Instructor's Guide
with Solutions

for

Moore, McCabe and Craig's

Introduction to the Practice of Statistics

Sixth Edition

Darryl K. Nester
Bluffton University

W.H. Freeman and Company
New York

© 2009 by W.H. Freeman and Company

All rights reserved.

Printed in the United States of America.

First printing

ISBN: 1-4292-1472-4
EAN: 978-1-4292-1472-8

W. H. Freeman and Company
41 Madison Avenue
New York, NY 10010
Houndmills, Basingstoke RG21 6XS, England

CONTENTS

Preface v
1 TO THE INSTRUCTOR 1
Philosophy and Goals 1
Calculators and Computers 3
Course Design 4
General Comments 7
Teaching Suggestions 9
Using Video 10
2 CHAPTER COMMENTS 16
Chapter 1: Looking at Data—Distributions 16
Chapter 2: Looking at Data—Relationships 20
Chapter 3: Producing Data 22
Chapter 4: Probability: The Study of Randomness 23
Chapter 5: From Probability to Inference 25
Chapter 6: Introduction to Inference 26
Chapter 7: Inference for Distributions 28
Chapter 8: Inference for Proportions 30
Chapter 9: Analysis of Two-Way Tables 31
Chapter 10: Inference for Regression 32
Chapter 11: Multiple Regression 33
Chapter 12: One-Way Analysis of Variance 33
Chapter 13: Two-Way Analysis of Variance 34
Chapter 14: Logistic Regression 34
Chapter 15: Nonparametric Tests 34
Chapter 16: Bootstrap Methods and Permutation Tests 35
Chapter 17: Statistics for Quality: Control and Capability 35
3 SAMPLE EXAMINATIONS 36
Sample examination I 36
Sample examination I solutions 38
Sample examination II 39
Sample examination II solutions 42
Sample final examination 44
Sample final examination solutions 47

4 SOLUTIONS TO EXERCISES 51
About these solutions 51
Acknowledgments 51
Using the table of random digits 51
Using statistical software 52
Chapter 1 Solutions 54
Chapter 2 Solutions 90
Chapter 3 Solutions 129
Chapter 4 Solutions 149
Chapter 5 Solutions 169
Chapter 6 Solutions 180
Chapter 7 Solutions 198
Chapter 8 Solutions 232
Chapter 9 Solutions 246
Chapter 10 Solutions 266
Chapter 11 Solutions 289
Chapter 12 Solutions 324
Chapter 13 Solutions 349
Chapter 14 Solutions 376
Chapter 15 Solutions 387
Chapter 16 Solutions 404
Chapter 17 Solutions 461

Preface

This Instructor's Guide tries to make it easier to teach from the sixth edition of *Introduction to the Practice of Statistics* (*IPS*). The most helpful material is without doubt the full solutions of all exercises, which make up the second part of the Guide. The first part of the Guide contains additional examples, sample examinations, and suggestions for using the *Against All Odds* and *Decisions Through Data* videos in class. It also contains brief discussions of the approach we have taken in *IPS* for presenting statistics to beginning students. Reasonable people can (and do) differ about the nature of basic instruction in statistics. Our comments in this Guide are one side of a many-sided conversation. Your position in this conversation will no doubt differ at some points. We hope we have at least made clear the reasons for our choices.

We welcome comments, suggestions for improvement of *IPS*, and reports of errors that escaped detection and can be fixed in future printings. You can reach us as follows:

Bruce Craig
Department of Statistics
Purdue University
West Lafayette, IN 47907-1267
Telephone: (765) 494-6043
Fax: (765) 494-0558
e-mail: bcraig@stat.purdue.edu

George P. McCabe
Department of Statistics
Purdue University
West Lafayette, IN 47907-1267
(765) 494-6047
(765) 494-0558
mccabe@stat.purdue.edu

Notes, comments, and corrections about this Instructor's Guide—especially the solutions—should be sent to:

Darryl K. Nester
Department of Mathematics
Bluffton University
Bluffton, OH 45817-2104
Telephone: (419) 358-3483
Fax: (419) 358-3074
e-mail: nesterd@bluffton.edu

1 TO THE INSTRUCTOR

Philosophy and Goals

The text is intended to be used in a first course in statistics for students with limited mathematical background. It covers the basic material presented in many courses at this level—data analysis, a little probability, and standard statistical methods. The emphasis is on understanding how to use statistics to address real problems. In this regard it differs from most introductions.

