum (total) of probabilities for all values of x always equals 1. For example, note that in Tables 4-3 and 4-5, adding the probabilities for all the possible outcomes yields a sum of 1.

The Binomial

A discrete variable that can result in only one of two outcomes is called **binomial**. For example, a coin flip is a binomial variable; but drawing a card from a standard deck of 52 is not. Whether a drug is either successful or unsuccessful in producing results is a binomial variable, as is whether a machine produces perfect or imperfect widgets.

Binomial experiments

Binomial experiments require the following elements:

- The experiment consists of a number of identical events (n).
- Each event has only one of two mutually exclusive outcomes. (These outcomes are called successes and failures.)
- The probability of a success outcome is equal to some percentage, which is identified as a **proportion**, π .
- This proportion, π , remains constant throughout all events and is defined as the ratio of number of successes to number of trials.
- The events are independent.
- Given all of the above, the **binomial** formula can be applied (x = number of favorable outcomes; n = number of events):

$$P(x) = \frac{n!}{x!(n-x)!} \pi^x (1-\pi)^{n-x}$$
$$n! = n(n-1)(n-2) \dots (3)(2)(1)$$

Example 6: A coin is flipped 10 times. What is the probability of getting exactly 5 heads? Using the binomial formula, where n (the number of events) is given as 10; x (the number of favorable outcomes) is given as 5; and the probability of landing a head in one flip is .5:

$$\begin{aligned}
 P(x) &= \frac{n!}{x!(n-x)!} \pi^x (1-\pi)^{n-x} \\
 &= \frac{10!}{5!(5)!} (.5)^5 (1-.5)^5 \\
 &= \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 (5 \cdot 4 \cdot 3 \cdot 2 \cdot 1)} (.0313)(.0313) \\
 &= 252 (.0313)(.0313) \\
 &= .246
 \end{aligned}$$

So the probability of getting exactly 5 heads in 10 flips is .246, or approximately 25 percent.

Binomial table

Because probabilities of binomial variables are so common in statistics, tables are used to alleviate having to continually use the formula. Refer to Table 1 in Appendix B, and you will find that given $n = 10$, $x = 5$, and $\pi = .5$, the probability is .2461.

Mean and standard deviation

The mean of the binomial probability distribution is determined by the following formula:

$$\mu = n\pi$$

where π is the proportion of favorable outcomes and n is the number of events.

The **standard deviation of the binomial probability distribution** is determined by this formula:

$$\sigma = \sqrt{n\pi(1-\pi)}$$

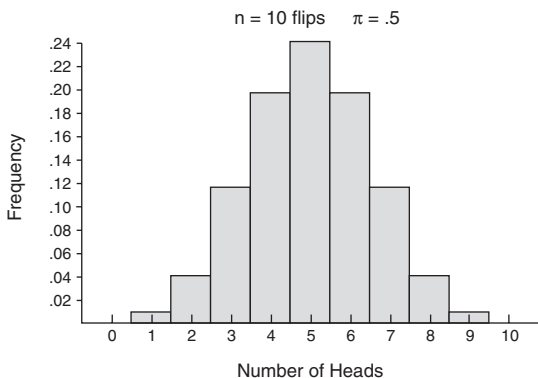
Example 7: What is the mean and standard deviation for a binomial probability distribution for 10 coin flips of a fair coin?

Because the proportion of favorable outcomes of a fair coin falling heads (or tails) is $\pi = .5$, simply substitute into the formulas:

$$\mu = n\pi = 10(.5) = 5$$
$$\sigma = \sqrt{n\pi(1 - \pi)} = \sqrt{10(.5)(.5)} = \sqrt{2.5} = 1.58$$

The probability distribution for the number of favorable outcomes is shown in Figure 4-2.

Figure 4-2 The binomial probability distribution of the number of heads resulting from 10 coin tosses.



Note that this distribution appears to display symmetry. Binomial probability distributions and use of the normal approximation to replace the probability histogram are discussed in Chapter 5.

Chapter Checkout

Q&A

1. A dart board is divided into twenty equal wedges, ignoring the bull's eye.
 - a. If only six of the twenty wedges are worth points, what is the probability of scoring on one throw (assuming you can hit the board)?
 - b. What is the chance of hitting at least one scoring region in three consecutive throws?

2. During a boring lecture in statistics class, you begin repeatedly flipping an old, worn quarter. Amazingly, the first ten flips are all heads.
 - a. What is the chance of this, if the coin is fair?
 - b. Suppose that the coin is not fair. If the probability of heads is 75 percent rather than 50 percent, what is the chance of your result?
 - c. How unfair would the coin have to be in order for your result to have had a 50 percent chance of occurring?
3. A woman has three sons. What is the probability that her next child will be a girl?
4. Consider an ordinary (six-sided) die, assumed to be fair.
 - a. Use the binomial distribution to determine the probability that, in 10 rolls of the die, you will see exactly 4 sixes.
 - b. If you roll the die 100 times, what is the chance of rolling exactly 40 sixes?

Answers: 1a. 0.3 b. 0.657 2a. 9.77×10^{-4} b. 0.0563 c. $p(\text{heads}) = 0.933$
3. 0.5 4a. 0.0543 b. 1.82×10^{-8}

Chapter 5

SAMPLING

Chapter Check-In

- ❑ Examining the differences between populations and samples
- ❑ Learning about sampling distributions, sampling errors, and the Central Limit Theorem
- ❑ Discussing in detail the properties of the normal distribution

As discussed in Chapter 1, the field of inferential statistics enables you to make educated guesses about the numerical characteristics of large groups. The logic of sampling gives you a way to test conclusions about such groups using only a small portion of its members.

Populations, Samples, Parameters, and Statistics

A **population** is a group of phenomena that have something in common. The term often refers to a group of people, as in the following examples:

All registered voters in Crawford County

All members of the International Machinists Union

All Americans who played golf at least once in the past year

But populations can refer to things as well as people:

All widgets produced last Tuesday by the Acme Widget Company

All daily maximum temperatures in July for major U.S. cities

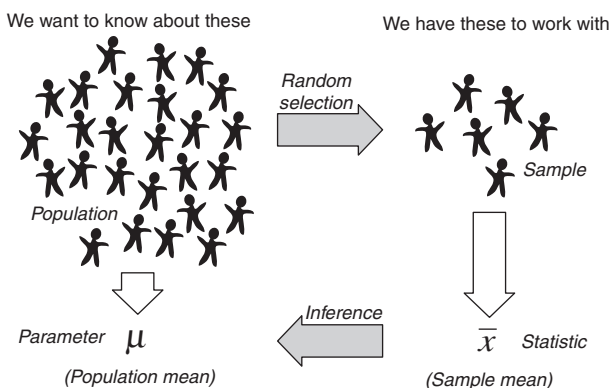
All basal ganglia cells from a particular rhesus monkey

Often, researchers want to know things about populations but don't have data for every person or thing in the population. If a company's customer service division wanted to learn whether its customers were satisfied, it would not be practical (or perhaps even possible) to contact every individual who purchased a product. Instead, the company might select a **sample** of the

population. A sample is a smaller group of members of a population selected to represent the population. In order to use statistics to learn things about the population, the sample must be **random**. A random sample is one in which every member of a population has an equal chance to be selected.

A **parameter** is a characteristic of a population. A **statistic** is a characteristic of a sample. Inferential statistics enables you to make an educated guess about a population parameter based on a statistic computed from a sample randomly drawn from that population. See Figure 5-1.

Figure 5-1 Illustration of the relationship between samples and populations.



For example, say you want to know the mean income of the subscribers to a particular magazine—a parameter of a population. You draw a random sample of 100 subscribers and determine that their mean income \bar{x} is \$27,500 (a statistic). You conclude that the population mean income μ is likely to be close to \$27,500 as well. This example is one of statistical inference.

Different symbols are used to denote statistics and parameters, as Table 5-1 shows.

Table 5-1 Comparison of Sample Statistics and Population Parameters

	<i>Sample Statistic</i>	<i>Population Parameter</i>
Mean	\bar{x}	μ
Standard deviation	s, sd	σ
Variance	s^2	σ^2

Sampling Distributions

Continuing with the earlier example, suppose that ten different samples of 100 people were drawn from the population, instead of just one. The income means of these ten samples would not be expected to be exactly the same, because of **sampling variability**. Sampling variability is the tendency of the same statistic computed from a number of random samples drawn from the same population to differ.

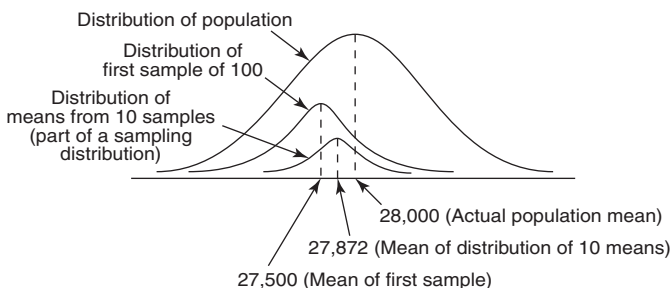
Suppose that the first sample of 100 magazine subscribers was “returned” to the population (made available to be selected again), another sample of 100 subscribers selected at random, and the mean income of the new sample computed. If this process were repeated ten times, it might yield the following sample means:

27,500	27,192	28,736	26,454	28,527
28,407	27,592	27,684	28,827	27,809

These ten values are part of a **sampling distribution**. The sampling distribution of a statistic—in this case, of a mean—is the distribution obtained by computing the statistic for a large number of samples drawn from the same population.

You can compute the sample mean of this sampling distribution by summing the ten sample means and dividing by ten, which gives a distribution mean of 27,872. Suppose that the mean income of the entire population of subscribers to the magazine is \$28,000. (You usually do not know what it is.) You can see in Figure 5-2 that the first sample mean (\$27,500) was not a bad estimate of the population mean and that the mean of the distribution of ten sample means (\$27,872) was even better.

Figure 5-2 Estimation of the population mean becomes progressively more accurate as more samples are taken.



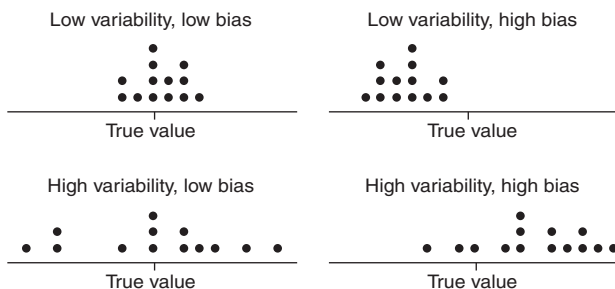
The more sample means included in the sampling distribution, the more accurate the mean of the sampling distribution becomes as an estimate of the population mean.

Random and Systematic Error

Two potential sources of error occur in statistical estimation—two reasons a statistic might misrepresent a parameter. **Random error** occurs as a result of sampling variability. The ten sample means in the preceding section differed from the true population mean because of random error. Some were below the true value; some above it. Similarly, the mean of the distribution of ten sample means was slightly lower than the true population mean. If ten more samples of 100 subscribers were drawn, the mean of that distribution—that is, the mean of those means—might be higher than the population mean.

Systematic error or **bias** refers to the tendency to consistently underestimate or overestimate a true value. Suppose that your list of magazine subscribers was obtained through a database of information about air travelers. The samples that you would draw from such a list would likely overestimate the population mean of all subscribers' income because lower-income subscribers are less likely to travel by air and many of them would be unavailable to be selected for the samples. This example would be one of bias. See Figure 5-3.

Figure 5-3 Random (sampling) error and systematic error (bias) distort the estimation of population parameters from sample statistics.



Central Limit Theorem

If the population of all subscribers to the magazine were normal, you would expect its sampling distribution of means to be normal as well. But what if the population were non-normal? The **Central Limit Theorem** states that even if a population distribution is strongly non-normal, its sampling distribution of means will be approximately normal for large sample sizes (over 30). The Central Limit Theorem makes it possible to use probabilities associated with the normal curve to answer questions about the means of sufficiently large samples.

According to the Central Limit Theorem, the mean of a sampling distribution of means is an unbiased estimator of the population mean.

$$\mu = \mu$$

Similarly, the standard deviation of a sampling distribution of means is

$$\sigma = \frac{\sigma}{\sqrt{n}}$$

Note that the larger the sample, the less variable the sample mean. The mean of many observations is less variable than the mean of few. The standard deviation of a sampling distribution of means is often called the **standard error** of the mean. Every statistic has a standard error, which is a measure of the statistic's random variability.

Example 1: If the population mean of number of fish caught per trip to a particular fishing hole is 3.2 and the population standard deviation is 1.8, what are the population mean and standard deviation of 40 trips?

$$\begin{aligned}\mu &= 3.2 \\ \sigma &= \frac{1.8}{\sqrt{40}} = .285\end{aligned}$$

Properties of the Normal Curve

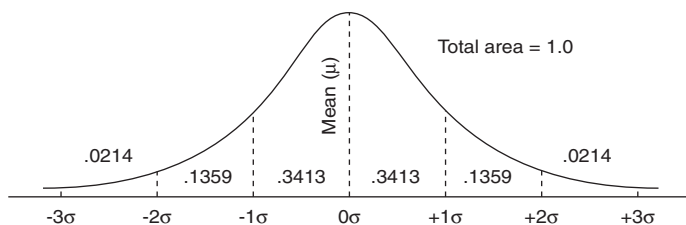
Known characteristics of the normal curve make it possible to estimate the probability of occurrence of any value of a normally distributed variable. Suppose that the total area under the curve is defined to be 1. You can multiply that number by 100 and say there is a 100 percent chance that

any value you can name will be somewhere in the distribution. (Remember, the distribution extends to infinity in both directions.) Similarly, because half of the area of the curve is below the mean and half is above it, you can say that there's a 50 percent chance that a randomly chosen value will be above the mean and the same chance that it will be below it.

It makes sense that the area under the normal curve is equivalent to the probability of randomly drawing a value in that range. The area is greatest in the middle, where the “hump” is, and thins out toward the tails. That's consistent with the fact that there are more values close to the mean in a normal distribution than far from it.

When the area of the standard normal curve is divided into sections by standard deviations above and below the mean, the area in each section is a known quantity (see Figure 5-4). As explained earlier, the area in each section is the same as the probability of randomly drawing a value in that range.

Figure 5-4 The normal curve and the area under the curve between σ units.

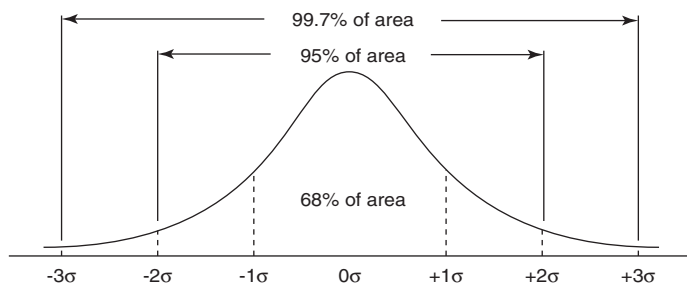


For example, $.3413$ of the curve falls between the mean and one standard deviation above the mean, which means that about 34 percent of all the values of a normally distributed variable are between the mean and one standard deviation above it. It also means that there is a $.3413$ chance that a value drawn at random from the distribution will lie between these two points.

Sections of the curve above and below the mean may be added together to find the probability of obtaining a value within (plus or minus) a given number of standard deviations of the mean (see Figure 5-5). For example, the amount of curve area between one standard deviation above the mean and one standard deviation below is $.3413 + .3413 = .6826$, which means

that approximately 68.26 percent of the values lie in that range. Similarly, about 95 percent of the values lie within two standard deviations of the mean, and 99.7 percent of the values lie within three standard deviations.

Figure 5-5 The normal curve and the area under the curve between σ units.



In order to use the area of the normal curve to determine the probability of occurrence of a given value, the value must first be **standardized**, or converted to a ***z*-score**. To convert a value to a *z*-score is to express it in terms of how many standard deviations it is above or below the mean. After the *z*-score is obtained, its corresponding probability may be looked up in a table. The formula to compute a *z*-score is

$$z = \frac{x - \mu}{\sigma}$$

where x is the value to be converted, μ is the population mean, and σ is the population standard deviation.

Example 2: A normal distribution of retail store purchases has a mean of \$14.31 and a standard deviation of 6.40. What percentage of purchases were under \$10? First, compute the *z*-score:

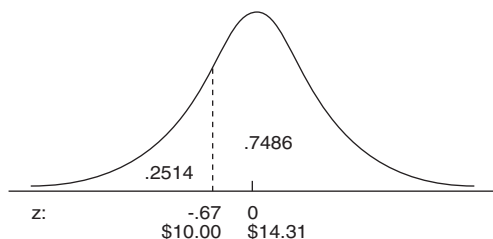
$$z = \frac{10 - 14.31}{6.40} = -.67$$

The next step is to look up the z -score in the table of standard normal probabilities (see Table 2 in Appendix B). The standard normal table lists the probabilities (curve areas) associated with given z -scores.

Table 2 from Appendix B gives the area of the curve below z or, in other words, the probability of obtaining a value of z or lower. Not all standard normal tables use the same format, however. Some list only positive z -scores and give the area of the curve between the mean and z . Such a table is slightly more difficult to use, but the fact that the normal curve is symmetric makes it possible to use it to determine the probability associated with any z -score, and vice versa.

To use Table 2 (the table of standard normal probabilities) in Appendix B, first look up the z -score in the left column, which lists z to the first decimal place. Then look along the top row for the second decimal place. The intersection of the row and column is the probability. In the example, you first find -0.6 in the left column and then $.07$ in the top row. Their intersection is $.2514$. The answer, then, is that about 25 percent of the purchases were under \$10 (see Figure 5-6).

Figure 5-6 Finding a probability using a z -score on the normal curve.



What if you had wanted to know the percentage of purchases above a certain amount? Because Table 2 gives the area of the curve below a given z , to obtain the area of the curve above z , simply subtract the tabled probability from 1. The area of the curve above a z of -0.67 is $1 - .2514 = .7486$. Approximately 75% of the purchases were above \$10.

Just as Table 2 can be used to obtain probabilities from z -scores, it can be used to do the reverse.

Example 3: Using the previous example, what purchase amount marks the lower 10 percent of the distribution?

Locate in Table 2 the probability of .1000, or as close as you can find, and read off the corresponding z -score. The figure that you seek lies between the tabled probabilities of .0985 and .1003, which correspond to z -scores of -1.29 and -1.28 . To be conservative, use the more extreme value of $z = -1.29$. Now, use the z formula, this time solving for x :

$$\begin{aligned}\frac{x - 14.31}{6.4} &= -1.29 \\ x - 14.31 &= (-1.29)(6.4) \\ x &= -8.256 + 14.31 = 6.054\end{aligned}$$

Approximately 10 percent of the purchases were below \$6.05.

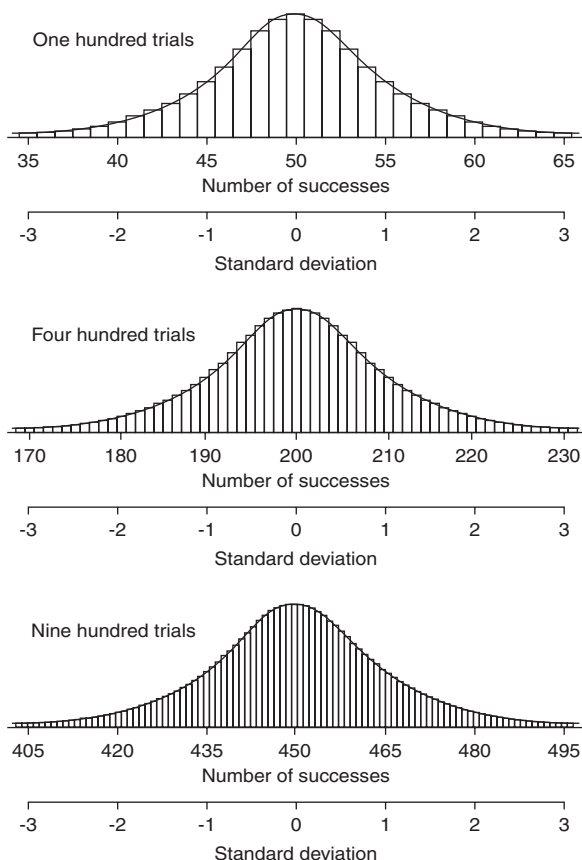
Normal Approximation to the Binomial

In Chapter 4, which discusses probability, you saw that some variables are continuous, that there is no limit to the number of times you could divide their intervals into still smaller ones, although you may round them off for convenience. Examples include age, height, and cholesterol level. Other variables are discrete, or made of whole units with no values between them. Some discrete variables are the number of children in a family, the sizes of televisions available for purchase, or the number of medals awarded at the Olympic Games.

Chapter 4 also introduces one kind of discrete variable, the binomial variable. A binomial variable can take only two values, often termed *successes* and *failures*. Examples include coin flips that come up either heads or tails, manufactured parts that either continue working past a certain point or do not, and basketball tosses that either fall through the hoop or do not.

You discovered that the outcomes of binomial trials have a frequency distribution, just as continuous variables do. The more binomial outcomes there are (for example, the more coins you flip simultaneously), the more closely the sampling distribution resembles a normal curve (see Figure 5-7). You can take advantage of this fact and use the table of standard normal probabilities (Table 2 in Appendix B) to estimate the likelihood of obtaining a given proportion of successes. You can do this by converting the test proportion to a z -score and looking up its probability in the standard normal table.

Figure 5-7 As the number of trials increases, the binomial distribution approaches the normal distribution.



The mean of the normal approximation to the binomial is

$$\mu = n\pi$$

and the standard deviation is

$$\sigma = \sqrt{n\pi(1 - \pi)}$$

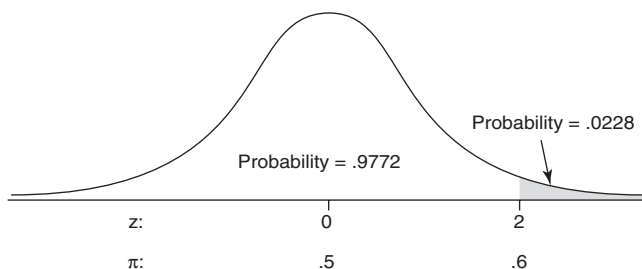
where n is the number of trials and π is the probability of success. The approximation will be more accurate the larger the n and the closer the proportion of successes in the population to .5.

Example 4: Assuming an equal chance of a new baby being a boy or a girl (that is, $\pi = 0.5$), what is the likelihood that 60 or more out of the next 100 births at a local hospital will be boys?

$$\begin{aligned} z &= \frac{x - \mu}{\sigma} = \frac{60 - ((100)(.5))}{\sqrt{(100)(.5)(1 - .5)}} \\ &= \frac{10}{5} = 2 \end{aligned}$$

According to Table 2, a z -score of 2 corresponds to a probability of .9772. As you can see in Figure 5-8, there is a .9772 chance that there will be 60 percent or fewer boys, which means that the probability that there will be more than 60 percent boys is $1 - .9772 = .0228$, or just over 2 percent. If the assumption that the chance of a new baby being a girl is the same as it being a boy is correct, the probability of obtaining 60 or fewer girls in the next 100 births is also .9772.

Figure 5-8 Finding a probability using a z -score on the normal curve.



Chapter Checkout

Q&A

1. Consider a fair, six-sided die. Using the normal approximation to the binomial distribution, determine the probability that in 10 successive rolls, there will be *exactly* 4 sixes.

2. The speed limit on the street in front of your house is 40 mph. In order to measure compliance with the limit, you decide to measure the mean speed of vehicles on the street. Each day after work, you measure the speed of 20 cars and take their average. Your data for five days is shown in the table below.

<i>Day</i>	<i>Average Speed (mph)</i>
Monday	38.3
Tuesday	40.1
Wednesday	34.2
Thursday	45.1
Friday	50.7

Using this information answer the following questions:

- Find the mean and standard deviation of these means. Is it likely that drivers comply, on average, with the law?
- Assuming the population mean of drivers does comply, what is the probability of your results? Take $\sigma = 7.5$ mph. Assume a normal distribution of speeds about μ .

Answers: 1. 0.0518 2a. 41.68, 6.38, yes. b. 0.59

Critical Thinking

4. On TV, surveys are often taken during sensational news stories to determine the “people’s opinion.” Usually, viewers are told to dial a phone number in order to vote yes or no on the issue. These phone numbers sometimes require the caller to pay a fee. Comment on this practice, in terms of sampling practice. Does it produce a random sample? If not, how could it be modified to produce a more random sample? Who is most likely to call these numbers?

Chapter 6

PRINCIPLES OF TESTING

Chapter Check-In

- ❑ Stating hypotheses and corresponding null hypotheses
- ❑ Learning how to apply simple statistical tests to determine if a null hypothesis is supported
- ❑ Understanding the types of statistical errors and estimating them
- ❑ Calculating test statistics, confidence intervals, and significance

Most people assume that the whole point of statistical analysis is to prove, or disprove, a certain claim. This chapter introduces you to techniques for testing claims based on the available data. These techniques are simple yet surprisingly powerful.

Stating Hypotheses

One common use of statistics is the testing of scientific hypotheses. First, the investigator forms a **research hypothesis** that states an expectation to be tested. Then the investigator derives a statement that is the opposite of the research hypothesis. This statement is called the **null hypothesis** (in notation: H_0). It is the null hypothesis that is actually tested, not the research hypothesis. If the null hypothesis can be rejected, that is taken as evidence in favor of the research hypothesis (also called the **alternative hypothesis**, H_a in notation). Because individual tests are rarely conclusive, it is usually not said that the research hypothesis has been “proved,” only that it has been supported.

An example of a research hypothesis comparing two groups might be the following:

Fourth-graders in Elmwood School perform differently in math than fourth-graders in Lancaster School.

Or in notation: $H_a: \mu_1 \neq \mu_2$

or sometimes: $H_a: \mu_1 - \mu_2 \neq 0$

The null hypothesis would be:

Fourth-graders in Elmwood School perform the same in math as fourth-graders in Lancaster School.

In notation: $H_0: \mu_1 = \mu_2$

or alternatively: $H_0: \mu_1 - \mu_2 = 0$

Some research hypotheses are more specific than that, predicting not only a difference but a difference in a particular direction:

Fourth-graders in Elmwood School are *better* in math than fourth-graders in Lancaster School.

In notation: $H_a: \mu_1 > \mu_2$

or alternatively: $H_a: \mu_1 - \mu_2 > 0$

The accompanying null hypothesis must be equally specific so that all possibilities are covered:

Fourth-graders in Elmwood School are *worse* in math, *or equal to*, fourth-graders in Lancaster School.

In notation: $H_0: \mu_1 \leq \mu_2$

or alternatively: $H_0: \mu_1 - \mu_2 \leq 0$

The Test Statistic

Hypothesis testing involves the use of distributions of known area, like the normal distribution, to estimate the probability of obtaining a certain value as a result of chance. The researcher is usually betting that the probability will be low because that means it's likely that the test result was not a mere coincidence but occurred because the researcher's theory is correct. It could mean, for example, that it's probably not just bad luck but faulty packaging equipment that caused you to get a box of raisin cereal with only five raisins in it.

Only two outcomes of a hypothesis test are possible: either the null hypothesis is rejected, or it is not. You have seen that values from normally distributed populations can be converted to z -scores and their probabilities looked up in Table 2 in Appendix B (see Chapter 5 for a complete discussion). The z -score is one kind of **test statistic** that is used to determine the probability of obtaining a given value. In order to test hypotheses, you must decide in advance what number to use as a cutoff for whether the null hypothesis will be rejected or not. This number is sometimes called the **critical** or **tabled value** because it is looked up in a table. It represents the level of probability that you will use to test the hypothesis. If the computed test statistic has a smaller probability than that of the critical value, the null hypothesis will be rejected.

For example, suppose you want to test the theory that sunlight helps prevent depression. One hypothesis derived from this theory might be that hospital admission rates for depression in sunny regions of the country are lower than the national average. Suppose that you know the national annual admission rate for depression to be 17 per 10,000. You intend to take the mean of a sample of admission rates from hospitals in sunny parts of the country and compare it to the national average.

Your research hypothesis is:

The mean annual admission rate for depression from the hospitals in sunny areas is less than 17 per 10,000.

In notation: $H_a: \mu_1 < 17 \text{ per } 10,000$

The null hypothesis is:

The mean annual admission rate for depression from the hospitals in sunny areas is equal to or greater than 17 per 10,000.

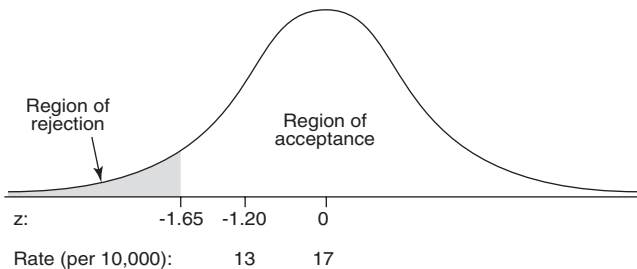
In notation: $H_0: \mu_1 \geq 17 \text{ per } 10,000$

Your next step is to choose a probability level for the test. You know that the sample mean must be lower than 17 per 10,000 in order to reject the null hypothesis, but how much lower? You settle on a probability level of 95 percent. That is, if the mean admission rate for the sample of sunny hospitals is so low that the chance of obtaining that rate from a sample selected at random from the national population is less than 5 percent, you'll reject the null hypothesis and conclude that there is evidence to support the hypothesis that exposure to the sun reduces the incidence of depression.

Next, you look up the critical z -score—the z -score that corresponds to your chosen level of probability—in the standard normal table. It is important to remember which end of the distribution you are concerned with. Table 2 in Appendix B lists the probability of obtaining a given z -score or lower. That is, it gives the area of the curve below the z -score. Because a computed test statistic in the lower end of the distribution will allow you to reject your null hypothesis, you look up the z -score for the probability (or area) of .05 and find that it is -1.65 . If you were hypothesizing that the mean in sunny parts of the country is greater than the national average, you would have been concerned with the upper end of the distribution instead and would have looked up the z -score associated with the probability (area) of .95, which is $z = 1.65$.

The critical z -score allows you to define the **region of acceptance** and the **region of rejection** of the curve (see Figure 6-1). If the computed test statistic is below the critical z -score, you can reject the null hypothesis and say that you have provided evidence in support of the alternative hypothesis. If it is above the critical value, you cannot reject the null hypothesis.

Figure 6-1 The z -score defines the boundary of the zones of rejection and acceptance.



Suppose that the mean admission rate for the sample of hospitals in sunny regions is 13 per 10,000 and suppose also that the corresponding z -score for that mean is -1.20 . The test statistic falls in the region of acceptance; so you cannot reject the null hypothesis that the mean in sunny parts of the country is significantly lower than the mean in the national average. There is a greater than a 5 percent chance of obtaining a mean admission

rate of 13 per 10,000 or lower from a sample of hospitals chosen at random from the national population, so you cannot conclude that your sample mean could not have come from that population.

One- and Two-tailed Tests

In the previous example, you tested a research hypothesis that predicted not only that the sample mean would be different from the population mean but that it would be different in a specific direction—it would be lower. This test is called a **directional** or **one-tailed test** because the region of rejection is entirely within one tail of the distribution.

Some hypotheses predict only that one value will be different from another, without additionally predicting which will be higher. The test of such a hypothesis is **non-directional** or **two-tailed** because an extreme test statistic in either tail of the distribution (positive or negative) will lead to the rejection of the null hypothesis of no difference.

Suppose that you suspect that a particular class's performance on a proficiency test is not representative of those people who have taken the test. The national mean score on the test is 74.

The research hypothesis is:

The mean score of the class on the test is not 74.

Or in notation: $H_a: \mu \neq 74$

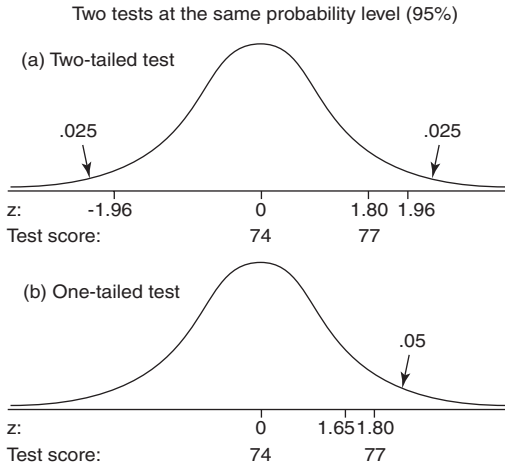
The null hypothesis is:

The mean score of the class on the test is 74.

In notation: $H_0: \mu = 74$

As in the last example, you decide to use a 95 percent probability level for the test. Both tests have a region of rejection, then, of five percent, or .05. In this example, however, the rejection region must be split between both tails of the distribution—.025 in the upper tail and .025 in the lower tail—because your hypothesis specifies only a difference, not a direction, as shown in Figure 6-2 (a). You will reject the null hypotheses of no difference if the class sample mean is either much higher or much lower than the population mean of 74. In the previous example, only a sample mean much lower than the population mean would have led to the rejection of the null hypothesis.

Figure 6-2 Comparison of (a) a two-tailed test and (b) a one-tailed test, at the same probability level (95%).



The decision of whether to use a one- or a two-tailed test is important because a test statistic that falls in the region of rejection in a one-tailed test may not do so in a two-tailed test, even though both tests use the same probability level. Suppose the class sample mean in your example was 77, and its corresponding z -score was computed to be 1.80. Table 2 in Appendix B shows the critical z -scores for a probability of .025 in either tail to be -1.96 and 1.96 . In order to reject the null hypothesis, the test statistic must be either smaller than -1.96 or greater than 1.96 . It is not, so you cannot reject the null hypothesis. Refer to Figure 6-2 (a).

Suppose, however, you had a reason to expect that the class would perform better on the proficiency test than the population, and you did a one-tailed test instead. For this test, the rejection region of .05 would be entirely within the upper tail. The critical z -value for a probability of .05 in the upper tail is 1.65. (Remember that Table 2 in Appendix B gives areas of the curve below z , so you look up the z -value for a probability of .95.) Your computed test statistic of $z = 1.80$ exceeds the critical value and falls in the region of rejection, so you reject the null hypothesis and say that your suspicion that the class was better than the population was supported. See Figure 6-2 (b).

In practice, you should use a one-tailed test only when you have good reason to expect that the difference will be in a particular direction. A two-tailed test is more conservative than a one-tailed test because a two-tailed test takes a more extreme test statistic to reject the null hypothesis.

Type I and II Errors

You have been using probability to decide whether a statistical test provides evidence for or against your predictions. If the likelihood of obtaining a given test statistic from the population is very small, you reject the null hypothesis and say that you have supported your hunch that the sample you are testing is different from the population.

But you could be wrong. Even if you choose a probability level of 95 percent, that means there is a 5 percent chance, or 1 in 20, that you rejected the null hypothesis when it was, in fact, correct. You can err in the opposite way, too; you might fail to reject the null hypothesis when it is, in fact, incorrect. These two errors are called Type I and Type II, respectively. Table 6-1 presents the four possible outcomes of any hypothesis test based on (1) whether the null hypothesis was accepted or rejected and (2) whether the null hypothesis was true in reality.

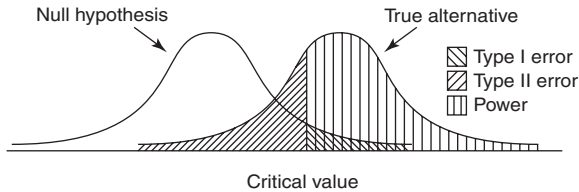
Table 6-1 Types of Statistical Errors

		<i>H₀ is actually:</i>	
		<i>True</i>	<i>False</i>
Test	reject H_0	Type I error	correct
Decision	accept H_0	correct	Type II error

A **Type I error** is often represented by the Greek letter alpha (α) and a Type II error by the Greek letter beta (β). In choosing a level of probability for a test, you are actually deciding how much you want to risk committing a Type I error—rejecting the null hypothesis when it is, in fact, true. For this reason, the area in the region of rejection is sometimes called the alpha level because it represents the likelihood of committing a Type I error.

In order to graphically depict a Type II, or β , error, it is necessary to imagine next to the distribution for the null hypothesis a second distribution for the true alternative (see Figure 6-3). If the alternative hypothesis is actually true, but you fail to reject the null hypothesis for all values of the test statistic falling to the left of the critical value, then the area of the curve of the alternative (true) hypothesis lying to the left of the critical value represents the percentage of times that you will have made a Type II error.

Figure 6-3 Graphical depiction of the relation between Type I and Type II errors, and the power of the test.



Type I and Type II errors are inversely related: As one increases, the other decreases. The Type I, or α (alpha), error rate is usually set in advance by the researcher. The Type II error rate for a given test is harder to know because it requires estimating the distribution of the alternative hypothesis, which is usually unknown.

A related concept is **power**—the probability that a test will reject the null hypothesis when it is, in fact, false. You can see from Figure 6-3 that power is simply 1 minus the Type II error rate (β). High power is desirable. Like β , power can be difficult to estimate accurately, but increasing the sample size always increases power.

Significance

How do you know how much confidence to put in the outcome of a hypothesis test? The statistician's criterion is the **statistical significance** of the test, or the likelihood of obtaining a given result by chance. This concept has already been spoken of, using several terms: probability, area of the curve, Type I error rate, and so forth. Another common representation of significance is the letter p (for probability) and a number between 0 and 1. There are several ways to refer to the significance level of a test, and it is important to be familiar with them. All of the following statements, for example, are equivalent:

The finding is significant at the .05 level.

The confidence level is 95 percent.

The Type I error rate is .05.

The alpha level is .05.

$\alpha = .05$.

There is a 95 percent certainty that the result is not due to chance.

There is a 1 in 20 chance of obtaining this result.

The area of the region of rejection is .05.

The p -value is .05.

$$p = .05.$$

The smaller the significance level p , the more stringent the test and the greater the likelihood that the conclusion is correct. The significance level is usually chosen in consideration of other factors that affect and are affected by it, like sample size, estimated size of the effect being tested, and consequences of making a mistake. Common significance levels are .05 (1 chance in 20), .01 (1 chance in 100), and .001 (1 chance in 1,000).

The result of a hypothesis test, as has been seen, is that the null hypothesis is either rejected or not. The significance level for the test is set in advance by the researcher in choosing a critical test value. When the computed test statistic is large (or small) enough to reject the null hypothesis, however, it is customary to report the observed (actual) p -value for the statistic.

If, for example, you intend to perform a one-tailed (lower tail) test using the standard normal distribution at $p = .05$, the test statistic will have to be smaller than the critical z -value of -1.65 in order to reject the null hypothesis. But suppose the computed z -score is -2.50 , which has an associated probability of .0062. The null hypothesis is rejected with room to spare. The observed significance level of the computed statistic is $p = .0062$; so you could report that the result was significant at $p < .01$. This result means that even if you had chosen the more stringent significance level of .01 in advance, you would still have rejected the null hypothesis, which is stronger support for your research hypothesis than rejecting the null hypothesis at $p = .05$.

It is important to realize that statistical significance and substantive, or practical, significance are not the same thing. A small, but important, real-world difference may fail to reach significance in a statistical test. Conversely, a statistically significant finding may have no practical consequence. This finding is especially important to remember when working with large sample sizes because any difference can be statistically significant if the samples are extremely large.

Point Estimates and Confidence Intervals

You have seen that the sample mean \bar{x} is an unbiased estimate of the population mean μ . Another way to say this is that \bar{x} is the best point estimate of the true value of μ . Some error is associated with this estimate, however—the true population mean may be larger or smaller than the sample mean.

Instead of a point estimate, you might want to identify a range of possible values p might take, controlling the probability that μ is not lower than the lowest value in this range and not higher than the highest value. Such a range is called a **confidence interval**.

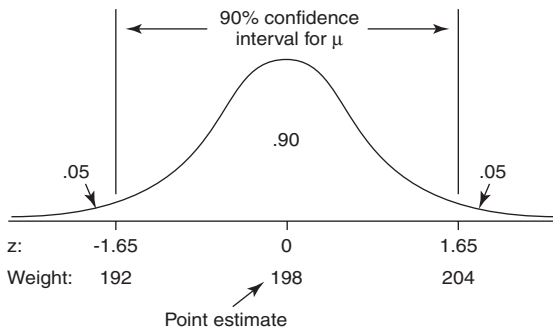
Example 1: Suppose that you want to find out the average weight of all players on the football team at Landers College. You are able to select ten players at random and weigh them. The mean weight of the sample of players is 198, so that number is your point estimate. The population standard deviation is $\sigma = 11.50$. What is a 90 percent confidence interval for the population weight, if you presume the players' weights are normally distributed?

This question is the same as asking what weight values correspond to the upper and lower limits of an area of 90 percent in the center of the distribution. You can define that area by looking up in Table 2 (in Appendix B) the z -scores that correspond to probabilities of .05 in either end of the distribution. They are -1.65 and 1.65 . You can determine the weights that correspond to these z -scores using the following formula:

$$(a, b) = \bar{x} \pm z \cdot \frac{\sigma}{\sqrt{n}}$$

The weight values for the lower and upper ends of the confidence interval are 192 and 204 (see Figure 6-4). A confidence interval is usually expressed by two values enclosed by parentheses, as in $(192, 204)$. Another way to express the confidence interval is as the point estimate plus or minus a margin of error; in this case, it is 198 ± 6 pounds. You are 90 percent certain that the true population mean of football player weights is between 192 and 204 pounds.

Figure 6-4 The relationship between point estimate, confidence interval, and z -score.