Many students view statistics (indeed, all of the mathematical sciences) as a collection of formulas that give correct results if the computations are performed accurately. The competence of the student is measured by the complexity of the formula that can be computed correctly. Unfortunately, this view is often reinforced by the way statistics is traditionally taught.

In contemporary practice, computations in statistics and other technical fields are always automated. Students, like users, should ideally employ statistical software, graphing calculators, or spreadsheets for computations and graphics. Even if the available technology is limited, students must prepare for reading and using statistical arguments with understanding. Our focus as teachers should be to help students use statistical tools and reasoning effectively by providing conceptual understanding and practical strategies. We are also free to move closer to the practice of statistics as a tool for learning about the real world.

Consider, for example, a pre-election poll taken for a major news organization. The results are reported as percentages with a margin of error. Typically, stratification is used, and perhaps even cluster sampling. The formulas for the standard error are complex. The most important issues for understanding and using the results, however, do not require knowing the details: How was the question phrased? How was the sample designed? What about undercoverage and nonresponse? What does the margin of error tell us in straightforward language? If students are able to analyze and interpret data produced by a simple random sample, then the transfer of the basic ideas to the more complex situation is easy. The meaning of a confidence interval, for example, does not depend on the formula used to calculate the standard error.

We describe in the Preface an emerging consensus on the nature of a first course among statisticians concerned with teaching, and we try to express that consensus in our writing. As a reminder, here are important ways in which *IPS* differs from more traditional texts:

- There is more attention to data analysis. Chapters 1 and 2 give quite full coverage. It is now becoming common to emphasize data in a first course, but many texts still begin with a brief treatment of "descriptive statistics." Note that we introduce distributions, specifically the Normal distributions, in Chapter 1 as models for the overall pattern of data long before probability appears. This is legitimate in itself,

and it helps break up the indigestible lump of formal probability that is a traditional stumbling block in learning statistics. Note also that the descriptive aspects of several-variable data (scatterplots, correlation, least-squares lines) are presented as part of data analysis rather than late in the book in conjunction with inference in these settings. We think that this order has important advantages. It emphasizes the centrality of data analysis in statistics and gives students a chance to build fluency in the strategies of data analysis in both one- and several-variable settings. It allows examples and exercises on relationships among variables that don't meet the narrower requirements of inference.

- There is more attention to designing data production. It is surprising to a practicing statistician how little attention these ideas, among the most influential aspects of statistics, receive in most first courses. Chapter 3 discusses sampling and experimental design, with attention to some of the practical issues involved. The importance of deliberate use of randomness and informal discussion of the distribution of results in repeated sampling motivate the introduction of probability and the very important idea of a sampling distribution.

- The introduction to probability in Chapter 4 begins with experience with randomness and is designed to allow instructors to cover only the essentials or to present the full traditional material by omitting or including the optional Section 4.5. For many groups of students, more statistics is a worthy exchange for less probability. Some of the most important probability facts (law of large numbers, central limit theorem) appear in the discussion of sampling distributions in Chapter 5. Some traditional details about binomial distributions are marked as optional. The criterion for optional material in Chapters 4 and 5 is simple: It is not required for the statistical inference presented later in the text.

- There is more emphasis on the reasoning of inference. Chapter 6 discusses this reasoning in some depth, with an optional section 6.4 comparing the "P-value" and "two types of error" approaches to testing. We emphasize P-values. This reflects common practice and helps students understand the output of statistical software.

- There is more attention to statistics in practice. Realism may be too much to claim in a beginning text. Nonetheless, we often give longer-than-usual discussions of examples in order to connect textbook procedures with real-world practice.

Upon completion of a course based on *IPS*, students should be able to think critically about data, to select and use graphical and numerical summaries, to apply standard statistical inference procedures, and to draw conclusions from such analyses. They are ready for more specialized statistics courses (such as applied regression or quality control); for "research methods" courses in many fields of study; and for projects, reports, or employment that require basic data analysis.