What would happen to the confidence interval if you wanted to be 95 percent certain of it? You would have to draw the limits (ends) of the intervals closer to the tails, in order to encompass an area of .95 between them instead of .90. That would make the low value lower and the high value higher, which would make the interval wider. The width of the confidence interval is related to the significance level, standard error, and n such that the following are true:

- The higher the percentage of accuracy (significance) desired, the wider the confidence interval.
- The larger the standard error, the wider the confidence interval.
- The larger the n , the smaller the standard error, and so the narrower the confidence interval.

All other things being equal, a smaller confidence interval is always more desirable than a larger one because a smaller interval means the population parameter can be estimated more accurately.

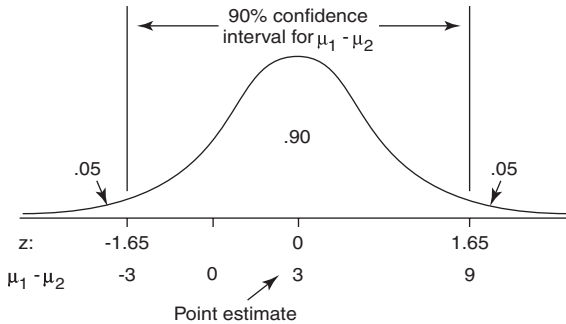
Estimating a Difference Score

Imagine that instead of estimating a single population mean μ , you wanted to estimate the difference between two population means μ_1 and μ_2 , such as the difference between the mean weights of two football teams. The statistic $\bar{x}_1 - \bar{x}_2$ has a sampling distribution just as the individual means do, and the rules of statistical inference can be used to calculate either a point estimate or a confidence interval for the difference between the two population means.

Suppose you wanted to know which was greater, the mean weight of Landers College's football team or the mean weight of Ingram College's team. You already have a point estimate of 198 pounds for Landers' team. Suppose that you draw a random sample of players from Ingram's team, and the sample mean is 195. The point estimate for the difference between the mean weights of Landers' team (μ_1) and Ingram's team (μ_2) is $198 - 195 = 3$.

But how accurate is that estimate? You can use the sampling distribution of the difference score to construct a confidence interval for $\mu_1 - \mu_2$, just as you did for Figure 6-4. Suppose that when you do so, you find that the confidence interval limits are $(-3, 9)$, which means that you are 90 percent certain that the mean for the Landers team is between three pounds lighter and nine pounds heavier than the mean for the Ingram team (see Figure 6-5).

Figure 6-5 The relationship between point estimate, confidence interval, and z-score, for a test of the difference of two means.



Suppose that instead of a confidence interval, you want to test the two-tailed hypothesis that the two team weights have different means. Your null hypothesis would be:

$$H_0: \mu_1 = \mu_2$$

or

$$H_0: \mu_1 - \mu_2 = 0$$

To reject the null hypothesis of equal means, the test statistic—in this example, z -score—for a difference in mean weights of 0 would have to fall in the rejection region at either end of the distribution. But you have already seen that it does not—only difference scores less than -3 or greater than 9 fall in the rejection region. For this reason, you would be unable to reject the null hypothesis that the two population means are equal.

This characteristic is a simple but important one of confidence intervals for difference scores. If the interval contains 0, you would be unable to reject the null hypothesis that the means are equal at the same significance level.

Univariate Tests: An Overview

To summarize, hypothesis testing of problems with one variable requires carrying out the following steps:

1. **State the null hypothesis and the alternative hypothesis.**
2. **Decide on a significance level for the test.**
3. **Compute the value of a test statistic.**

- 4. Compare the test statistic to a critical value from the appropriate probability distribution corresponding to your chosen level of significance and observe whether the test statistic falls within the region of acceptance or the region of rejection.**

Thus far, you have used the test statistic z and the table of standard normal probabilities (Table 2 in Appendix B) to carry out your tests. There are other test statistics and other probability distributions. The general formula for computing a test statistic for making an inference about a single population is

$$\text{test statistic} = \frac{\text{observed sample statistic} - \text{tested value}}{\text{standard error}}$$

where *observed sample statistic* is the statistic of interest from the sample (usually the mean), *tested value* is the hypothesized population parameter (again, usually the mean), and *standard error* is the standard deviation of the sampling distribution divided by the positive square root of n .

The general formula for computing a test statistic for making an inference about a difference between two populations is

$$\text{test statistic} = \frac{\text{statistic}_1 - \text{statistic}_2 - \text{tested value}}{\text{standard error}}$$

where *statistic*₁ and *statistic*₂ are the statistics from the two samples (usually the means) to be compared, *tested value* is the hypothesized difference between the two population parameters (0 if testing for equal values), and *standard error* is the standard error of the sampling distribution, whose formula varies according to the type of problem.

The general formula for computing a confidence interval is

$$(a, b) = \bar{x} \pm \text{test statistic} \cdot \text{standard error}$$

where a and b are the lower and upper limits of the confidence interval, \bar{x} is the point estimate (usually the sample mean), *test statistic* is the critical value from the table of the appropriate probability distribution (upper or positive value if z) corresponding to half of the desired alpha level, and *standard error* is the standard error of the sampling distribution.

Why must the alpha level be halved before looking up the critical value when computing a confidence interval? Because the rejection region is split

between both tails of the distribution, as in a two-tailed test. For a confidence interval at $\alpha = .05$, you would look up the critical value corresponding to an upper-tailed probability of .025.

Chapter Checkout

Q&A

1. An advertiser wants to know if the average age of people watching a particular TV show regularly is less than 24 years.
 - a. State the alternative and null hypotheses.
 - b. Is this a one- or two-tailed test?
 - c. A random survey of 50 viewers determines that their mean age is 19 years, with a standard deviation of 1.7 years. Find the 90 percent confidence interval of the age of the viewers.
 - d. What is the significance level for rejecting the null hypothesis?
2. True or False: Statistical tests should always be performed on a null hypothesis.
3. True or False: A result with a high level of significance is always very important.

Answers: 1a. $H_a: \mu < 24$; $H_0: \mu \geq 24$ b. one c. (18.6, 19.4) d. .0016
2. True 3. False

Critical Thinking

1. Discuss the relationship between statistical tests and the assumption of causality. For example, suppose a researcher shows that, to a level $\alpha = 0.001$, the students at Middletown High are shorter than the students at Clearview High. Would you say that going to Middletown High causes you to be short? What, if anything, would you conclude if research showed, to $\alpha = 0.001$, that people living near a power plant die at a younger age than people living elsewhere?

Chapter 7

UNIVARIATE INFERENCE TESTS

Chapter Check-In

- ❑ Learning about several statistical tests, including the z -test and t -test
- ❑ Calculating confidence levels and confidence intervals for each test
- ❑ Applying these tests to examples

Chapter 6 explained how to formulate hypotheses and test them with the z -test. In Chapter 6, you also saw how a general test statistic is constructed and can be used to test a hypothesis against any probability distribution.

In this chapter, you take a closer look at some of the most common statistical tests used in scientific and sociological research: the z - and t -tests. These tests are used to test hypotheses concerning population means or proportions. Each test is explained in a separate section with a simple format so that you can easily look up how to calculate the appropriate test statistic. The formula for finding the confidence interval of your data is given as well. Most useful of all, each section contains several simple examples to illustrate how to apply the test to your data.

One-sample z -test

Requirements: normally distributed population, σ known

Test for population mean

Hypothesis test

Formula:

$$z = \frac{\bar{X} - \Delta}{\frac{\sigma}{\sqrt{n}}}$$

where \bar{x} is the sample mean, Δ is a specified value to be tested, σ is the population standard deviation, and n is the size of the sample. Look up the significance level of the z -value in the standard normal table (Table 2 in Appendix B).

Example 1 (one-tailed test): A herd of 1,500 steers was fed a special high-protein grain for a month. A random sample of 29 were weighed and had gained an average of 6.7 pounds. If the standard deviation of weight gain for the entire herd is 7.1, what is the likelihood that the average weight gain per steer for the month was at least 5 pounds?

null hypothesis: $H_0: \mu < 5$

alternative hypothesis: $H_a: \mu \geq 5$

$$z = \frac{6.7 - 5}{\frac{7.1}{\sqrt{29}}} = \frac{1.7}{1.318} = 1.289$$

tabled value for $z \leq 1.28$ is .8997

$$1 - .8997 = .1003$$

So the probability that the herd gained at least 5 pounds per steer is $p < .1003$. Should the null hypothesis of a weight gain of less than 5 pounds for the population be rejected? That depends on how conservative you want to be. If you had decided beforehand on a significance level of $p < .05$, the null hypothesis could not be rejected.

Example 2 (two-tailed test): In national use, a vocabulary test is known to have a mean score of 68 and a standard deviation of 13. A class of 19 students takes the test and has a mean score of 65.

Is the class typical of others who have taken the test? Assume a significance level of $p < .05$.

There are two possible ways that the class may differ from the population. Its scores may be lower than, or higher than, the population of all students taking the test; therefore, this problem requires a two-tailed test. First, state the null and alternative hypotheses:

null hypothesis: $H_0: \mu = 68$

alternative hypothesis: $H_a: \mu \neq 68$

Because you have specified a significance level, you can look up the critical z -value in Table 2 of Appendix B before computing the statistic. This

is a two-tailed test; so the .05 must be split such that .025 is in the upper tail and another .025 in the lower. The z -value that corresponds to $-.025$ is -1.96 , which is the lower critical z -value. The upper value corresponds to $1 - .025$, or .975, which gives a z -value of 1.96 . The null hypothesis of no difference will be rejected if the computed z statistic falls outside of the range of -1.96 to 1.96 .

Next, compute the z statistic:

$$z = \frac{65 - 68}{\frac{13}{\sqrt{19}}} = \frac{-3}{2.982} = -1.006$$

Because -1.006 is between -1.96 and 1.96 , the null hypothesis of population mean is 68 and cannot be rejected. That is, the class can be considered as typical of others who have taken the test.

Confidence interval for population mean using z

Formula:

$$(a, b) = \bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

where a and b are the limits of the confidence interval, \bar{x} is the sample mean, $z_{\alpha/2}$ is the upper (or positive) z -value from the standard normal table corresponding to half of the desired alpha level (because all confidence intervals are two-tailed), σ is the population standard deviation, and n is the size of the sample.

Example 3: A sample of 12 machine pins has a mean diameter of 1.15 inches, and the population standard deviation is known to be .04. What is a 99 percent confidence interval of diameter width for the population?

First, determine the z -value. A 99 percent confidence level is equivalent to $p < .01$. Half of .01 is .005. The z -value corresponding to an area of .005 is 2.58. The interval may now be calculated:

$$\begin{aligned} (a, b) &= 1.15 \pm 2.58 \cdot \frac{.04}{\sqrt{12}} \\ &= 1.15 \pm .31 \\ &= (1.12, 1.18) \end{aligned}$$

It is 99 percent certain that the population mean of pin diameters lies between 1.12 and 1.18 inches. Note that this is not the same as saying that 99 percent of the machine pins have diameters between 1.12 and 1.18 inches, which would be an incorrect conclusion from this test.

Choosing a sample size

Because surveys cost money to administer, researchers often want to calculate how many subjects will be needed to determine a population mean using a fixed confidence interval and significance level. The formula is

$$n = \left(\frac{z_{\alpha/2} \sigma}{w} \right)^2$$

where n is the number of subjects needed, $z_{\alpha/2}$ is the critical z -value corresponding to the desired significance level, σ is the population standard deviation, and w is the desired confidence interval width.

Example 4: How many subjects will be needed to find the average age of students at Fisher College plus or minus a year, with a 95 percent significance level and a population standard deviation of 3.5?

$$n = \left(\frac{(2)(1.96)(3.5)}{2} \right)^2 = \left(\frac{13.72}{2} \right)^2 = 47.06$$

Rounding up, a sample of 48 students would be sufficient to determine students' mean age plus or minus one year. Note that the confidence interval width is always double the "plus or minus" figure.

One-sample t -test

Requirements: normally distributed population, σ is unknown

Test for population mean

Hypothesis test

Formula:

$$t = \frac{\bar{x} - \Delta}{\frac{s}{\sqrt{n}}}$$

where \bar{x} is the sample mean, Δ is a specified value to be tested, s is the sample standard deviation, and n is the size of the sample. When the standard deviation of the sample is substituted for the standard deviation of the population, the statistic does not have a normal distribution; it has what is called the t -distribution (see Table 3 in Appendix B). Because there is a different t -distribution for each sample size, it is not practical to list a separate area-of-the-curve table for each one. Instead, critical t -values for common alpha levels (.05, .01, .001, and so forth) are usually given in a single table for a range of sample sizes. For very large samples, the t -distribution approximates the standard normal (z) distribution.

Values in the t -table are not actually listed by sample size but by degrees of freedom (df). The number of degrees of freedom for a problem involving the t -distribution for sample size n is simply $n - 1$ for a one-sample mean problem.

Example 5 (one-tailed test): A professor wants to know if her introductory statistics class has a good grasp of basic math. Six students are chosen at random from the class and given a math proficiency test. The professor wants the class to be able to score at least 70 on the test. The six students get scores of 62, 92, 75, 68, 83, and 95. Can the professor be at least 90 percent certain that the mean score for the class on the test would be at least 70?

null hypothesis: $H_0: \mu < 70$

alternative hypothesis: $H_a: \mu \geq 70$

First, compute the sample mean and standard deviation (see Chapter 2).

$$\begin{array}{r} 62 \\ 92 \\ 75 \\ 68 \\ 83 \\ \hline 95 \\ \hline 475 \end{array} \quad \begin{array}{l} \bar{x} = \frac{475}{6} = 79.17 \\ s = 13.17 \end{array}$$

Next, compute the t -value:

$$t = \frac{79.17 - 70}{\frac{13.17}{\sqrt{6}}} = \frac{9.17}{5.38} = 1.71$$

To test the hypothesis, the computed t -value of 1.71 will be compared to the critical value in the t -table. But which do you expect to be larger and which smaller? One way to reason about this is to look at the formula and see what effect different means would have on the computation. If the sample mean had been 85 instead of 79.17, the resulting t -value would have been larger. Because the sample mean is in the numerator, the larger it is, the larger the resulting figure will be. At the same time, you know that a higher sample mean will make it more likely that the professor will conclude that the math proficiency of the class is satisfactory and that the null hypothesis of less-than-satisfactory class math knowledge can be

rejected. Therefore, it must be true that the larger the computed t -value, the greater the chance that the null hypothesis can be rejected. It follows, then, that if the computed t -value is larger than the critical t -value from the table, the null hypothesis can be rejected.

A 90 percent confidence level is equivalent to an alpha level of .10. Because extreme values in one rather than two directions will lead to rejection of the null hypothesis, this is a one-tailed test, and you do not divide the alpha level by 2. The number of degrees of freedom for the problem is $6 - 1 = 5$. The value in the t -table for $t_{.10,5}$ is 1.476. Because the computed t -value of 1.71 is larger than the critical value in the table, the null hypothesis can be rejected, and the professor can be 90 percent certain that the class mean on the math test would be at least 70.

Note that the formula for the one-sample t -test for a population mean is the same as the z -test, except that the t -test substitutes the sample standard deviation s for the population standard deviation σ and takes critical values from the t -distribution instead of the z -distribution. The t -distribution is particularly useful for tests with small samples ($n < 30$).

Example 6 (two-tailed test): A Little League baseball coach wants to know if his team is representative of other teams in scoring runs. Nationally, the average number of runs scored by a Little League team in a game is 5.7. He chooses five games at random in which his team scored 5, 9, 4, 11, and 8 runs. Is it likely that his team's scores could have come from the national distribution? Assume an alpha level of .05.

Because the team's scoring rate could be either higher than or lower than the national average, the problem calls for a two-tailed test. First, state the null and alternative hypotheses:

null hypothesis: $H_0: \mu = 5.7$

alternative hypothesis: $H_a: \mu \neq 5.7$

Next compute the sample mean and standard deviation:

$$\begin{array}{r} 5 \\ 9 \\ 4 \\ 11 \\ \underline{8} \\ 37 \end{array} \quad \begin{array}{l} \bar{x} = \frac{37}{5} = 7.4 \\ s = 2.88 \end{array}$$

Next, the t -value:

$$t = \frac{7.4 - 5.7}{\frac{2.88}{\sqrt{5}}} = \frac{1.7}{1.29} = 1.32$$

Now, look up the critical value from the t -table (Table 3 in Appendix B). You need to know two things in order to do this: the degrees of freedom and the desired alpha level. The degrees of freedom is $5 - 1 = 4$. The overall alpha level is .05, but because this is a two-tailed test, the alpha level must be divided by two, which yields .025. The tabled value for $t_{.025,4}$ is 2.776. The computed t of 1.32 is smaller than the t from Table 3, so you cannot reject the null hypothesis that the mean of this team is equal to the population mean. The coach can conclude that his team fits in with the national distribution on runs scored.

Confidence interval for population mean using t

Formula:

$$(a, b) = \bar{x} \pm t_{\alpha/2, df} \cdot \frac{s}{\sqrt{n}}$$

where a and b are the limits of the confidence interval, \bar{x} is the sample mean, $t_{\alpha/2, df}$ is the value from the t -table corresponding to half of the desired alpha level at $n - 1$ degrees of freedom, s is the sample standard deviation, and n is the size of the sample.

Example 7: Using the previous example, what is a 95 percent confidence interval for runs scored per team per game?

First, determine the t -value. A 95 percent confidence level is equivalent to an alpha level of .05. Half of .05 is .025. The t -value corresponding to an area of .025 at either end of the t -distribution for 4 degrees of freedom ($t_{.025,4}$) is 2.776. The interval may now be calculated:

$$\begin{aligned} (a, b) &= 5.7 \pm 2.78 \frac{2.88}{\sqrt{5}} \\ &= 5.7 \pm 3.58 \\ &= (2.12, 9.28) \end{aligned}$$

The interval is fairly wide, mostly because n is small.

Two-sample z-test for Comparing Two Means

Requirements: two normally distributed but independent populations, σ is known

Hypothesis test

Formula:

$$z = \frac{\bar{x}_1 - \bar{x}_2 - \Delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

where \bar{x}_1 and \bar{x}_2 are the means of the two samples, Δ is the hypothesized difference between the population means (0 if testing for equal means), σ_1 and σ_2 are the standard deviations of the two populations, and n_1 and n_2 are the sizes of the two samples.

Example 8 (two-tailed test): The amount of a certain trace element in blood is known to vary with a standard deviation of 14.1 ppm (parts per million) for male blood donors and 9.5 ppm for female donors. Random samples of 75 male and 50 female donors yield concentration means of 28 and 33 ppm, respectively. What is the likelihood that the population means of concentrations of the element are the same for men and women?

Null hypothesis: $H_0: \mu_1 = \mu_2$

or $H_0: \mu_1 - \mu_2 = 0$

alternative hypothesis: $H_a: \mu_1 \neq \mu_2$

or: $H_a: \mu_1 - \mu_2 \neq 0$

$$z = \frac{28 - 33 - 0}{\sqrt{\frac{14.1^2}{75} + \frac{9.5^2}{50}}} = \frac{-5}{\sqrt{2.65 + 1.81}} = -2.37$$

The computed z -value is negative because the (larger) mean for females was subtracted from the (smaller) mean for males. But because the hypothesized difference between the populations is 0, the order of the samples in this computation is arbitrary— \bar{x}_1 could just as well have been the female sample mean and \bar{x}_2 the male sample mean, in which case z would be 2.37 instead of -2.37 . An extreme z -score in either tail of the distribution (plus or minus) will lead to rejection of the null hypothesis of no difference.

The area of the standard normal curve corresponding to a z -score of -2.37 is .0089. Because this test is two-tailed, that figure is doubled to yield a probability of .0178 that the population means are the same. If the test had been conducted at a pre-specified significance level of $\alpha < .05$, the null hypothesis of equal means could be rejected. If the specified significance level had been the more conservative (more stringent) $\alpha < .01$, however, the null hypothesis could not be rejected.

In practice, the two-sample z -test is not often used because the two population standard deviations σ_1 and σ_2 are usually unknown. Instead, sample standard deviations and the t -distribution are used.

Two-sample t -test for Comparing Two Means

Requirements: two normally distributed but independent populations, σ is unknown

Hypothesis test

Formula:

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where \bar{x}_1 and \bar{x}_2 are the means of the two samples, Δ is the hypothesized difference between the population means (0 if testing for equal means), s_1 and s_2 are the standard deviations of the two samples, and n_1 and n_2 are the sizes of the two samples. The number of degrees of freedom for the problem is the smaller of $n_1 - 1$ and $n_2 - 1$.