Calculators and Computers

The practice of statistics requires a good deal of graphing and numerical calculation. Doing some graphing and calculating "by hand" may build understanding of methods. On the other hand, graphics and calculations are always automated in statistical practice. Moreover, struggling with computational aspects of a procedure often interferes with a full understanding of the concepts. Students are easily frustrated by their inability to complete problems correctly. Automating the arithmetic greatly improves their ability to complete problems. We therefore favor automating calculations and graphics as much as your resources and setting allow.

All students should have a calculator that does "two-variable statistics," that is, that calculates not only $\bar{x}$ and s but the correlation r and the least-squares regression line from keyed-in data. Even if you use computer software, students should have a calculator for use at home and on exams. Two-variable statistics calculators sell for less than $20. We don't discuss anachronistic "computing formulas" (e.g., for standard deviation and correlation) that presuppose a four-function calculator.

IPS is designed to be used in courses where students have access to computing facilities, and also where they do not have such access. Many of the examples used in the text involve computations that can readily be performed with a two-variable statistics calculator rather than a computer. Each chapter contains exercises with small amounts of data or with the results of some computations given that are suitable for students without a computer. Although the use of a computer for a course based on *IPS* is not required, most people who *do* statistics today use a computer. Any serious attempt to introduce the practice of statistics must recognize this fact. You will find examples of output from a variety of technological tools (statistical software, spreadsheets, graphing calculators) throughout the text. The output appears because any statistics student should become accustomed to looking at output and using the fruits of the study of statistics to recognize terms and results in output from any source.

As computing continues to become cheaper and more readily available, and as statistics software becomes more "user-friendly," use of computers in courses at this level will increase. Typical students have some prior acquaintance with personal computers. For students with this background, modern statistical software speeds work and allows more elaborate work without the added burden that learning software once presented. We encourage use of software, and some exercises are not realistic without it.

As a practical encouragement to software use, the CD-ROM packaged with the text includes most of the data (all of the larger sets of data) used in exercises in *IPS*. The data appear as plain text (ASCII) files and also formatted for Excel, JMP, Minitab, SPSS, S-PLUS, and TI-83 or -84 calculators. The files are named by their location in the book:

`eg01_004.txt` = Example 1.4 (plain text)
`ex01_013.xls` = Exercise 1.13 (Excel format)
`ta01_002.mtp` = Table 1.2 (Minitab format)

Course Design

Think of the content of *IPS* this way: The nonoptional parts of Chapters 1–8 form a core. Chapters 9 (Inference for Two-Way Tables), 10 (Inference for Regression), 12 (One-Way ANOVA), 15 (Nonparametric Tests), and 16 (Bootstrap Methods) offer a selection of further topics that are independent of each other. Depending on the length of your course and the quantitative background of your students, you will probably elect one or more of these. Chapters 11 (Multiple Regression), 13 (Two-Way ANOVA), 14 (Logistic Regression), and 17 (Process Control and Capability) are brief introductions to more advanced topics. Each of these chapters requires at least one of the elective chapters as a prerequisite, and all go better with software.

IPS is an elementary but serious introduction to modern statistics for general college audiences. This means that the material is presented in a way that it can be learned by students who are not particularly skillful in mathematics. *IPS* does require more attentive reading (words, not equations) than some other texts. The amount of time required for learning—and hence, the amount of material that can be covered—depends on the skill level and (more important) the maturity of the students in a particular course.

We have personally used *IPS* for general undergraduates in nontechnical disciplines (agriculture, health sciences, consumer studies, etc.); for graduate students in nonquantitative disciplines (education, social sciences, retailing, etc.) who will be analyzing data for their research; and as a first course for majors in actuarial science, mathematics, and statistics. These are quite varied groups. Because general undergraduates are taking a required course they would not otherwise elect to take, motivation is the big problem in teaching them. More data and less formal probability are helpful. Graduate students often have a weak quantitative background but are mature and hard-working. Many of them will take another course in statistics, but some will not. They need some exposure to analysis of count data, regression, and analysis of variance. Accomplishing this for students with moderate mathematics skills requires skipping or treating lightly many of the optional sections and encouraging a great deal of work on the part of the students. The practical relevance of the examples and exercises is an important factor for sustaining the motivation and drive that these students bring with them to the first lecture. Our course for potential undergraduate majors is software- and data-intensive. We end with multiple regression, including some material not in *IPS*.

The sample outlines for courses using *IPS* are aids for instructors, not strict rules. You should adapt the pace and extent of your course to your students.

Outline 1. This is a semester course for students with low or moderate skills and motivation. We suggest this outline for general, less quantitative undergraduates. For students with moderate mathematics skills, you can add a little more depth for some topics and assign or discuss some of the more challenging exercises. We are always tempted to go faster than this outline suggests, and when we do we find that the students don't come with us. In particular, each exam, viewed as an opportunity to solidify learning, uses a week. We suggest the following: Spend one class on active review. Consider distributing a sample exam in class and asking the students to work on the first problem for about 5 minutes—long enough to determine whether they know how to approach it. Then, discuss that problem together. Continue through the sample exam in this manner. Working under conditions similar to an exam concentrates the mind. The exam itself occupies a second class period, and returning and discussing it fills most of a third. This isn't lost time: exams are learning tools. We recommend omitting all starred subsections for this audience, but use your judgment.

Week	Material to be covered
1	Chapter 1, Section 1, Section 2 – measures of center
2	Chapter 1, Sections 2 and 3
3	Chapter 2, Sections 1, 2, and 3
4	Chapter 2, Sections 4 and 5
5	Review Chapters 1 and 2; Exam I on Chapters 1 and 2
6	Chapter 3, Sections 1, 2, and 3
7	Chapter 3, Section 4; Chapter 4, Sections 1, 2, and 3
8	Chapter 4, Section 4; Chapter 5, Section 1
9	Chapter 5, Section 2
10	Review Chapters 3, 4, and 5; Exam II on Chapters 3, 4, and 5
11	Chapter 6, Section 1; start Section 2
12	Chapter 6, Sections 2 and 3
13	Chapter 7, Sections 1 and 2
14	Chapter 8, Sections 1 and 2
15	Review and extended examples
Comprehensive final exam	

Outline 2. A semester course for students with either higher skills or stronger motivation. We follow this outline for nonquantitative graduate students. Omit optional material except where noted.

Week	Material to be covered
1	Chapter 1, Sections 1, 2, and 3
2	Chapter 2, Sections 1, 2, and 3
3	Chapter 2, Sections 4 and 5
4	Chapter 3, Sections 1, 2, 3, and 4
5	Review Chapters 1 to 3; Exam I on Chapters 1 to 3
6	Chapter 4, Sections 1, 2, and 3
7	Chapter 4, Section 4; Chapter 5, Section 1
8	Chapter 5, Section 2
9	Review Chapters 4 and 5; Exam II on Chapters 4 and 5
10	Chapter 6, Sections 1 and 2
11	Chapter 6, Sections 2 and 3, power from Section 4
12	Chapter 7, Sections 1, 2, and 3
13	Chapters 8 and 9
14	Chapter 12
15	Review and extended examples
Comprehensive final exam	

Outline 3. A semester course for students with adequate quantitative skills who may need more probability than the previous outlines offer. As outlined, ending with multiple regression, the course is quite ambitious. Use software to accelerate the later chapters.

Week	Material to be covered
1	Chapter 1, Sections 1, 2, and 3
2	Chapter 2, Sections 1, 2, and 3
3	Chapter 2, Sections 4 and 5
4	Chapter 3, Sections 1, 2, 3, and 4
5	Exam I on Chapters 1 to 3; Chapter 4, Sections 1 and 2
6	Chapter 4, Sections 3 and 4
7	Chapter 4, Section 5
8	Chapter 5, Sections 1 (including optional parts) and 2
9	Chapter 6, Sections 1 and 2
10	Chapter 6, Sections 2 and 3; start Chapter 7
11	Chapter 7, Sections 1, 2, and 3
12	Exam on Chapters 4 to 7; start Chapter 8
13	Chapters 8 and 9
14	Chapter 10
15	Chapter 11
Comprehensive final exam	

Outline 4. A one-quarter course for students with low to moderate skills. It covers the essentials, ending with the *t* procedures (and the introduction to inference in practice) in Chapter 7 of *IPS*. This provides a good foundation for any of a number of more specialized courses in statistics or related areas. This outline is quite tentative; the number of class meetings per week varies in institutions using the quarter system. You may be able to assign more than is shown. The main point of this outline is the goal of completing Chapter 7 in even the briefest introduction. *IPS* contains sufficient material for a sequence of two-quarter courses, particularly taking into account the supplementary chapters. You have considerable flexibility in choosing from Chapters 8–17 for the second quarter.