Example 9 (one-tailed test): An experiment is conducted to determine whether intensive tutoring (covering a great deal of material in a fixed amount of time) is more effective than paced tutoring (covering less material in the same amount of time). Two randomly chosen groups are tutored separately and then administered proficiency tests. Use a significance level of $\alpha < .05$.

null hypothesis: $H_0: \mu_1 \leq \mu_2$

or $H_0: \mu_1 - \mu_2 \leq 0$

alternative hypothesis: $H_a: \mu_1 > \mu_2$

or: $H_a: \mu_1 - \mu_2 > 0$

<i>Group</i>	<i>Method</i>	<i>n</i>	\bar{x}	<i>s</i>
1	intensive	12	46.31	6.44
2	paced	10	42.79	7.52

$$t = \frac{46.31 - 42.79 - 0}{\sqrt{\frac{6.44^2}{12} + \frac{7.52^2}{10}}} = \frac{3.52}{\sqrt{3.46 + 5.66}} = 1.166$$

The degrees of freedom parameter is the smaller of $(12 - 1)$ and $(10 - 1)$, or 9. Because this is a one-tailed test, the alpha level (.05) is not divided by two. The next step is to look up $t_{.05,9}$ in the t -table (Table 3 in Appendix B), which gives a critical value of 1.833. The computed t of 1.166 does not exceed the tabled value, so the null hypothesis cannot be rejected. This test has not provided statistically significant evidence that intensive tutoring is superior to paced tutoring.

Confidence interval for comparing two means

Formula: $(a, b) = \bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2, df} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

where a and b are the limits of the confidence interval, \bar{x}_1 and \bar{x}_2 are the means of the two samples, $t_{\alpha/2, df}$ is the value from the t -table corresponding to half of the desired alpha level, s_1 and s_2 are the standard deviations of the two samples, and n_1 and n_2 are the sizes of the two samples. The degrees of freedom parameter for looking up the t -value is the smaller of $n_1 - 1$ and $n_2 - 1$.

Example 10: Estimate a 90 percent confidence interval for the difference between the number of raisins per box in two brands of breakfast cereal.

<i>Brand</i>	<i>n</i>	\bar{x}	<i>s</i>
A	6	102.1	12.3
B	9	93.6	7.52

The difference between \bar{x}_1 and \bar{x}_2 is $102.1 - 93.6 = 8.5$. The degrees of freedom is the smaller of $(6 - 1)$ and $(9 - 1)$, or 5. A 90 percent confidence interval is equivalent to an alpha level of .10, which is then halved to give .05. According to Table 3, the critical value for $t_{.05,5}$ is 2.015. The interval may now be computed.

$$\begin{aligned}
 (a, b) &= 8.5 \pm 2.015 \cdot \sqrt{\frac{(12.3)^2}{6} + \frac{(7.52)^2}{9}} \\
 &= 8.5 \pm 2.015 \cdot \sqrt{25.22 + 6.28} \\
 &= 8.5 \pm 11.31 \\
 &= (-2.81, 19.81)
 \end{aligned}$$

You can be 90 percent certain that Brand A cereal has between 2.81 fewer and 19.81 more raisins per box than Brand B. The fact that the interval contains 0 means that if you had performed a test of the hypothesis that the two population means are different (using the same significance level), you would not have been able to reject the null hypothesis of no difference.

Pooled variance method

If the two population distributions can be assumed to have the same variance—and therefore the same standard deviation— s_1 and s_2 can be pooled together, each weighted by the number of cases in each sample. Although using pooled variance in a t -test is generally more likely to yield significant results than using separate variances, it is often hard to know whether the variances of the two populations are equal. For this reason, the pooled variance method should be used with caution. The formula for the pooled estimator of σ^2 is

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

where s_1 and s_2 are the standard deviations of the two samples and n_1 and n_2 are the sizes of the two samples.

The formula for comparing the means of two populations using pooled variance is

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \Delta}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

where \bar{x}_1 and \bar{x}_2 are the means of the two samples, Δ is the hypothesized difference between the population means (0 if testing for equal means), S_p^2 is the pooled variance, and n_1 and n_2 are the sizes of the two samples. The number of degrees of freedom for the problem is

$$df = n_1 + n_2 - 2$$

Example 11 (two-tailed test): Does right- or left-handedness affect how fast people type? Random samples of students from a typing class are given a typing speed test (words per minute), and the results are compared. Significance level for the test: .10. Because you are looking for a difference between the groups in either direction (right-handed faster than left, or vice versa), this is a two-tailed test.

null hypothesis: $H_0: \mu_1 = \mu_2$

or: $H_0: \mu_1 - \mu_2 = 0$

alternative hypothesis: $H_a: \mu_1 \neq \mu_2$

or: $H_a: \mu_1 - \mu_2 \neq 0$

Group	-handed	n	\bar{x}	s
1	right	16	55.8	5.7
2	left	9	59.3	4.3

First, calculate the pooled variance:

$$\begin{aligned}
 S_p^2 &= \frac{(16-1)5.7^2 + (9-1)4.3^2}{16+9-2} \\
 &= \frac{487.35 + 147.92}{23} \\
 &= 27.62
 \end{aligned}$$

Next, calculate the t -value:

$$t = \frac{55.8 - 59.3 - 0}{\sqrt{27.62 \left(\frac{1}{16} + \frac{1}{9} \right)}} = \frac{-3.5}{\sqrt{4.80}} = -1.598$$

The degrees-of-freedom parameter is $16 + 9 - 2$, or 23. This test is a two-tailed one, so you divide the alpha level (.10) by two. Next, you look up $t_{.05,23}$ in the t -table (Table 3 in Appendix B), which gives a critical value of 1.714. This value is larger than the absolute value of the computed t of -1.598 , so the null hypothesis of equal population means cannot be rejected. There is no evidence that right- or left-handedness has any effect on typing speed.

Paired Difference *t*-test

Requirements: a set of paired observations from a normal population

This *t*-test compares one set of measurements with a second set from the same sample. It is often used to compare “before” and “after” scores in experiments to determine whether significant change has occurred.

Hypothesis test

Formula:
$$t = \frac{\bar{x} - \Delta}{\frac{s}{\sqrt{n}}}$$

where \bar{x} is the mean of the change scores, Δ is the hypothesized difference (0 if testing for equal means), s is the sample standard deviation of the differences, and n is the sample size. The number of degrees of freedom for the problem is $n - 1$.

Example 12 (one-tailed test): A farmer decides to try out a new fertilizer on a test plot containing 10 stalks of corn. Before applying the fertilizer, he measures the height of each stalk. Two weeks later, he measures the stalks again, being careful to match each stalk’s new height to its previous one. The stalks would have grown an average of six inches during that time even without the fertilizer. Did the fertilizer help? Use a significance level of .05.

null hypothesis: $H_0: \Delta \leq 6$

alternative hypothesis: $H_a: \Delta > 6$

Stalk	1	2	3	4	5	6	7	8	9	10
Before height	35.5	31.7	31.2	36.3	22.8	28.0	24.6	26.1	34.5	27.7
After height	45.3	36.0	38.6	44.7	31.4	33.5	28.8	35.8	42.9	35.0

Subtract each stalk’s “before” height from its “after” height to get the change score for each stalk; then compute the mean and standard deviation of the change scores and insert these into the formula.

$$45.3 - 35.5 = 9.8$$

$$36.0 - 31.7 = 4.3$$

$$38.6 - 31.2 = 7.4$$

$$44.7 - 36.3 = 8.4$$

$$31.4 - 22.8 = 8.6$$

$$33.5 - 28.0 = 5.5$$

$$28.8 - 24.6 = 4.2$$

$$35.8 - 26.1 = 9.7$$

$$42.9 - 34.5 = 8.4$$

$$35.0 - 27.7 = 7.3$$

$$73.6/10 = 7.36$$

$$s^2 = \frac{(9.8 - 7.8)^2 + (4.3 - 7.36)^2 + \dots + (7.3 - 7.36)^2}{\sqrt{10 - 1}} = 4.216$$

$$s = \sqrt{4.216}$$

$$s = 2.05$$

$$t = \frac{7.36 - 6}{\frac{2.05}{\sqrt{10}}} = \frac{1.36}{.65} = 2.098$$

The problem has $n - 1$, or $10 - 1 = 9$ degrees of freedom. The test is one-tailed because you are asking only whether the fertilizer increases growth, not reduces it. The critical value from the t -table for $t_{0.05,9}$ is 1.833. Because the computed t -value of 2.098 is larger than 1.833, the null hypothesis can be rejected. The test has provided evidence that the fertilizer caused the corn to grow more than if it had not been fertilized. The amount of actual increase was not large (1.36 inches over normal growth), but it was statistically significant.

Test for a Single Population Proportion

Requirements: binomial population, sample $n\pi_0 \geq 10$, and sample $n(1 - \pi_0) \geq 10$, where π_0 is the hypothesized proportion of successes in the population.

Hypothesis test

Formula:

$$z = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1 - \pi_0)}{n}}}$$

where $\hat{\pi}$ is the sample proportion, π_0 is the hypothesized proportion, and n is the sample size. Because the distribution of sample proportions is approximately normal for large samples, the z statistic is used. The test is most accurate when π (the population proportion) is close to .5 and least accurate when π is close to 0 or 1.

Example 13 (one-tailed test): The sponsors of a city marathon have been trying to encourage more women to participate in the event. A sample of 70 runners is taken, of which 32 are women. The sponsors would like to be 90 percent certain that at least 40 percent of the participants are women. Were their recruitment efforts successful?

null hypothesis: $H_0: \pi < .4$

alternative hypothesis: $H_0: \pi \geq .4$

The proportion of women runners in the sample is 32 out of 70, or 45.7 percent. The z -value may now be calculated:

$$z = \frac{.457 - .40}{\sqrt{\frac{.40(1 - .40)}{70}}} = \frac{.057}{\sqrt{.003}} = .973$$

From the z -table, you find that the probability of a z -value less than .97 is .834, not quite the .90 required to reject the null hypothesis, so it cannot be concluded at that level of significance that the population of runners is at least 40 percent women.

Confidence interval for a single population proportion

Formula:
$$\hat{\pi} \pm z_{\alpha/2} \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$$

where $\hat{\pi}$ is the sample proportion, $z_{\alpha/2}$ is the upper z -value corresponding to half of the desired alpha level, and n is the sample size.

Example 14: A sample of 100 voters selected at random in a congressional district prefer Candidate Smith to Candidate Jones by a ratio of 3 to 2. What is a 95 percent confidence interval of the percentage of voters in the district who prefer Smith?

A ratio of 3 to 2 is equivalent to a proportion of $3/5 = .60$. A 95 percent confidence interval is equivalent to an alpha level of .05, half of which is .025. The critical z -value corresponding to an upper probability of $1 - .025$ is 1.96. The interval may now be computed:

$$\begin{aligned}(a, b) &= .60 \pm 1.96 \sqrt{\frac{.60(1-.60)}{100}} \\ &= .60 \pm 1.96 \sqrt{.002} \\ &= .60 \pm .096 \\ &= (.504, .696)\end{aligned}$$

It is 95 percent certain that between 50.4 percent and 69.6 percent of the voters in the district prefer Candidate Smith. Note that the problem could have been figured for Candidate Jones by substituting the proportion .40 for Smith's proportion of .60.

Choosing a sample size

In the previous problem, you estimated that the percentage of voters in the district who prefer Candidate Smith is 60 percent plus or minus about 10 percent. Another way to say this is that the estimate has a "margin of error" of ± 10 percent, or a confidence interval width of 20 percent. That's a pretty wide range. You may wish to make the margin smaller.

Because the width of the confidence interval decreases at a known rate as the sample size increases, it is possible to determine the sample size needed to estimate a proportion with a fixed confidence interval. The formula is

$$n = \left(\frac{2z_{\alpha/2}}{w} \right)^2 \cdot p^* (1 - p^*)$$

where n is the number of subjects needed, $z_{\alpha/2}$ is the z -value corresponding to half of the desired significance level, w is the desired confidence interval width, and p^* is an estimate of the true population proportion. A p^* of .50 will result in a higher n than any other proportion estimate but is often used when the true proportion is not known.

Example 15: How large a sample is needed to estimate the preference of district voters for Candidate Smith with a margin of error of ± 4 percent, at a 95 percent significance level?

You will conservatively estimate the (unknown) true population proportion of preference for Smith at .50. If it's really larger (or smaller) than that, you will overestimate the size of the sample needed, but $p^* = .50$ is playing it safe.

$$\begin{aligned} n &= \left(\frac{2 \cdot 1.96}{.08} \right)^2 \cdot .50(1 - .50) \\ &= 49^2 \cdot .25 \\ &= 600.25 \end{aligned}$$

A sample of about 600 voters would be needed to estimate the percentage of voters in the district who prefer Smith and to be 95 percent certain that the estimate is within (plus or minus) 4 percent of the true population percentage.

Test for Comparing Two Proportions

Requirements: two binomial populations, $\pi_0 \geq 5$ and $n(1 - \pi_0) \geq 5$ (for each sample), where π_0 is the hypothesized proportion of successes in the population.

Difference test

Hypothesis test

Formula:

$$z = \frac{\hat{\pi}_1 - \hat{\pi}_2 - \Delta}{\sqrt{\hat{\pi}(1 - \hat{\pi})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$\hat{\pi} = \frac{x_1 + x_2}{n_1 + n_2}$$

and where $\hat{\pi}_1$ and $\hat{\pi}_2$ are the sample proportions, Δ is their hypothesized difference (0 if testing for equal proportions), n_1 and n_2 are the sample sizes, and x_1 and x_2 are the number of “successes” in each sample. As in the test for a single proportion, the z distribution is used to test the hypothesis.

Example 16 (one-tailed test): A swimming school wants to determine whether a recently hired instructor is working out. Sixteen out of 25 of Instructor A’s students passed the lifeguard certification test on the first try. In comparison, 57 out of 72 of more experienced Instructor B’s students passed the test on the first try. Is Instructor A’s success rate worse than Instructor B’s? Use $\alpha = .10$.

null hypothesis: $H_0: \pi_1 \geq \pi_2$

alternative hypothesis: $H_a: \pi_1 < \pi_2$

First, you need to compute the values for some of the terms in the formula. The sample proportion $\hat{\pi}_1$ is $16/25 = .640$. The sample proportion $\hat{\pi}_2$ is $57/72 = .792$. Next, compute $\hat{\pi}$:

$$\hat{\pi} = \frac{16 + 57}{25 + 72} = \frac{73}{97} = .753$$

Finally, the main formula:

$$\begin{aligned} z &= \frac{.640 - .792 - 0}{\sqrt{.753(1 - .753)\left(\frac{1}{25} + \frac{1}{72}\right)}} \\ &= \frac{-.152}{\sqrt{(.186)(.054)}} = \frac{-.152}{.100} = -1.52 \end{aligned}$$

The standard normal (z) table shows that the lower critical z -value for $\alpha = .10$ is approximately -1.28 . The computed z must be lower than -1.28 in order to reject the null hypothesis of equal proportions. Because the computed z is -1.517 , the null hypothesis can be rejected. It can be concluded (at this level of significance) that Instructor A’s success rate is worse than Instructor B’s.

Confidence interval for comparing two proportions

Formula: $(a, b) = \hat{\pi}_1 - \hat{\pi}_2 \pm z_{\alpha/2} \cdot s(D)$

where $s(D) = \sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}$

and where a and b are the limits of the confidence interval of $\pi_1 - \pi_2$, $\hat{\pi}_1$ and $\hat{\pi}_2$ are the sample proportions, $z_{\alpha/2}$ is the upper z -value corresponding to half of the desired alpha level, and n_1 and n_2 are the sizes of the two samples.

Example 17: A public health researcher wants to know how two high schools, one in the inner city and one in the suburbs, differ in the percentage of students who smoke. A random survey of students gives the following results:

<i>Population</i>	<i>n</i>	<i>Smokers</i>
1 (inner-city)	125	47
2 (suburban)	153	52

What is a 90 percent confidence interval for the difference between the smoking rates in the two schools?

The proportion of smokers in the inner-city school is $\hat{\pi}_1 = 47/125 = .376$. The proportion of smokers in the suburban school is $\hat{\pi}_2 = 52/153 = .340$. Next solve for $s(D)$:

$$\begin{aligned} s(D) &= \sqrt{\frac{.376(1-.376)}{125} + \frac{.340(1-.340)}{153}} \\ &= \sqrt{\frac{.235}{125} + \frac{.224}{153}} = \sqrt{.003} = .058 \end{aligned}$$

A 90 percent confidence interval is equivalent to $\alpha = .10$, which is halved to give $.05$. The upper tabled value for $z_{.05}$ is 1.65. The interval may now be computed:

$$\begin{aligned} (a, b) &= (.376 - .340) \pm (1.65)(.058) \\ &= .036 \pm .096 \\ &= (-.060, .132) \end{aligned}$$

The researcher can be 90 percent certain that the true population proportion of smokers in the inner-city high school is between 6 percent lower and 13.2 percent higher than the proportion of smokers in the suburban high school. Thus, since the confidence interval contains zero, there is no difference between the two types of schools at $\alpha = .10$.

Chapter Checkout

Q&A

1. Suppose you count the number of days in May on which it rains in your town and find it to be 15 (out of 31). Checking the weather database, you find that for your town, the average number of rainy days in May is 10.
 - a. What type of test would you use to determine if this May was unusually wet to $\alpha = 0.05$?
 - b. Perform the test. What is the value of your test statistic (that is, z or t)?
2. A biologist believes that polluted water may be causing frogs in a lake to be smaller than normal. In an unpolluted lake, she measures 15 adult frogs and finds a mean length of 7.6 inches with a standard deviation of 1.5 inches. In the polluted lake, she measures 23 adult frogs and finds a mean of 6.5 inches and a standard deviation of 2.3 inches.
 - a. State the null and alternative hypotheses.
 - b. Calculate the appropriate test statistic.
3. A high school track coach makes the sprinters on his team begin weight lifting, in the hopes that their speed will increase. Before beginning, each runs a 40 yard dash. After a month of training, each runs the dash again. Their times are given below.

<i>Runner</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>
Before training	4.8	5.1	5.5	4.9	5.6	6.0	5.8	5.3	6.1	4.7
After training	4.7	5.0	5.1	5.0	5.2	5.1	5.6	5.1	5.5	5.0

- a.** State the alternative and null hypotheses.
b. To what level can the coach be sure that his runners have improved?

Answers: **1a.** test for single population proportion **b.** 1.92
2a. $H_a: \mu_{\text{pol}} < \mu_{\text{unpol}}$; $H_o: \mu_{\text{pol}} \geq \mu_{\text{unpol}}$ **b.** 1.784 **3a.** $H_a: \mu_{\text{after}} < \mu_{\text{before}}$; $H_o: \mu_{\text{after}} \geq \mu_{\text{before}}$ **b.** $p < .025$

Chapter 8

BIVARIATE RELATIONSHIPS

Chapter Check-In

- Checking for correlation between two variables
- Learning about regression
- Using the chi-square (χ^2) test to determine whether two variables are independent

So far, you have been working with problems involving a single variable. Many times, however, you want to know something about the relationship between two variables.

Correlation

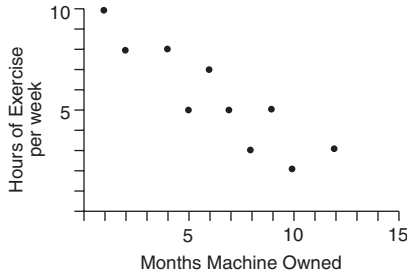
Consider Table 8-1, which contains measurements on two variables for ten people: the number of months the person has owned an exercise machine and the number of hours the person spent exercising in the past week.

Table 8-1 Exercise Data for Ten People

<i>Person</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>
<i>Months Owned</i>	5	10	4	8	2	7	9	6	1	12
<i>Hours Exercised</i>	5	2	8	3	8	5	5	7	10	3

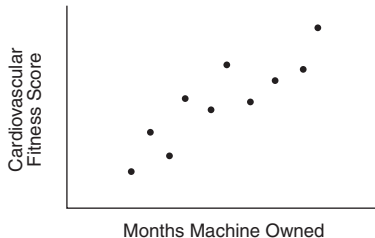
If you display these data pairs as points in a scatter plot (see Figure 8-1), you can see a definite trend. The points appear to form a line that slants from the upper left to the lower right. As you move along that line from left to right, the values on the vertical axis (hours of exercise) get smaller, while the values on the horizontal axis (months owned) get larger. Another way to express this is to say that the two variables are inversely related: The more months the machine was owned, the less the person tended to exercise.

Figure 8-1 The data in Table 8-1 is an example of negative correlation.



These two variables are **correlated**. More than that, they are correlated in a particular direction—negatively. For an example of a positive correlation, suppose that instead of displaying “hours of exercise” on the vertical axis, you put the person’s score on a test that measures cardiovascular fitness (see Figure 8-2). The pattern of these data points suggests a line that slants from lower left to upper right, which is the opposite of the direction of slant in the first example. Figure 8-2 shows that the longer the person has owned the exercise machine, the better his or her cardiovascular fitness tends to be. This might be true in spite of the fact that time spent exercising decreases the longer the machine has been owned because purchasers of exercise machines might be starting from a point of low fitness, which may improve only gradually.

Figure 8-2 An example of positive correlation.

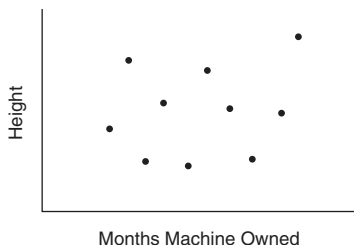


If two variables are positively correlated, as the value of one increases, so does the value of the other. If they are negatively (or inversely) correlated, as the value of one increases, the value of the other decreases.

A third possibility remains: that as the value of one variable increases, the value of the other neither increases nor decreases. Figure 8-3 is a scatter

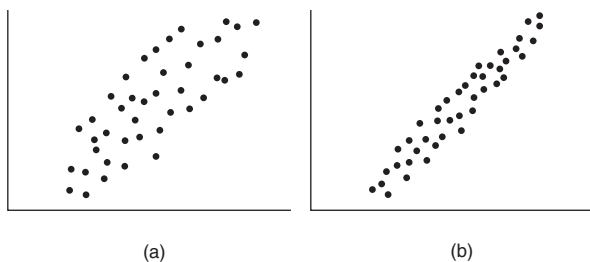
plot of months exercise machine owned (horizontal axis) by person's height (vertical axis). No line trend can be seen in the plot. New owners of exercise machines may be short or tall, and the same is true of people who have had their machines longer. These two variables appear to be uncorrelated.

Figure 8-3 An example of uncorrelated data.



You can go even farther in expressing the relationship between variables. Compare the two scatter plots in Figure 8-4. Both plots show a positive correlation because, as the values on one axis increase, so do the values on the other. But the data points in Figure 8-4 (b) are more closely packed than the data points in Figure 8-4 (a), which are more spread out. If a line were drawn through the middle of the trend, the points in Figure 8-4 (b) would be closer to the line than the points in Figure 8-4 (a). In addition to direction (positive or negative), correlations can also have strength, which is a reflection of the closeness of the data points to a perfect line. Figure 8-4 (b) shows a stronger correlation than Figure 8-4 (a).

Figure 8-4 (a) Weak and (b) strong correlations.



Pearson's product moment coefficient (r), commonly referred to as the correlation coefficient, is a quantitative measure of correlation between

two interval-level variables. The coefficient r can take values from -1.0 to 1.0 . The sign of r indicates whether the correlation is positive or negative. The magnitude (absolute value) of r indicates the strength of the correlation, or how close the array of data points is to a straight line.

The computing formula for r is

$$r = \frac{\sum xy - \frac{1}{n}(\sum x)(\sum y)}{(n-1)s_x s_y}$$

where $\sum xy$ is the sum of the xy cross-products (each x multiplied by its paired y), n is the size of the sample (the number of data pairs), $\sum x$ and $\sum y$ are the sums of the x and y values, and s_x and s_y are the sample standard deviations of x and y .

Example 1: Use Table 8-2 to compute r for the relationship between months of exercise-machine ownership and hours of exercise per week. The first step is to compute the components required in the main formula. Let x be months of ownership and y be hours of exercise, although you could also do the reverse.

Table 8-2 Determining the Correlation Coefficient for Table 8-1

x	y	xy	x^2	y^2
5	5	25	25	25
10	2	20	100	4
4	8	32	16	64
8	3	24	64	9
2	8	16	4	64
7	5	35	49	25
9	5	45	81	25
6	7	42	36	49
1	10	10	1	100
12	3	36	144	9
$\Sigma x = 64$	$\Sigma y = 56$	$\Sigma xy = 285$	$\Sigma x^2 = 520$	$\Sigma y^2 = 374$

$$(\sum x)^2 = 64^2 = 4096$$

$$(\sum y)^2 = 56^2 = 3136$$

Now, compute the sample standard deviations for x and y using the formula from Chapter 3:

$$\begin{aligned} s_x &= \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}} \\ &= \sqrt{\frac{520 - \frac{4096}{10}}{10-1}} \\ &= \sqrt{\frac{110.4}{9}} = \sqrt{12.267} = 3.502 \end{aligned}$$

$$\begin{aligned} s_y &= \sqrt{\frac{\sum y^2 - \frac{(\sum y)^2}{n}}{n-1}} \\ &= \sqrt{\frac{60.4}{9}} = \sqrt{6.711} = 2.591 \end{aligned}$$

Finally, r may be computed:

$$\begin{aligned} r_{xy} &= \frac{\sum xy - \frac{1}{n}(\sum x)(\sum y)}{(n-1)s_x s_y} \\ &= \frac{285 - \frac{1}{10}(64)(56)}{(10-1)(3.502)(2.591)} \\ &= \frac{-73.4}{81.663} = -0.899 \end{aligned}$$

A correlation of $r = -.899$ is almost as strong as the maximum negative correlation of -1.0 , reflecting the fact that your data points fall relatively close to a straight line.

Finding the significance of r

You might want to know how significant an r of $-.899$ is. The formula to test the null hypothesis that R (the population correlation) = 0 is

$$t = \frac{r\sqrt{(n-2)}}{\sqrt{(1-r^2)}}$$

where r is the sample correlation coefficient, and n is the size of the sample (the number of data pairs). The probability of t may be looked up in Table 3 (in Appendix B) using $n-2$ degrees of freedom.

$$t = \frac{-.899\sqrt{(10-2)}}{\sqrt{1-(-1.899)^2}} = \frac{-2.543}{.438} = -5.807$$

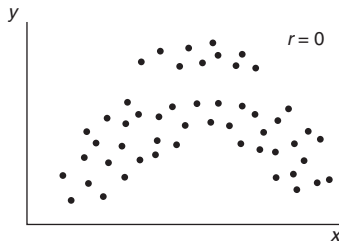
The probability of obtaining a t of -5.807 with 8 df (drop the sign when looking up the value of t) is lower than the lowest listed probability of .0005. If the correlation between months of exercise-machine ownership and hours of exercise per week were actually 0, you would expect an r of $-.899$ or lower in fewer than one out of a thousand random samples.

To evaluate a correlation coefficient, first determine its significance. If the probability that the coefficient resulted from chance is not acceptably low, the analysis should end there; neither the coefficient's sign nor its magnitude may reflect anything other than sampling error. If the coefficient is statistically significant, the sign should give an indication of the direction of the relationship, and the magnitude indicates its strength. Remember, however, that all statistics become significant with a high enough n .

Even if it's statistically significant, whether a correlation of a given magnitude is substantively or practically significant depends greatly on the phenomenon being studied. Generally, correlations tend to be higher in the physical sciences, where relationships between variables often obey uniform laws, and lower in the social sciences, where relationships may be harder to predict. A correlation of .4 between a pair of sociological variables may be more meaningful than a correlation of .7 between two variables in physics.

Bear in mind also that the correlation coefficient measures only straight-line relationships. Not all relationships between variables trace a straight line. Figure 8-5 shows a curvilinear relationship such that values of y increase along with values of x up to a point, then decrease with higher values of x . The correlation coefficient for this plot is 0, the same as for the plot in Figure 8-3. This plot, however, shows a relationship that Figure 8-3 does not.

Correlation does not imply causation. The fact that Variable A and Variable B are correlated does not necessarily mean that A caused B or that B caused A (though either may be true). If you were to examine a database of demographic information, for example, you would find that the number of churches in a city is correlated with the number of violent crimes in the city. The reason is not that church attendance causes crime, but that these two variables both increase as a function of a third variable: population.

Figure 8-5 Data that can cause trouble with correlation analysis.

Also note that the scales used to measure two variables have no effect on their correlation. If you had converted hours of exercise per week to minutes per week, and/or months machine owned to days owned, the correlation coefficient would have been the same.

The coefficient of determination

The square of the correlation coefficient r is called the **coefficient of determination** and can be used as a measure of the proportion of variability that two variables share, or how much one can be “explained” by the other. The coefficient of determination for this example is $-0.899^2 = .808$. Approximately 80 percent of the variability of each variable is shared with the other variable.

Simple Linear Regression

Although correlation is not concerned with causation in relationships among variables, a related statistical procedure called **regression** often is. Regression is used to assess the contribution of one or more “causing” variables (called **independent** variables) to one “caused” (or **dependent**) variable. It can also be used to predict the value of one variable from the values of others. When there is only one independent variable and when the relationship can be expressed as a straight line, the procedure is called simple linear regression.

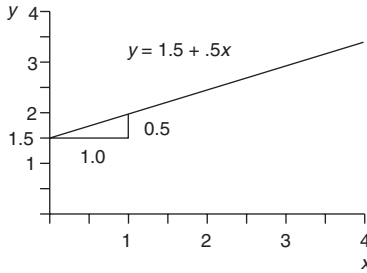
Any straight line in two-dimensional space can be represented by this equation:

$$y = a + bx$$

where y is the variable on the vertical axis, x is the variable on the horizontal axis, a is the y -value where the line crosses the vertical axis (often called

the **intercept**), and b is the amount of change in y corresponding to a one-unit increase in x (often called the **slope**). Figure 8-6 gives an example.

Figure 8-6 A straight line.



Returning to the exercise example, you observed that the scatter plot of points in Figure 8-1 resembles a line. The regression procedure fits the best possible straight line to an array of data points. If no single line can be drawn such that all the points fall on it, what is the “best” line? The best line is the one that minimizes the distance of all the data points to the line, a fact that will become clearer after you compute the line for the example.

Regression is an inferential procedure, meaning that it can be used to draw conclusions about populations based on samples randomly drawn from those populations. Suppose that your ten exercise-machine owners were randomly selected to represent the population of all exercise-machine owners. In order to use this sample to make educated guesses about the relationship between the two variables (months of machine ownership and time spent exercising) in the population, you need to rewrite the equation above to reflect the fact that you will be estimating population parameters:

$$y = \beta_0 + \beta_1 x$$

All you have done is replace the intercept (a) with β_0 and the slope (b) with β_1 . The formula to compute the parameter estimate $\hat{\beta}_1$ is

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$$

where

$$S_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

and

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

The formula to compute the parameter estimate $\hat{\beta}_0$ is

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

where \bar{y} and \bar{x} are the two sample means.

You have already computed the quantities that you need to substitute into these formulas for the exercise example—except for the mean of $x(\bar{x})$, which is $64/10 = 6.4$, and the mean of $y(\bar{y})$, which is $56/10 = 5.6$. First, compute the estimate of the slope:

$$S_{xy} = 285 - \frac{(64)(56)}{10} = -73.4$$

$$S_{xx} = 520 - \frac{4096}{10} = 110.4$$

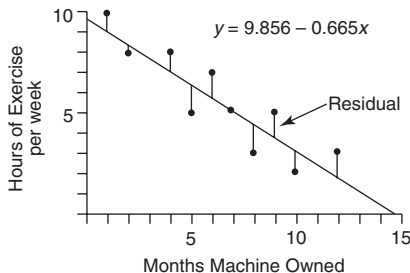
$$\hat{\beta}_1 = \frac{-73.4}{110.4} = -0.665$$

Now the intercept may be computed:

$$\begin{aligned} \hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x} \\ &= 5.6 - (-0.665(6.4)) \\ &= 5.6 - (-4.256) \\ &= 9.856 \end{aligned}$$

So the regression equation for the example is $y = 9.856 - 0.665(x)$. When you plot this line over the data points, the result looks like that shown in Figure 8-7.

Figure 8-7 Illustration of residuals.



The vertical distance from each data point to the regression line is the error, or **residual**, of the line's accuracy in estimating that point. Some points have positive residuals (they lie above the line); some have negative ones (they lie below it). If all the points fell on the line, there would be no error and no residuals. The mean of the sample residuals is always 0 because the regression line is always drawn such that half of the error is above it and half below. The equations that you used to estimate the intercept and slope determine a line of "best fit" by minimizing the sum of the squared residuals. This method of regression is called **least squares**.

Because regression estimates usually contain some error (that is, all points do not fall on the line), an error term (ϵ , the Greek letter epsilon) is usually added to the end of the equation:

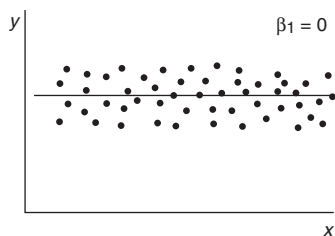
$$y = \beta_0 + \beta_1 x + \epsilon$$

The estimate of the slope β_1 for the exercise example was -0.665 . The slope is negative because the line slants down from left to right, as it must for two variables that are negatively correlated, reflecting that one variable decreases as the other increases. When the correlation is positive, β_1 is positive, and the line slants up from left to right.

Confidence interval for the slope

Example 2: What if the slope is 0, as in Figure 8-8? That means that y has no linear dependence on x , or that knowing x does not contribute anything to your ability to predict y .

Figure 8-8 Example of uncorrelated data, so the slope is zero.



It is often useful to compute a confidence interval for a regression slope. If it contains 0, you would be unable to conclude that x and y are related. The formula to compute a confidence interval for β_1 is

$$(a, b) = \hat{\beta}_1 \pm t_{\alpha/2, n-2} \cdot \frac{s}{\sqrt{S_{xx}}}$$

where

$$s = \sqrt{\frac{\sum(y - \hat{y})^2}{n - 2}}$$

and

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$$

and where $\sum(y - \hat{y})^2$ is the sum of the squared residuals, $t_{\alpha/2, n-2}$ is the critical value from the t -table corresponding to half the desired alpha level at $n - 2$ degrees of freedom, and n is the size of the sample (the number of data pairs). The test for this example will use an alpha of .05. Table 3 (in Appendix B) shows that $t_{0.025, 8} = 2.306$.

Compute the quantity $\sum(y - \hat{y})^2$ by subtracting each predicted y -value (\hat{y}) from each actual y -value, squaring it, and summing the squares (see Table 8-3). The predicted y -value (\hat{y}) is the y -value that would be predicted from each given x , using the formula $y = 9.856 - 0.665(x)$.

Table 8-3 Determining the Residuals for the Data in Table 8-1

x	y		\hat{y}	=	<i>residual</i>	<i>residual</i> ²
5	5	-	6.530	=	-1.530	2.341
10	2	-	3.205	=	-1.205	1.452
4	8	-	7.195	=	0.805	.648
8	3	-	4.535	=	-1.535	2.356
2	8	-	8.525	=	-0.525	.276
7	5	-	5.200	=	-0.200	.040
9	5	-	3.870	=	1.130	1.277
6	7	-	5.865	=	1.135	1.288
1	10	-	9.190	=	0.810	.656
12	3	-	1.875	=	<u>1.125</u>	<u>1.266</u>
					0	11.600

Now, compute s :

$$s = \sqrt{\frac{11,600}{10-2}} = \sqrt{1,450} = 1,204$$

You have already determined that $S_{xx} = 110.4$; you can proceed to the main formula:

$$\begin{aligned}(a, b) &= -0.665 \pm 2,306 \frac{1,204}{\sqrt{110.4}} \\ &= -0.665 \pm 2,306 \frac{1,204}{10,507} \\ &= -0.665 \pm 2,306(0.115) \\ &= -0.665 \pm 0,264 \\ &= (-0,929, -0,401)\end{aligned}$$

You can be 95 percent certain that the population parameter β_1 (the slope) is no lower than -0.929 and no higher than -0.401 . Because this interval does not contain 0, you would be able to reject the null hypothesis that $\beta_1 = 0$ and conclude that these two variables are indeed related in the population.

Confidence interval for prediction

You have learned that you could predict a y -value from a given x -value. Because there is some error associated with your prediction, however, you might want to produce a confidence interval rather than a simple point estimate. The formula for a prediction interval for y for a given x is

$$(a, b) = \hat{y} \pm t_{\alpha/2, n-2} (s) \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{S_{xx}}}$$

where $S_{xx} = \sum X^2 - \frac{(\sum x)^2}{n}$

and where \hat{y} is the y -value predicted for x using the regression equation, $t_{\alpha/2, n-2}$ is the critical value from the t -table corresponding to half the desired alpha level at $n-2$ degrees of freedom, and n is the size of the sample (the number of data pairs).

Example 3: What is a 90 percent confidence interval for the number of hours spent exercising per week if the exercise machine is owned 11 months?

The first step is to use the original regression equation to compute a point estimate for y :

$$\begin{aligned} y &= 9.856 - 0.665(11) \\ &= 9.856 - 7.315 \\ &= 2.541 \end{aligned}$$

For a 90 percent confidence interval, you need to use $t_{05,8}$ which Table 3 (in Appendix B) shows to be 1.860. You have already computed the remaining quantities, so you can proceed with the formula:

$$\begin{aligned} (a, b) &= 2.541 \pm 1.860(1.204) \sqrt{1 + \frac{1}{10} + \frac{(11 - 6.4)^2}{110.40}} \\ &= 2.541 \pm 1.860(1.204) \sqrt{1 + .10 + \frac{21.16}{110.40}} \\ &= 2.541 \pm 1.860(1.204) \sqrt{1.292} \\ &= 2.541 \pm 2.545 \\ &= (-0.004, 5.086) \end{aligned}$$

You can be 90 percent confident that the population mean for the number of hours spent exercising per week when x (number of weeks machine owned) = 11 is between about 0 and 5.

Assumptions and cautions

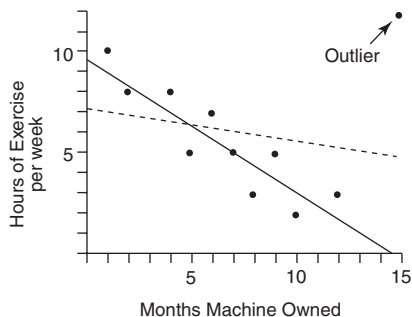
The use of regression for parametric inference assumes that the errors (ϵ) are (1) independent of each other and (2) normally distributed with the same variance for each level of the independent variable. Figure 8-9 shows a violation of the second assumption. The errors (residuals) are greater for higher values of x than for lower values.

Figure 8-9 Data with increasing variance as x increases.



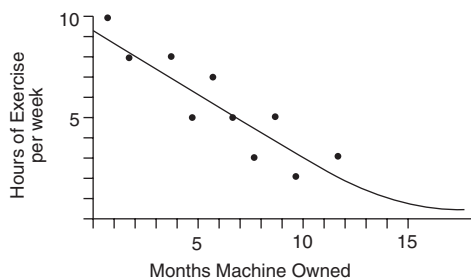
Least squares regression is sensitive to outliers, or data points that fall far from most other points. If you were to add the single data point $x = 15$, $y = 12$ to the exercise data, the regression line would change to the dotted line shown in Figure 8-10. You need to be wary of outliers because they can influence the regression equation greatly.

Figure 8-10 Least squares regression is sensitive to outliers.



It can be dangerous to extrapolate in regression—to predict values beyond the range of your data set. The regression model assumes that the straight line extends to infinity in both directions, which is often not true. According to the regression equation for the example, people who have owned their exercise machines longer than around 15 months do not exercise at all. It is more likely, however, that “hours of exercise” reaches some minimum threshold and then declines only gradually, if at all (see Figure 8-11).

Figure 8-11 Extrapolation beyond the data is dangerous.



Chi-Square (χ^2)

The statistical procedures that you have reviewed thus far are appropriate only for variables at the interval and ratio levels of measurement. The **chi-square** (χ^2) test can be used to evaluate a relationship between two nominal or ordinal variables. It is one example of a **non-parametric test**. Non-parametric tests are used when assumptions about normal distribution in the population cannot be met or when the level of measurement is ordinal or less. These tests are less powerful than parametric tests.

Suppose that 125 children are shown three television commercials for breakfast cereal and are asked to pick which they liked best. The results are shown in Table 8-4.

You would like to know if the choice of favorite commercial was related to whether the child was a boy or a girl or if these two variables are independent. The totals in the margins will allow you to determine the overall probability of (1) liking commercial A, B, or C, regardless of gender, and (2) being either a boy or a girl, regardless of favorite commercial. If the two variables are independent, then you should be able to use these probabilities to predict approximately how many children should be in each cell. If the actual count is very different from the count that you would expect if the probabilities are independent, the two variables must be related.

Table 8-4 Commercial Preference for Boys and Girls

	<i>A</i>	<i>B</i>	<i>C</i>	<i>Totals</i>
Boys	30	29	16	75
Girls	12	33	5	50
Totals	42	62	21	125

Consider the first (upper left) cell of the table. The overall probability of a child in the sample being a boy is $75 \div 125 = .6$. The overall probability of liking Commercial A is $42 \div 125 = .336$. The multiplication rule (covered in Chapter 4) states that the probability of both of two independent events occurring is the product of their two probabilities. Therefore, the probability of a child both being a boy and liking Commercial A is $.6 \times .336 = .202$. The expected number of children in this cell, then, is $.202 \times 125 = 25.2$.

There is a faster way of computing the expected count for each cell: Multiply the row total by the column total and divide by n . The expected count for the first cell is, therefore, $(75 \times 42) \div 125 = 25.2$. If you perform this operation for each cell, you get the expected counts (in parentheses) shown in Table 8-5.

Table 8-5 Chi-Square Results for Table 8-4

	A	B	C	Totals
Boys	30 (25.2)	29 (37.2)	16 (12.6)	75
Girls	12 (16.8)	33 (24.8)	5 (8.4)	50
Totals	42	62	21	125

Note that the expected counts properly add up to the row and column totals. You are now ready for the formula for χ^2 , which compares each cell's actual count to its expected count:

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

The formula describes an operation that is performed on each cell and which yields a number. When all the numbers are summed, the result is χ^2 . Now, compute it for the six cells in the example:

$$\begin{aligned} \chi^2 &= \frac{(30 - 25.2)^2}{25.2} + \frac{(29 - 37.2)^2}{37.2} + \frac{(16 - 12.6)^2}{12.6} + \\ &\quad \frac{(12 - 16.8)^2}{16.8} + \frac{(33 - 24.8)^2}{24.8} + \frac{(5 - 8.4)^2}{8.4} \\ &= .914 + 1.808 + .917 + 1.371 + 2.711 + 1.376 \\ &= 9.097 \end{aligned}$$

The larger χ^2 , the more likely that the variables are related; note that the cells that contribute the most to the resulting statistic are those in which the expected count is very different from the actual count.

Chi-square has a probability distribution, the critical values for which are listed in Table 4 in Appendix B. As with the t -distribution, χ^2 has a degrees-of-freedom parameter, the formula for which is

$$(\# \text{rows} - 1) \times (\# \text{columns} - 1)$$

or in your example:

$$(2 - 1) \times (3 - 1) = 1 \times 2 = 2$$

In Table 4 in Appendix B, a chi-square of 9.097 with two degrees of freedom falls between the commonly used significance levels of .05 and .01. If you had specified an alpha of .05 for the test, you could, therefore, reject the null hypothesis that gender and favorite commercial are independent. At $\alpha = .01$, however, you could not reject the null hypothesis.

The χ^2 test does not allow you to conclude anything more specific than that there is some relationship in your sample between gender and commercial liked (at $\alpha = .05$). Examining the observed vs. expected counts in each cell might give you a clue as to the nature of the relationship and which levels of the variables are involved. For example, Commercial B appears to have been liked more by girls than boys. But χ^2 tests only the very general null hypothesis that the two variables are independent.

Chapter Checkout

Q&A

For the problems 1-3, consider this data on the dose of Medicine X administered and the length of time of continued illness.

Dose	0	1	2	3	4	5	6
Time to cure (days)	15.4	12.2	13.7	9.2	9.9	6.1	4.1

- Make a scatter plot of the data. Is this negative or positive correlation?
 - Calculate the correlation coefficient.
 - Find the significance of the correlation coefficient.
- Apply linear regression to the data to find the slope and intercept.
 - Calculate the 95 percent confidence interval for the slope of your regression.

- 3. a.** Predict the result if the dose was 1.5, and give the 90 percent confidence interval for your prediction.
- b.** Predict the result if the dose was 8, and give the 90 percent confidence interval.

Answers **1a.** negative **b.** -0.9511 **c.** $t = -6.885$, $p < .0005$ **2a.** $\beta_1 = -1.78$; $\beta_0 = 15.42$ **b.** -1.78 ± 0.66 **3a.** 12.76 ± 3.05 **b.** 1.19 ± 3.94

Critical Thinking

- 1.** Imagine that the experiment in problems 1-3 is carried further, and the results are as follows:

Dose	7	8	9	10	11
Time to cure (days)	4.0	3.9	4.2	3.8	4.5

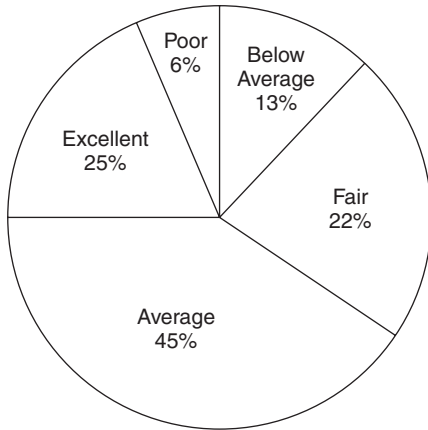
Add this data to your scatter plot and comment on what is happening. What about your prediction in Problem 3b.?

CQR REVIEW

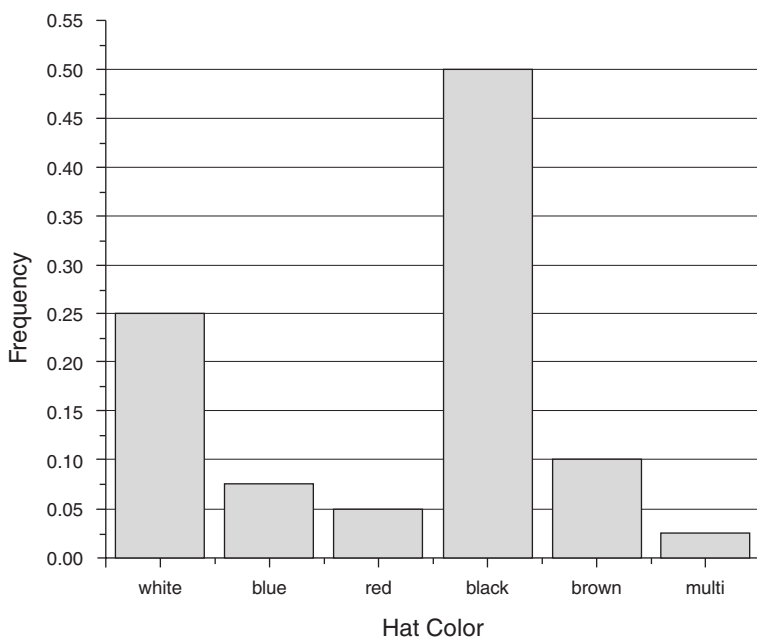
Use this CQR Review to practice what you've learned in this book. After you work through the review questions, you're well on your way to achieving your goal of understanding basic statistical methods.

Chapter 2

1. What is wrong with this pie chart?



2. Study the following frequency histogram of the inventory at a hat store. The total number of hats is 1,000. Then answer these questions:
 - a. How many hats are blue?
 - b. How many hats are white or black?
 - c. If a hat is chosen randomly, what is the chance that it is multi-colored?



3. True or False: “Median” and “mean” are two terms for the same quantity.
4. True or False: “Variance” and “standard deviation” are two terms for the same quantity.

Chapter 4

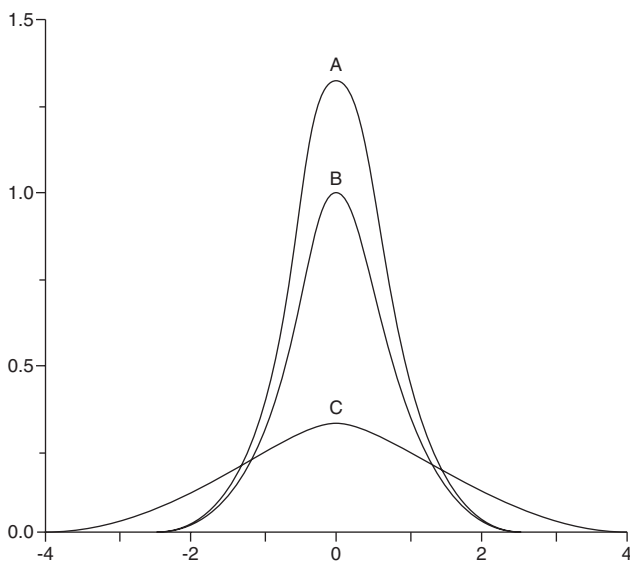
5. True or False: The binomial distribution applies only to events that have two outcomes: “success” or “failure.”
6. A fair die, with six sides, is tossed twice. The chance of rolling two sixes is
 - a. $2/6$
 - b. $1/6$
 - c. $1/12$
 - d. $1/36$
 - e. $1/4$

7. Two fair dice, with six sides each, are rolled. The probability that the sum of the dice equals 12 is
- 0
 - 1
 - $1/6$
 - $1/12$
 - $1/36$
8. Two fair dice, with six sides each, are rolled. The probability that the sum of the dice equals 7 is
- 0
 - 1
 - $1/6$
 - $1/12$
 - $1/36$
9. A population has a mean of 1 and a variance of 4. The best you can say about the probability of randomly choosing a value greater than 7 from this sample is that it is
- 0
 - less than 0.15 percent
 - less than 0.3 percent
 - less than 32 percent
 - less than 16 percent
10. A slot machine has four windows, and each can contain one of eight different symbols. To win the jackpot, all four windows must contain the same symbol. The chance of a jackpot is
- 0
 - $1/8$
 - $(1/8)^3$
 - $1/4$
 - 1

Chapter 5

11. True or False: In a normally distributed population, the standard deviation may never be larger than the mean.
12. True or False: For all practical purposes, μ and \bar{x} are the same thing.

13. You think that the mean of your population falls in the range between -1 and 1 and want to test this idea. Your null hypothesis is H_0 :
- $-1 \geq \mu \geq 1$
 - $-1 \leq \mu \leq 1$
 - $\mu \neq 0$
 - $\mu \neq 1, \mu \neq -1$
 - $-1 \geq \bar{x} \geq 1$
14. Which normal curve pictured below has the greatest standard deviation?



Chapter 6

15. True or False: If two variables are correlated, then they must be causally related.
16. Rejecting the null hypothesis when it is actually true is
- no error
 - a Type I error
 - a Type II error
 - neither a Type I nor a Type II error
 - impossible

Chapter 7

17. A machine for filling beer bottles is designed to put 12 ounces of beer in each bottle, with a standard deviation of 0.05 ounce. How sure can you be that your beer contains at least 11.9 ounces?
18. Suppose that there are only two candidates (A and B) for president and that in a particular state 5,000,000 votes will be cast.
- What percentage of the total votes should a TV station count, to be sure to within ± 1 percent that candidate A has won, at a 95 percent confidence level?
 - The first 20,000 votes are rapidly counted. Candidate A has received 52 percent and candidate B has received 48 percent of the vote. What is the width of the 99.9 percent confidence interval on candidate A's result, as a percentage of the whole 5,000,000 votes?
 - Comment on the methods of projection in part **a**. Is it a good idea to use the first votes cast to project the winner?
19. At a major airport, a researcher studies the average length of delays (excluding flights that leave on time). His data is shown below, for ten randomly chosen days.

Day	1	2	3	4	5	6	7	8	9	10
Average delay (hours)	2.0	1.8	0.32	0.41	2.3	0.72	0.22	0.28	0.57	3.2

- Calculate the mean and standard deviation of the delays, as well as the median.
- What is the 95 percent confidence interval on the average length of delay?
- Compare the mean and median from part **a**. What is the p value for a delay of the median size?
- Do you think this population is normally distributed?

Chapter 8

20. True or False: If the correlation coefficient for a set of data is zero, then there is no relationship between the dependent and independent variables.
21. True or False: The χ^2 test only tells you if two variables are independent.

- 22.** A microbiologist is testing a new antibiotic she developed against a common one. She applies each drug to different plates of bacterial cultures and measures the fraction of bacteria that dies after one day. Her results are given in the following table.

<i>Dose of Drug (mg)</i>	0	0.5	1.0	1.5	2.0	2.5	3.0
New drug (% dead)	2.0	12.3	27.1	32.2	48.9	56.6	60.6
Old drug (% dead)	3.1	6.0	16.5	12.3	25.9	29.0	42.5

Using the information in the preceding table, answer these questions:

- Plot the results for each drug and calculate the coefficient of correlation for each drug.
 - What is the significance level of each of these correlations?
 - Calculate the slope for each plot.
 - Do the 95 percent confidence intervals for the two slopes overlap?
- 23.** Consider the following data on the color of a dog and the average annual income of the dog's owner. Compute a χ^2 value for this data and determine the p -level to which these variables are related.

	\$0-\$25,000	\$25,001-50,000	\$50,001-75,000	\$75,000+	Totals
White	10	12	17	23	62
Black	15	28	13	11	67
Orange/golden	7	9	11	15	42
Mixed	24	26	32	42	124
Total	56	75	73	91	295

Answers: 1. More than 100% is shown. 2a. 75 b. 750 c. 0.025 3. False 4. False 5. True 6. d 7. e 8. c 9. b 10. c 11. False 12. False 13. a 14. c 15. False 16. b 17. 97.72% 18a. 0.192% b. 2.2% 19a. 1.18; 1.06; .65 b. (0.42,1.94) c. $.05 < p < .1$ d. no 20. False 21. True 22a. .995; .949 b. $p < .0005$; $.0005 < p < .001$ c. 20.45; 12.28 d. no 23. $\chi^2 = 17.04$; $.025 < p < .05$.

CQR RESOURCE CENTER

CQR Resource Center offers the best resources available in print and online to help you study and review the core concepts of statistics and statistical methods. You can find additional resources, plus study tips and tools to help test your knowledge, at www.cliffsnotes.com.

Books

There are many books available on basic statistical methods. Many use humor to get across the concepts. Although most of these concepts were covered in *CQR Statistics*, if you're still struggling, one of these may help.

Excel 2000 For Windows For Dummies, by Greg Harvey. Once you've got your data, how can you analyze it efficiently? After all, if you've got more than ten samples, you don't want to be calculating by hand. You need a spreadsheet, and Excel performs good statistical analysis, including all of the basic techniques from *CQR Statistics*. This book shows you how to use Excel effectively and how to perform the statistical tests you are interested in. Hungry Minds, \$21.99.

The Cartoon Guide to Statistics, by Larry Gonick and Woollcott Smith. This book offers much of the same information as *CQR Statistics*—but with cartoons, too! It may sound silly, but it can really help to illustrate the concepts. Harper Collins, \$16.00.

Chances Are: The Only Statistics Book You'll Ever Need, by Steve Slavin. This book is designed to make statistics accessible to people with little mathematical background. If your math skills are holding you up, check this book out. Madison Books, \$16.95.

How to Lie with Statistics, by Darrell Huff. This is a classic, and any list of statistics book would be incomplete without it. Huff shows how badly analyzed and interpreted data can be used to sell products or push agendas, providing good examples of the misuses and abuses of statistics. Worth a look, to keep from repeating the errors he points out. W.W. Norton & Co., \$9.25.

Hungry Minds also has three Web sites that you can visit to read about all the books we publish:

- www.cliffsnotes.com
- www.dummies.com
- www.hungryminds.com

Internet

Visit the following Web sites for more information about statistics:

American Statistical Association—www.amstat.org—This site is really more interesting than you might think. There is a nice discussion of “What statisticians do,” career opportunities in statistics, and also some educational resources. The Association offers online courses and maintains a list of other Web sites related to statistics (click on Center for Statistics Education or Links & Resources).

Portraits of Statisticians—www.york.ac.uk/depts/maths/histstat/people/—If you want to see where all this statistics stuff came from, well, here are pictures or portraits of famous statisticians, past and present. An interesting site put together by the mathematics department at the University of York in the United Kingdom.

Concepts and Applications of Inferential Statistics—<http://faculty.vassar.edu/~lowry/webtext.html>—This site is a complete statistics textbook! If you want to review the concepts in the CQR *Statistics* book or move beyond into some more advanced concepts, this is worth a look. The explanations are clear and examples abound. The site is maintained by Professor Richard Lowry at Vassar College.

Statistical Methods (Index)—<http://bubl.ac.uk/link/s/statisticalmethods.htm>—This is actually a page from the BUBL Catalog on Internet Resources, a larger site maintained by the Centre for Digital Library Research at Strathclyde University in the United Kingdom. On this page, you’ll find many links to pages dedicated to statistical methods. The BUBL page is well organized, with a description of the page that each link leads to. If the Web pages mentioned in CQR *Statistics* material don’t have what you’re looking for, try one of the links on this page.

World Lecture Hall: Statistics—www.utexas.edu/world/lecture/statistics/—From the University of Texas, this site contains links to statistics course Web pages from universities around the world. Some of the links are non-functional, but most lead to a page with lecture notes, exercises, and other statistics material.

Note: The Internet is loaded with statistical information. Most it, however, is descriptive statistics. These pages include all sorts of government statistics—from the U.S. Bureau of Transportation (www.bts.gov), to the Bureau of Justice (www.ojp.usdoj.gov/bjs/). There is a vast amount of information, to which you could apply the statistical methods you learned with this book.

Next time you're on the Internet, don't forget to drop by www.cliffsnotes.com. We created an online Resource Center that you can use today, tomorrow, and beyond.

Glossary

addition rule for mutually exclusive random events, the chance of at least one of them occurring is the sum of their individual probabilities.

alternative hypothesis a research hypothesis; the hypothesis that is supported if the null hypothesis is rejected.

bar chart a graphic that displays how data fall into different categories or groups.

bell-shaped curve symmetrical, single-peaked frequency distribution. Also called the normal curve or gaussian curve.

bias the consistent underestimation or overestimation of a true value, because of preconceived notion of the person sampling the population.

bimodal curve with two equal scores of highest frequency.

binomial event with only two possible outcomes.

binomial probability distribution for binomial events, the frequency of the number of favorable outcomes. For a large number of trials, the binomial distribution approaches the normal distribution.

bivariate involving two variables, especially, when attempting to show a correlation between two variables, the analysis is said to be bivariate.

box plot (box-and-whiskers) a graphic display of data indicating symmetry and central tendency.

Central Limit Theorem a rule that states that the sampling distribution of means from any population will be normal for large sample n .

chi-square a probability distribution used to test the independence of two nominal variables.

class frequency the number of observations that fall into each class interval.

class intervals categories or groups contained in frequency graphics.

coefficient of determination a measure of the proportion of each other's variability that two variables share.

confidence interval the range of values that a population parameter could take at a given level of significance.

confidence level the probability of obtaining a given result by chance.

continuous variable a variable that can be measured with whole numbers and fractional (or decimal) parts thereof.

correlated two (or more) quantities that change together in a consistent manner. Thus, if the value of one variable is known, the other can be immediately determined from their relationship.

correlation coefficient a measure of the degree to which two variables are linearly related.

critical value the value of a computed statistic used as a threshold to decide whether the null hypothesis will be rejected.

data numerical information about variables; the measurements or observations to be analyzed with statistical methods.

degrees of freedom a parameter used to help select the critical value in some probability distributions.

dependent events events such that the outcome of one has an effect on the probability of the outcome of the other.

dependent variable a variable that is caused or influenced by another.

descriptive statistics numerical data that describe phenomena.

deviation the distance of a value in a population (or sample) from the mean value of the population (or sample).

directional test a test of the prediction that one value is higher than another; also called a one-tailed test.

discrete variable a variable that can be measured only by means of whole numbers; or one which assumes only a certain set of definite values, and no others.

disjoint occurrence both outcomes unable to happen at the same time.

distribution a collection of measurements; how scores tend to be dispersed about a measurement scale.

dot plot a graphic that displays the variability in a small set of measures.

double counting a mistake encountered in calculating the probability of at least one of several events occurring, when the events are not mutually exclusive. In this case, the addition rule does not apply.

empirical rule a rule that is founded on observation, without a theoretical basis. Or a “rule of thumb.”

frequency distribution the frequency of occurrence of the values of a variable. For each possible value of the variable, there is an associated frequency with which the variable assumes that value.

frequency histogram a graphic that displays how many measures fall into different classes, giving the frequency at which each category is seen observed.

frequency polygon a graphic presentation of frequency of a phenomenon that typically uses straight lines and points.

grouped data data that has been sorted into categories, usually in order to construct a frequency histogram.

grouped measures a set of values that belong to the same class.

histogram a graphic presentation of frequency of a phenomenon.

independent events events such that the outcome of one has no effect on the probability of the outcome of the other.

independent variable a variable that causes, or influences, another variable.

inference conclusion about a population parameter based upon analysis of a sample statistic. Inferences are always stated with a confidence level.

intercept the value of y at which a line crosses the vertical axis.

interquartile range set of measures lying between the lower quartile (25th percentile) and the upper quartile (75th percentile), inclusive.

interval a scale using numbers to rank order; its intervals are equal but with an arbitrary 0 point.

joint occurrence both outcomes happening simultaneously; $P(AB)$.

least squares any line- or curve-fitting model that minimizes the squared distance of data points to the line.

lower quartile (Q_1), the 25th percentile of a set of measures.

mean the sum of the measures in a distribution divided by the number of measures; the average.

measures of central tendency descriptive measures that indicate the center of a set of values, for example, mean, median, and mode.

measures of variation descriptive measures that indicate the dispersion of a set of values, for example, variance, standard deviation, and standard error of the mean.

median the middle measure in an ordered distribution.

middle quartile (Q_2), the 50th percentile of a set of measures; the median.

mode most frequent measure in a distribution; the high point on a frequency distribution.

mound-shaped curve see *bell-shaped curve*.

multiplication rule the probability of two or more independent (hence, not mutually exclusive) events all occurring is the product of their individual probabilities.

mutually exclusive events such that the occurrence of one precludes the occurrence of the other.

negative relationship a relationship between two variables such that when one increases, the other decreases.

negatively skewed curve a probability or frequency distribution that is not normal, but rather is shifted such that the mean is less than the mode.

nominal a scale using numbers, symbols, or names to designate different subclasses.

non-directional test a test of the prediction that two values are equal or a test that they are not equal; a two-tailed test.

non-parametric test statistical test used when assumptions about normal distribution in the population cannot be met, or when the level of measurement is ordinal or less. For example, the χ -square test.

normal distribution smooth bell-shaped curve symmetrical about the mean such that its shape and area obey the empirical rule.

null hypothesis the reverse of the research hypothesis. The null hypothesis is directly tested by statistical analysis so that it is either rejected or not rejected, with a confidence level. If the null hypothesis is rejected, the alternative hypothesis is supported.

numerical statistics statistical parameters presented as numbers (as opposed to pictorial statistics).

ogive a graphic that displays a running total.

one-tailed test a test of the prediction that one value is higher than another.

ordinal a scale using numbers or symbols to rank order; its intervals are unspecified.

outlier a data point that falls far from most other points; a score extremely divergent from the other measures of a set.

parameter a characteristic of a population. The goal of statistical analysis is usually to estimate population parameters, using statistics from a sample of the population.

Pearson's product moment coefficient identical to the correlation coefficient.

percentile the value in an ordered set of measurements such that $P\%$ of the measures lie below that value.

pictorial statistics statistical parameters that are presented as graphs or charts (as opposed to simply as numbers).

pie chart a graphic that displays parts of the whole, in the form of a circle with its area divided appropriately.

point estimate a number computed from a sample to represent a population parameter.

population a group of phenomena that have something in common. The population is the larger group, whose properties (parameters) are estimated

by taking a smaller sample from within the population, and applying statistical analysis to the sample.

positive relationship a relationship between two variables such that when one increases, the other increases, or when one decreases, the other decreases.

positively skewed curve a probability or frequency distribution that is not normal, but rather is shifted such that the mean is greater than the mode.

power the probability that a test will reject the null hypothesis when it is, in fact, false.

probability a quantitative measure of the chances for a particular outcome or outcomes.

probability distribution a smooth curve indicating the frequency distribution for a continuous random variable.

proportion for a binomial random event, the probability of a successful (or favorable) outcome in a single trial.

qualitative variable phenomenon measured in kind, that is, non-numerical units. For example, color is a qualitative variable, because it cannot be expressed simply as a number.

quantitative variable phenomenon measured in amounts, that is, numerical units. For example, length is a quantitative variable.

random an event for which there is no way to know, before it occurs, what the outcome will be. Instead, only the probabilities of each possible outcome can be stated.

random error error that occurs as a result of sampling variability, through no direct fault of the sampler. It is a reflection of the fact that the sample is smaller than the population; for larger samples, the random error is smaller.

range difference between the largest and smallest measures of a set.

ratio a scale using numbers to rank order; its intervals are equal, and the scale has an absolute 0 point.

region of acceptance the area of a probability curve in which a computed test statistic will lead to acceptance of the null hypothesis.

region of rejection the area of a probability curve in which a computed test statistic will lead to rejection of the null hypothesis.

regression a statistical procedure used to estimate the linear dependence of one or more independent variables on a dependent variable.

relative frequency the ratio of class frequency to total number of measures.

relative frequency principle of probability if a random event is repeated a large number of times, then the proportion of times that a particular outcome occurs is the probability of that outcome occurring in a single event.

research hypothesis a prediction or expectation to be tested. If the null hypothesis is rejected, then the research hypothesis (also called alternative hypothesis) is supported.

residual the vertical distance between a predicted value y and its actual value.

sample a group of members of a population selected to represent that population. A sample to which statistical analysis is applied should be randomly drawn from the population, to avoid bias.

sampling distribution the distribution obtained by computing a statistic for a large number of samples drawn from the same population.

sampling variability the tendency of the same statistic computed from a number of random samples drawn from the same population to differ.

scatter plot a graphic display used to illustrate degree of correlation between two variables.

skewed a distribution displaced at one end of the scale and a tail strung out at the other end.

slope a measure of a line's slant.

standard deviation a measure of data variation; the square root of the variance.

standard error a measure of the random variability of a statistic, such as the mean (i.e., standard error of the mean). The standard error of the mean is equal to the standard deviation divided by the square root of the sample size (n).

standardize to convert to a z -score.

statistic a characteristic of a sample. A statistic is an estimate of a population parameter. For larger samples, the statistic is a better estimate of the parameter.

statistical significance the probability of obtaining a given result by chance. High statistical significance does not necessarily imply importance.

statistics a branch of mathematics that describes and reasons from numerical observations; or descriptive measures of a sample.

stem-and-leaf graphic display that shows actual scores as well as distribution of classes.

symmetry a shape such that one side is the exact mirror image of the other.

symmetric distribution a probability or frequency distribution that has the property in which the mean, median, and mode are all the same value.

systematic error the consistent underestimation or overestimation of a true value, due to poor sampling technique.

***z*-distribution** a probability distribution often used when the population standard deviation is not known or when the sample size is small.

tabled value the value of a computed statistic used as a threshold to decide whether the null hypothesis will be rejected.

test statistic a computed quantity used to decide hypothesis tests.

two-tailed test a test of the prediction that two values are equal, or a test that they are not equal.

Type I error rejecting a null hypothesis that is, in fact, true.

Type II error failing to reject a null hypothesis that is, in fact, false.

upper quartile (Q_3), the 75th percentile of a set of measures.

value a measurement or classification of a variable.

variable an observable characteristic of a phenomenon that can be measured or classified.

variance a measure of data variation; the mean of the squared deviation scores about the means of a distribution.

***z*-score** a unit of measurement obtained by subtracting the mean and dividing by the standard deviation.

Appendix A

COMMON MISTAKES

It may seem that there are many ways to make errors in working a statistics problem. In fact, most errors on statistics exams can be reduced to a short list of common oversights. If you learn to avoid the mistakes listed here, you can greatly reduce your chances of making an error on an exam.

- **Forgetting to convert between standard deviation (σ and s) and variance (σ^2 and s^2).** Some formulas use one; some use the other. Square the standard deviation to get the variance, or take the positive square root of the variance to get the standard deviation.
- **Misstating one-tailed and two-tailed hypotheses.** If the hypothesis predicts simply that one value will be higher than another, it requires a one-tailed test. If, however, it predicts that two values will be different—that is, one value will be either higher *or* lower than another or that they will be equal—then use a two-tailed test. Make sure your null and alternative hypotheses together cover all possibilities—greater than, less than, and equal to.
- **Failing to split the alpha level for two-tailed tests.** If the overall significance level for the test is .05, then you must look up the critical (tabled) value for a probability of .025. The alpha level is always split when computing confidence intervals.
- **Misreading the standard normal (z) table.** As explained in Chapter 5, all standard normal tables do not have the same format, and it is important to know what area of the curve (or probability) the table presents as corresponding to a given z -score. Table 2 in Appendix B gives the area of the curve lying at or below z . The area to the right of z (or the probability of obtaining a value above z) is simply 1 minus the tabled probability.

- **Using n instead of $n - 1$ degrees of freedom in one-sample t -tests.** Remember that you must subtract 1 from n in order to get the degrees-of-freedom parameter that you need in order to look up a value in the t -table.
- **Confusing confidence level with confidence interval.** The **confidence level** is the significance level of the test or the likelihood of obtaining a given result by chance. The **confidence interval** is a range of values between the lowest and highest values that the estimated parameter could take at a given confidence level.
- **Confusing interval width with margin of error.** A confidence interval is always a point estimate plus or minus a margin of error. The interval width is double that margin of error. If, for example, a population parameter is estimated to be 46 percent plus or minus 4 percent, the interval width is 8 percent.
- **Confusing statistics with parameters.** Parameters are characteristics of the population that you usually do not know; they are designated with Greek symbols (μ and σ). Statistics are characteristics of samples that you are usually able to compute. Although statistics correspond to parameters (\bar{x} is the mean of a sample, as μ is the mean of a population), the two are not interchangeable; hence, you need to be careful and know which variables are parameters and which are statistics. You compute statistics in order to estimate parameters.
- **Confusing the addition rule with the multiplication rule.** When determining probability, the multiplication rule applies if all favorable outcomes must occur in a series of events. The addition rule applies when *at least one* success must occur in a series of events.
- **Forgetting that the addition rule applies to mutually exclusive outcomes.** If outcomes can occur together, the probability of their joint occurrence must be subtracted from the total “addition-rule” probability. (See Chapter 4 for more information.)
- **Forgetting to average the two middle values of an even-numbered set when figuring median and (sometimes) quartiles.** If an ordered series contains an even number of measures, the median is always the mean of the two middle measures.

- **Forgetting where to place the points of a frequency polygon.** The points of a frequency polygon are always at the center of each of the class intervals, not at the ends.

Appendix B

TABLES

This appendix includes the following tables:

Table 1. Binomial Probabilities, $P(x)$ for $n \leq 20$

Table 2. Standard Normal Probabilities

Table 3. t Distribution Critical Values

Table 4. χ^2 Critical Values

Table 1 Binomial Probabilities, $P(x)$ for $n \leq 20$

$n=2$											
$x \downarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	$x \uparrow$
0	.9025	.8100	.7225	.6400	.5625	.4900	.4225	.3600	.3025	.2500	2
1	.0950	.1800	.2550	.3200	.3750	.4200	.4550	.4800	.4950	.5000	1
2	.0025	.0100	.0225	.0400	.0625	.0900	.1225	.1600	.2025	.2500	0
0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	0.50	$x \uparrow$
$n=3$											
$x \downarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	$x \uparrow$
0	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664	.1250	3
1	.1354	.2430	.3251	.3840	.4219	.4410	.4436	.4320	.4084	.3750	2
2	.0071	.0270	.0574	.0960	.1406	.1890	.2389	.2880	.3341	.3750	1
3	.0001	.0010	.0034	.0080	.0156	.0270	.0429	.0640	.0911	.1250	0
0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	0.50	$x \uparrow$
$n=4$											
$x \downarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	$x \uparrow$
0	.8145	.6561	.5220	.4096	.3164	.2401	.1785	.1296	.0915	.0625	4
1	.1715	.2916	.3685	.4096	.4219	.4116	.3845	.3456	.2995	.2500	3
2	.0135	.0486	.0975	.1536	.2109	.2646	.3105	.3456	.3675	.3750	2
3	.0005	.0036	.0115	.0256	.0469	.0756	.1115	.1536	.2005	.2500	1
4	.0000	.0001	.0005	.0016	.0039	.0081	.0150	.0256	.0410	.0625	0
0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	0.50	$x \uparrow$

$n=5$

$x \downarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	.7738	.5905	.4437	.3277	.2373	.1681	.1160	.0778	.0503	.0313
1	.2036	.3281	.3915	.4096	.3955	.3602	.3124	.2592	.2059	.1563
2	.0214	.0729	.1382	.2048	.2637	.3087	.3364	.3456	.3369	.3125
3	.0011	.0081	.0244	.0512	.0879	.1323	.1811	.2304	.2757	.3125
4	.0000	.0005	.0022	.0064	.0146	.0284	.0488	.0768	.1128	.1563
5	.0000	.0000	.0001	.0003	.0010	.0024	.0053	.0102	.0185	.0313
	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50
										$x \uparrow$

 $n=6$

$x \downarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	.7351	.5314	.3771	.2621	.1780	.1176	.0754	.0467	.0277	.0156
1	.2321	.3543	.3993	.3932	.3560	.3025	.2437	.1866	.1359	.0938
2	.0305	.0984	.1762	.2458	.2966	.3241	.3280	.3110	.2780	.2344
3	.0021	.0146	.0415	.0819	.1318	.1852	.2355	.2765	.3032	.3125
4	.0001	.0012	.0055	.0154	.0330	.0595	.0951	.1382	.1861	.2344
5	.0000	.0001	.0004	.0015	.0044	.0102	.0205	.0369	.0609	.0938
6	.0000	.0000	.0000	.0001	.0002	.0007	.0018	.0041	.0083	.0156
	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50
										$x \uparrow$

$n=7$

$x \downarrow$	π										$x \uparrow$
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	
0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	.0078	7
1	.2573	.3720	.3960	.3670	.3115	.2471	.1848	.1306	.0872	.0547	6
2	.0406	.1240	.2097	.2753	.3115	.3177	.2985	.2613	.2140	.1641	5
3	.0036	.0230	.0617	.1147	.1730	.2269	.2679	.2903	.2918	.2734	4
4	.0002	.0026	.0109	.0287	.0577	.0972	.1442	.1935	.2388	.2734	3
5	.0000	.0002	.0012	.0043	.0115	.0250	.0466	.0774	.1172	.1641	2
6	.0000	.0000	.0001	.0004	.0013	.0036	.0084	.0172	.0320	.0547	1
7	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0016	.0037	.0078	0
	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	$x \uparrow$

 $n=8$

$x \downarrow$	π										$x \uparrow$
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	
0	.6634	.4305	.2725	.1678	.1001	.0576	.0319	.0168	.0084	.0039	8
1	.2793	.3826	.3847	.3355	.2670	.1977	.1373	.0896	.0548	.0313	7
2	.0515	.1488	.2376	.2936	.3115	.2965	.2587	.2090	.1569	.1094	6
3	.0054	.0331	.0839	.1468	.2076	.2541	.2786	.2787	.2568	.2188	5
4	.0004	.0046	.0185	.0459	.0865	.1361	.1875	.2322	.2627	.2734	4
5	.0000	.0004	.0026	.0092	.0231	.0467	.0808	.1239	.1719	.2188	3
6	.0000	.0000	.0002	.0011	.0038	.0100	.0217	.0413	.0703	.1094	2
7	.0000	.0000	.0000	.0001	.0004	.0012	.0033	.0079	.0164	.0313	1
8	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0017	.0039	0
	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	$x \uparrow$

$n=9$

$x \downarrow$	π										$x \uparrow$
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	
0	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020	9
1	.2985	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176	8
2	.0629	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703	7
3	.0077	.0446	.1069	.1762	.2336	.2668	.2716	.2508	.2119	.1641	6
4	.0006	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461	5
5	.0000	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2461	4
6	.0000	.0001	.0006	.0028	.0087	.0210	.0424	.0743	.1160	.1641	3
7	.0000	.0000	.0000	.0003	.0012	.0039	.0098	.0212	.0407	.0703	2
8	.0000	.0000	.0000	.0000	.0001	.0004	.0013	.0035	.0083	.0176	1
9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0020	0
0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50		$x \uparrow$

$n=10$

$x \downarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	
0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010	10
1	.3151	.3874	.3474	.2684	.1877	.1211	.0725	.0403	.0207	.0098	9
2	.0746	.1937	.2759	.3020	.2816	.2335	.1757	.1209	.0763	.0439	8
3	.0105	.0574	.1298	.2013	.2503	.2668	.2522	.2150	.1665	.1172	7
4	.0010	.0112	.0401	.0881	.1460	.2001	.2377	.2508	.2384	.2051	6
5	.0001	.0015	.0085	.0264	.0584	.1029	.1536	.2007	.2340	.2461	5
6	.0000	.0001	.0012	.0055	.0162	.0368	.0689	.1115	.1596	.2051	4
7	.0000	.0000	.0001	.0008	.0031	.0090	.0212	.0425	.0746	.1172	3
8	.0000	.0000	.0000	.0001	.0004	.0014	.0043	.0106	.0229	.0439	2
9	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016	.0042	.0098	1
10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	0
	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	$x \uparrow$

$n=12$

$x \downarrow$	π												$x \uparrow$
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50			
0	.5404	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008	.0002	.0002	12	
1	.3413	.3766	.3012	.2062	.1267	.0712	.0368	.0174	.0075	.0029	.0029	11	
2	.0988	.2301	.2924	.2835	.2323	.1678	.1088	.0639	.0339	.0161	.0161	10	
3	.0173	.0852	.1720	.2362	.2581	.2397	.1954	.1419	.0923	.0537	.0537	9	
4	.0021	.0213	.0683	.1329	.1936	.2311	.2367	.2128	.1700	.1208	.1208	8	
5	.0002	.0038	.0193	.0532	.1032	.1585	.2039	.2270	.2225	.1934	.1934	7	
6	.0000	.0005	.0040	.0155	.0401	.0792	.1281	.1766	.2124	.2256	.2256	6	
7	.0000	.0000	.0006	.0033	.0115	.0291	.0591	.1009	.1489	.1934	.1934	5	
8	.0000	.0000	.0001	.0005	.0024	.0078	.0199	.0420	.0762	.1208	.1208	4	
9	.0000	.0000	.0000	.0001	.0004	.0015	.0048	.0125	.0277	.0537	.0537	3	
10	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0025	.0068	.0161	.0161	2	
11	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0029	.0029	1	
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0002	0	
	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	0.50	$x \uparrow$	

$n=14$

$x \downarrow$	π														$x \uparrow$
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50					
0	.4877	.2288	.1028	.0440	.0178	.0068	.0024	.0008	.0002	.0001	14				
1	.3593	.3559	.2539	.1539	.0832	.0407	.0181	.0073	.0027	.0009	13				
2	.1229	.2570	.2912	.2501	.1802	.1134	.0634	.0317	.0141	.0056	12				
3	.0259	.1142	.2056	.2501	.2402	.1943	.1366	.0845	.0462	.0222	11				
4	.0037	.0349	.0998	.1720	.2202	.2290	.2022	.1549	.1040	.0611	10				
5	.0004	.0078	.0352	.0860	.1468	.1963	.2178	.2066	.1701	.1222	9				
6	.0000	.0013	.0093	.0322	.0734	.1262	.1759	.2066	.2088	.1833	8				
7	.0000	.0002	.0019	.0092	.0280	.0618	.1082	.1574	.1952	.2095	7				
8	.0000	.0000	.0003	.0020	.0082	.0232	.0510	.0918	.1398	.1833	6				
9	.0000	.0000	.0000	.0003	.0018	.0066	.0183	.0408	.0762	.1222	5				
10	.0000	.0000	.0000	.0000	.0003	.0014	.0049	.0136	.0312	.0611	4				
11	.0000	.0000	.0000	.0000	.0000	.0002	.0010	.0033	.0093	.0222	3				
12	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0019	.0056	2				
13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0009	1				
14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	0				
	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	$x \uparrow$				

$n=16$

$x \downarrow$	π															$x \uparrow$	
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50							
0	.4401	.1853	.0743	.0281	.0100	.0033	.0010	.0003	.0001	.0000	16						
1	.3706	.3294	.2097	.1126	.0535	.0228	.0087	.0030	.0009	.0002	15						
2	.1463	.2745	.2775	.2111	.1336	.0732	.0353	.0150	.0056	.0018	14						
3	.0359	.1423	.2285	.2463	.2079	.1465	.0888	.0468	.0215	.0085	13						
4	.0061	.0514	.1311	.2001	.2252	.2040	.1553	.1014	.0572	.0278	12						
5	.0008	.0137	.0555	.1201	.1802	.2099	.2008	.1623	.1123	.0667	11						
6	.0001	.0028	.0180	.0550	.1101	.1649	.1982	.1983	.1684	.1222	10						
7	.0000	.0004	.0045	.0197	.0524	.1010	.1524	.1889	.1969	.1746	9						
8	.0000	.0001	.0009	.0055	.0197	.0487	.0923	.1417	.1812	.1964	8						
9	.0000	.0000	.0001	.0012	.0058	.0185	.0442	.0840	.1318	.1746	7						
10	.0000	.0000	.0000	.0002	.0014	.0056	.0167	.0392	.0755	.1222	6						
11	.0000	.0000	.0000	.0000	.0002	.0013	.0049	.0142	.0337	.0667	5						
12	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0040	.0115	.0278	4						
13	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0029	.0085	3						
14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018	2						
15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0090	.0001	.0002	1						
	.095	.090	.085	.080	.075	.070	.065	.060	.055	.050							

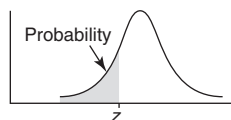
$n=18$

$x \downarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	$x \uparrow$
0	.3972	.1501	.0536	.0180	.0056	.0016	.0004	.0001	.0000	.0000	18
1	.3763	.3002	.1704	.0811	.0338	.0126	.0042	.0012	.0003	.0001	17
2	.1683	.2835	.2556	.1723	.0958	.0458	.0190	.0069	.0022	.0006	16
3	.0473	.1680	.2406	.2297	.1704	.1046	.0547	.0246	.0095	.0031	15
4	.0093	.0700	.1592	.2153	.2130	.1681	.1104	.0614	.0291	.0117	14
5	.0014	.0218	.0787	.1507	.1988	.2017	.1664	.1146	.0666	.0327	13
6	.0002	.0052	.0301	.0816	.1436	.1873	.1941	.1655	.1181	.0708	12
7	.0000	.0010	.0091	.0350	.0820	.1376	.1792	.1892	.1657	.1214	11
8	.0000	.0002	.0022	.0120	.0376	.0811	.1327	.1734	.1864	.1669	10
9	.0000	.0000	.0004	.0033	.0139	.0386	.0794	.1284	.1694	.1855	9
10	.0000	.0000	.0001	.0008	.0042	.0149	.0385	.0771	.1248	.1669	8
11	.0000	.0000	.0000	.0001	.0010	.0046	.0151	.0374	.0742	.1214	7
12	.0000	.0000	.0000	.0000	.0002	.0012	.0047	.0145	.0354	.0708	6
13	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0045	.0134	.0327	5
14	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0039	.0117	4
15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0009	.0031	3
16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006	2
17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	1
	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	

$n=20$

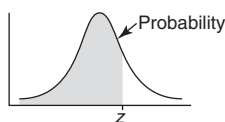
$x \downarrow$	π										$x \uparrow$
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	
0	.3585	.1216	.0388	.0115	.0032	.0008	.0002	.0000	.0000	.0000	20
1	.3774	.2702	.1368	.0576	.0211	.0068	.0020	.0005	.0001	.0000	19
2	.1887	.2852	.2293	.1369	.0669	.0278	.0100	.0031	.0008	.0002	18
3	.0596	.1901	.2428	.2054	.1339	.0716	.0323	.0123	.0040	.0011	17
4	.0133	.0898	.1821	.2182	.1897	.1304	.0738	.0350	.0139	.0046	16
5	.0022	.0319	.1028	.1746	.2023	.1789	.1272	.0746	.0365	.0148	15
6	.0003	.0089	.0454	.1091	.1686	.1916	.1712	.1244	.0746	.0370	14
7	.0000	.0020	.0160	.0545	.1124	.1643	.1844	.1659	.1221	.0739	13
8	.0000	.0004	.0046	.0222	.0609	.1144	.1614	.1797	.1623	.1201	12
9	.0000	.0001	.0011	.0074	.0271	.0654	.1158	.1597	.1771	.1602	11
10	.0000	.0000	.0002	.0020	.0099	.0308	.0686	.1171	.1593	.1762	10
11	.0000	.0000	.0000	.0005	.0030	.0120	.0336	.0710	.1185	.1602	9
12	.0000	.0000	.0000	.0001	.0008	.0039	.0136	.0355	.0727	.1201	8
13	.0000	.0000	.0000	.0000	.0002	.0010	.0045	.0146	.0366	.0739	7
14	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0049	.0150	.0370	6
15	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0049	.0148	5
16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0046	4
17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	3
18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	2
0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50		

Table 2 Standard Normal Probabilities

 (Table entry is probability at or below z .)


z	.00	.01	.02	.03	.04
- 3.4	.0003	.0003	.0003	.0003	.0003
- 3.3	.0005	.0005	.0005	.0004	.0004
- 3.2	.0007	.0007	.0006	.0006	.0006
- 3.1	.0010	.0009	.0009	.0009	.0008
- 3.0	.0013	.0013	.0013	.0012	.0012
- 2.9	.0019	.0018	.0018	.0017	.0016
- 2.8	.0026	.0025	.0024	.0023	.0023
- 2.7	.0035	.0034	.0033	.0032	.0031
- 2.6	.0047	.0045	.0044	.0043	.0041
- 2.5	.0062	.0060	.0059	.0057	.0055
- 2.4	.0082	.0080	.0078	.0075	.0073
- 2.3	.0107	.0104	.0102	.0099	.0096
- 2.2	.0139	.0136	.0132	.0129	.0125
- 2.1	.0179	.0174	.0170	.0166	.0162
- 2.0	.0228	.0222	.0217	.0212	.0207
- 1.9	.0287	.0281	.0274	.0268	.0262
- 1.8	.0359	.0351	.0344	.0336	.0329
- 1.7	.0446	.0436	.0427	.0418	.0409
- 1.6	.0548	.0537	.0526	.0516	.0505
- 1.5	.0668	.0655	.0643	.0630	.0618
- 1.4	.0808	.0793	.0778	.0764	.0749
- 1.3	.0968	.0951	.0934	.0918	.0901
- 1.2	.1151	.1131	.1112	.1093	.1075
- 1.1	.1357	.1335	.1314	.1292	.1271
- 1.0	.1587	.1562	.1539	.1515	.1492
- 0.9	.1841	.1814	.1788	.1762	.1736
- 0.8	.2119	.2090	.2061	.2033	.2005
- 0.7	.2420	.2389	.2358	.2327	.2296
- 0.6	.2743	.2709	.2676	.2643	.2611
- 0.5	.3085	.3050	.3015	.2981	.2946
- 0.4	.3446	.3409	.3372	.3336	.3300
- 0.3	.3821	.3783	.3745	.3707	.3669
- 0.2	.4207	.4168	.4129	.4090	.4052
- 0.1	.4602	.4562	.4522	.4483	.4443
- 0.0	.5000	.4960	.4920	.4880	.4840

<i>.05</i>	<i>.06</i>	<i>.07</i>	<i>.08</i>	<i>.09</i>	<i>z</i>
.0003	.0003	.0003	.0003	.0002	- 3.4
.0004	.0004	.0004	.0004	.0003	- 3.3
.0006	.0006	.0005	.0005	.0005	- 3.2
.0008	.0008	.0008	.0007	.0007	- 3.1
.0011	.0011	.0011	.0010	.0010	- 3.0
.0016	.0015	.0015	.0014	.0014	- 2.9
.0022	.0021	.0021	.0020	.0019	- 2.8
.0030	.0029	.0028	.0027	.0026	- 2.7
.0040	.0039	.0038	.0037	.0036	- 2.6
.0054	.0052	.0051	.0049	.0048	- 2.5
.0071	.0069	.0068	.0066	.0064	- 2.4
.0094	.0091	.0089	.0087	.0084	- 2.3
.0122	.0119	.0116	.0113	.0110	- 2.2
.0158	.0154	.0150	.0146	.0143	- 2.1
.0202	.0197	.0192	.0188	.0183	- 2.0
.0256	.0250	.0244	.0239	.0233	- 1.9
.0322	.0314	.0307	.0301	.0294	- 1.8
.0401	.0392	.0384	.0375	.0367	- 1.7
.0495	.0485	.0475	.0465	.0455	- 1.6
.0606	.0594	.0582	.0571	.0559	- 1.5
.0735	.0721	.0708	.0694	.0681	- 1.4
.0885	.0869	.0853	.0838	.0823	- 1.3
.1056	.1038	.1020	.1003	.0985	- 1.2
.1251	.1230	.1210	.1190	.1170	- 1.1
.1469	.1446	.1423	.1401	.1379	- 1.0
.1711	.1685	.1660	.1635	.1611	- 0.9
.1977	.1949	.1922	.1894	.1867	- 0.8
.2266	.2236	.2206	.2177	.2148	- 0.7
.2578	.2546	.2514	.2483	.2451	- 0.6
.2912	.2877	.2843	.2810	.2776	- 0.5
.3264	.3228	.3192	.3156	.3121	- 0.4
.3632	.3594	.3557	.3520	.3483	- 0.3
.4013	.3974	.3936	.3897	.3859	- 0.2
.4404	.4364	.4325	.4286	.4247	- 0.1
.4801	.4761	.4721	.4681	.4641	- 0.0

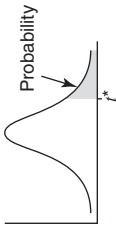


<i>z</i>	<i>.00</i>	<i>.01</i>	<i>.02</i>	<i>.03</i>	<i>.04</i>
0.0	.5000	.5040	.5080	.5120	.5160
0.1	.5398	.5438	.5478	.5517	.5557
0.2	.5793	.5832	.5871	.5910	.5948
0.3	.6179	.6217	.6255	.6293	.6331
0.4	.6554	.6591	.6628	.6664	.6700
0.5	.6915	.6950	.6985	.7019	.7054
0.6	.7257	.7291	.7324	.7357	.7389
0.7	.7580	.7611	.7642	.7673	.7704
0.8	.7881	.7910	.7939	.7967	.7995
0.9	.8159	.8186	.8212	.8238	.8264
1.0	.8413	.8438	.8461	.8485	.8508
1.1	.8643	.8665	.8686	.8708	.8729
1.2	.8849	.8869	.8888	.8907	.8925
1.3	.9032	.9049	.9066	.9082	.9099
1.4	.9192	.9207	.9222	.9236	.9251
1.5	.9332	.9345	.9357	.9370	.9382
1.6	.9452	.9463	.9474	.9484	.9495
1.7	.9554	.9564	.9573	.9582	.9591
1.8	.9641	.9649	.9656	.9664	.9671
1.9	.9713	.9719	.9726	.9732	.9738
2.0	.9772	.9778	.9783	.9788	.9793
2.1	.9821	.9826	.9830	.9834	.9838
2.2	.9861	.9864	.9868	.9871	.9875
2.3	.9893	.9896	.9898	.9901	.9904
2.4	.9918	.9920	.9922	.9925	.9927
2.5	.9938	.9940	.9941	.9943	.9945
2.6	.9953	.9955	.9956	.9957	.9959
2.7	.9965	.9966	.9967	.9968	.9969
2.8	.9974	.9975	.9976	.9977	.9977
2.9	.9981	.9982	.9982	.9983	.9984
3.0	.9987	.9987	.9987	.9988	.9988
3.1	.9990	.9991	.9991	.9991	.9992
3.2	.9993	.9993	.9994	.9994	.9994
3.3	.9995	.9995	.9995	.9996	.9996
3.4	.9997	.9997	.9997	.9997	.9997

<i>.05</i>	<i>.06</i>	<i>.07</i>	<i>.08</i>	<i>.09</i>	<i>z</i>
.5199	.5239	.5279	.5319	.5359	0.0
.5596	.5636	.5675	.5714	.5753	0.1
.5987	.6026	.6064	.6103	.6141	0.2
.6368	.6406	.6443	.6480	.6517	0.3
.6736	.6772	.6808	.6844	.6879	0.4
.7088	.7123	.7157	.7190	.7224	0.5
.7422	.7454	.7486	.7517	.7549	0.6
.7734	.7764	.7794	.7823	.7852	0.7
.8023	.8051	.8078	.8106	.8133	0.8
.8289	.8315	.8340	.8365	.8389	0.9
.8531	.8554	.8577	.8599	.8621	1.0
.8749	.8770	.8790	.8810	.8830	1.1
.8944	.8962	.8980	.8997	.9015	1.2
.9115	.9131	.9147	.9162	.9177	1.3
.9265	.9279	.9292	.9306	.9319	1.4
.9394	.9406	.9418	.9429	.9441	1.5
.9505	.9515	.9525	.9535	.9545	1.6
.9599	.9608	.9616	.9625	.9633	1.7
.9678	.9686	.9693	.9699	.9706	1.8
.9744	.9750	.9756	.9761	.9767	1.9
.9798	.9803	.9808	.9812	.9817	2.0
.9842	.9846	.9850	.9854	.9857	2.1
.9878	.9881	.9884	.9887	.9890	2.2
.9906	.9909	.9911	.9913	.9916	2.3
.9929	.9931	.9932	.9934	.9936	2.4
.9946	.9948	.9949	.9951	.9952	2.5
.9960	.9961	.9962	.9963	.9964	2.6
.9970	.9971	.9972	.9973	.9974	2.7
.9978	.9979	.9979	.9980	.9981	2.8
.9984	.9985	.9985	.9986	.9986	2.9
.9989	.9989	.9989	.9990	.9990	3.0
.9992	.9992	.9992	.9993	.9993	3.1
.9994	.9994	.9995	.9995	.9995	3.2
.9996	.9996	.9996	.9996	.9997	3.3
.9997	.9997	.9997	.9997	.9998	3.4

Table 3 *t* Distribution Critical Values

(Table entry is the point t^* with given probability p lying above it.)

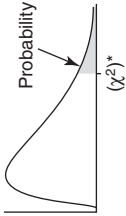


<i>df</i>	<i>p</i>											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015

17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291

Table 4 χ^2 Critical Values

(Table entry is the point χ^2 with given probability p lying above it.)



df	p											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72

16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2

Index

A

- addition rule, 41, 115, 145
- alternative hypothesis, 62, 145
- assumptions, regression, 108–109
- average, 24. *See also* mean
- axes
 - frequency distribution, 15
 - histograms, 14
 - intercept, 102–103
 - slope, 103

B

- bar charts, 8, 9–10, 21, 145
- bell-shaped curves, 16, 145
- bias (systematic errors), 53, 145
- bimodal curves, 17, 145
- binomial events, 145
- binomial probabilities, 118–127, 145
- binomial tables, 47
- binomials
 - discrete variables, 46–48
 - normal approximation, 58–60
- bivariate relationships, 145
 - correlation, 96–102
- box plots, 18–20, 145
- box-and-whiskers, 18–20

C

- Central Limit Theorem, 54, 145
- central tendency. *See* measures of central tendency
- charts
 - bar charts, 8, 9–10
 - box plots, 18–20
 - dot plots, 10
 - frequency distributions, 15–17
 - frequency polygons, 14–15
 - ogive, 11–12
 - pie charts, 8, 10–11
 - scatter plots, 20–21
 - stem-and-leaf, 17–18

- chi-square (χ^2), 110–112, 145
 - critical values, 134–135
- class frequency, 145
- class intervals, 145
- classes, histograms
 - frequency, 13
 - intervals, 12
- classic theory of probability, 37–38
- coefficient of determination, 102
- conditional probability, 44
- confidence intervals, 70–72, 145
 - populations, 78, 82
 - regression, 107–108
 - slope, 105–107
- confidence levels, 145
- continuous variables, probability distribution and, 46, 145
- correlation, 96–102, 145
- correlation coefficient, 145. *See also* Pearson's product moment coefficient
- critical value, test statistic, 64, 145
- curves
 - frequency distribution, 16–17
 - frequency distribution, bimodal, 17
 - negatively skewed, 147
 - positively skewed, 148
 - variation and, 31–32

D

- data, 8, 146
- dependent events, probability of, 39–40, 146
- dependent variables, 146
- descriptive statistics, 146
- deviation from the mean, 146
 - percentile, 33
 - standard deviation, 31
 - variation, 30–31
- df (degrees of freedom), 80, 115, 146
- directional tests, 66–67, 146
- discrete variables, 45–48, 146
- disjoint outcomes, 38, 146
- distributions, 52–53, 146
- dot plot charts, 10, 146
- double counting, probability and, 42–43, 146

E

- empirical rule, 146
- errors
 - random, 53
 - systematic, 53
 - Type I, 150
 - Type II, 150
- estimates, hypotheses, 72–73

F

- failures, binomial variables, 58
- fair coins, probability and, 7
- fixed coins, probability and, 7
- formulae
 - central tendency, 25
 - shortcuts, variance/deviation, 32–33
- frequency distributions, 15–17, 146
- frequency histogram, 12–13, 146
- frequency polygon, 14–15, 116, 146
- frequency, class frequency, 145

G

- graphics, pictorial statistics, 3
- graphs, histograms, 8. *See also* charts
- grouped data, 25–28, 146
- grouped measures, 146
 - mean and, 25–26

H

- histograms, 8, 146
 - frequency histogram, 12–13
 - relative frequency histograms, 13–14
- horizontal bar charts, 8
- hypotheses
 - alternative hypothesis, 62
 - confidence intervals, 70–72
 - estimating difference score, 72–73
 - null hypothesis, 62, 63
 - point estimates, 70–72
 - research hypothesis, 62–63
 - stating, 62–63
 - statistical significance, 69–70
 - test statistic, 63–66
 - Type I errors, 68–69, 150
 - Type II errors, 68–69, 150
 - univariate tests, 73–75

I

- independent events, probability of, 39, 146
- independent variables, 146
- inference, 4, 146
- inferential tests, univariate, 76–95
- intercept, 102–103, 146
- interval scale, 34, 147
- IQR (interquartile range), 18, 147
 - deviation, 34
 - variation, 34

J

- joint occurrences, probability of, 40–41, 147
- J-shaped curves, 16–17

L

- least squares, regression, 105, 147
- lower quartile (Q_1), 147
 - box plots, 18, 19
 - deviation, 33–34
 - variation, 33–34

M

- margin of error, 115
- mean, 3, 147. *See also* deviation from the mean
 - binomials, 47
 - central tendency, 23, 24
 - comparing, t -tests, 84–87
 - comparing, z -tests, 83–84
 - grouped measures, 25–26
 - standard error, 54
 - symmetric distribution, 28
- measurement scales, 34–35
- measures of central tendency, 21–29, 147
- measures of variation, 29–34, 147
- median value, 115, 147
 - box plots, 18, 19
 - central tendency, 23, 24–25
 - grouped data, 26–28
 - outliers and, 24
 - symmetric distribution, 28
- middle quartile (Q_2), 147
 - deviation, 33–34
 - variation, 33–34
- mode, 147
 - central tendency, 23, 25
 - symmetric distribution, 28

mound-shaped curves, 16, 147
 multiplication rule, joint occurrences and,
 40–41, 115, 147
 mutually exclusive outcomes, 38, 115, 147

N

negative relationships, 21, 147
 negatively skewed curves, 147
 nominal scale, 34, 147
 non-directional tests, 66–67, 147
 non-mutually exclusive outcomes, 42
 non-parametric tests, 110–112, 147
 normal approximation to binomials, 58–60
 normal curves, 31–32
 properties, 54–58
 z -score, 56–57
 normal distributions, 31–32, 147
 normal tables, 114
 null hypothesis, 62, 63, 147
 numerical statistics, 3, 148

O

ogive, 11–12, 22, 148
 one-sample t -test, 79–82
 one-sample z -test, 76–79
 one-tailed tests, 66–67, 114, 148
 ordinal scale, 34, 148
 organization, 4
 outliers, 148
 central tendency, 24
 stem-and-leaf, 17

P

paired difference t -test, 87–89
 parameters, 50–51, 115, 148
 Pearson's product moment coefficient (r),
 98–102, 148
 percentile, 33, 148
 pictorial statistics, 3–4, 148
 pictures. *See* graphs, histograms
 pie charts, 8, 10–11, 148
 point estimates, hypotheses, 70–72, 148
 pooled variance method, population, 86–87
 population, 5, 50–51, 148
 Central Limit Theorem, 54
 confidence interval, 78, 82
 single population proportion, 89–91
 positive relationships, 21, 148
 positively skewed curves, 29, 148
 power, hypothesis errors, 69, 148

predictions, 5–6
 confidence level, 107–108
 probability, 6–7, 37–38, 148
 conditional, 44
 dependent events, 39–40
 double counting, 42–43
 independent events, 39
 joint occurrences, 40–41
 non-mutually exclusive outcomes, 42
 relative frequency principle, 38, 149
 simple events, 38–39
 probability distributions, 44–45, 148
 continuous variables, 46
 discrete variables, 45–46
 process of statistics gathering, 4
 properties of normal curve, 54–58
 proportion, 148
 binomial experiments, 46
 proportional frequency, histograms, 14

Q

qualitative data, 15, 148
 quantitative data, 15, 148
 quartiles, 115

R

r (Pearson's product moment), 98–102
 random error, 53, 149
 randomness, 148
 samples, 51
 ranges, 149
 histograms, 12
 variation, 30
 rank order, ordinal scale, 34
 ratio scale, 35, 38, 149
 region of acceptance, 65, 149
 region of rejection, 65, 149
 regression, 102–109, 149
 relative frequency histograms, 13–14, 149
 relative frequency principle of probability,
 38, 149
 research hypothesis, 62–63, 149
 residual, regression, 105, 149
 resources, 142–144
 results comparison, 6

S

samples, 5, 50–51, 149
 central limit theorem, 54
 sampling distributions, 52–53, 149

sampling variability, 52–53, 149
scales. *See* measurement scales
scatter plots, 20–21, 149
 correlation, 96–97
single population proportion, 89–91
skewed curves, 16, 28, 149
slope, 103, 105–107, 149
spread, 30. *See also* measures of variation
square root, 31
standard deviation, 31–32, 47, 114, 149
standard error, 54, 149
standard normal probabilities, 128–131
standard normal tables, 114
standardization, 149
statistical insurance methods, 4
statistical significance, hypotheses,
 69–70, 150
statistics, 50–51, 115, 149, 150
 descriptive statistics, 146
 samples, 51
 types, 3–4
stem-and-leaf chart, 17–18, 150
successes, binomial variables, 58
sum notation, 25
symmetric distribution, 28, 150
symmetry, 150
 box plot, 18
 curves, 16
systematic error, 53, 150

T

t -distribution, 79, 150
 critical values, 132–133
 t -tests, 79–89, 115
tabled value, 64, 150
test statistics, 63–67, 150
testing, 73–95
two-tailed tests, 66–67, 114, 150
Type I errors, 68–69, 150
Type II errors, 68–69, 150

U

univariate tests, 73–75
upper quartile (Q_3), 150
 box plots, 18, 19
 deviation, 34
 variation, 34

V

values, 150
 axes, 15
 pairs, interval scale, 34–35
variables, 150
variance, 114, 150
 variation, 30–31
variation. *See also* measures of variation
 lower quartile (Q_1), 33–34
 middle quartile (Q_2), 33–34
vertical bar charts, 9

W

Web sites, 143–144

X

x axis, 15
 slope, 103

Y

y axis, 15
 slope, 102–103

Z

z -score, 150
 normal curves and, 56–57
 z -tests
 comparing two means, 83–84
 one-sample, 76–79