Week	Material to be covered
1	Chapter 1, Sections 1 and 2
2	Chapter 1, Section 3; Chapter 2, Section 1
3	Chapter 2, Sections 2, 3, 4, and 5
4	Chapter 3, Sections 1, 2, and 3; Review Chapters 1 to 3
5	Exam on Chapters 1 to 3; Chapter 4, Sections 1 and 2
6	Chapter 4, Sections 3 and 4
7	Chapter 5, Sections 1 and 2
8	Chapter 6, Sections 1 and 2
9	Chapter 6, Section 3; Chapter 7, Section 1
10	Chapter 7, Sections 2 and 3
Comprehensive final exam	

General Comments

When students complete a course based on *IPS* they should be able to analyze real data and draw conclusions. Since the emphasis is on *doing* statistics rather than talking about statistics, much of the class time should be spent discussing how statistical analysis tells us something about data. The methodologies presented are all based upon sound statistical theory. Details of how the theory leads to particular formulas are neither interesting to the students nor are they essential to the central purpose of learning to do statistics. A minimum amount of time should be spent on examining formulas and procedures in the abstract. Theoretical ideas need to be understood in the context of how they *apply* to the analysis of data.

Consider, for example, the problem of outliers. They are a fact of life for anyone applying statistics to real problems. It is not sufficient to take a strictly mathematical view and assume that they do not exist. A procedure based on Normal distributions may give very misleading conclusions if applied to data contaminated by outliers. In the practice of statistics it is important to identify outliers and, with an understanding of the area of application, to do something about them. On the other hand, it is not important at this level to know that a *t* distribution can be represented at the ratio of a standard Normal to the square root of a chi-square.

To teach students how to use statistical methods with understanding, you will spend most of your class time doing statistics. This means discussing data, not just numbers devoid of any real meaning. It means saying something about the context in which the data were collected and the field of application. Instructors in statistics courses typically are comfortable with the rather elementary mathematical and statistical ideas needed to discuss the methods presented in the text. On the other hand, few of us are experts in the various substantive fields from which the examples and exercises are drawn. Most of these, however, require only a basic understanding, most of which is presented in stating the example or exercise.

Using data from real problems provides many opportunities for the students to become actively involved in the class. Incomplete knowledge of all of the fields covered by the examples and exercises can be turned into a very effective teaching tool. Students have diverse backgrounds and often know a great deal more than we think. It is a very pleasant experience to leave a class session having learned something from the students. Once, when discussing corn yield data, George was unable to remember whether or not the fields were irrigated, but he did remember that they were in Nebraska. He mentioned this, and one woman in the class immediately raised her hand and explained that if it was corn in Nebraska, then it certainly was irrigated. This sort of participation occurs naturally in a teaching environment where the students are encouraged to think about data rather than numbers.

If we could offer just one piece of advice to teachers using *IPS*, it would be this: **A number or a graph, or a formula such as "Reject H_0," is not an adequate answer to a statistical problem.** Insist that students state a brief conclusion in the context of the specific problem setting. We are dealing with data, not just with numbers.

There is a tendency today to try to design statistics courses specifically for students in one particular major or general field. This is particularly true for students in business and the social and behavioral sciences. Even in general courses some students express the opinion that they would prefer examples and homework based on problems coming exclusively from their own field of interest. If we are attempting to educate rather than to simply train students, this view is shortsighted. It overlooks, for example, the fact that most students will change careers several times during their working lives. Ideas learned well in one context are easily translated to others. Barry Bonds' 73 home runs do have something in common with the agricultural production in a drought year. One of the interesting things about statistics as a field is that the same fundamental ideas occur in many diverse settings. Different settings facilitate learning of the big ideas. Furthermore, many fields are not as homogeneous as they might appear at a first glance. Often, interesting problems require the assimilation of material from diverse areas. This is particularly true of business. In summary, do not apologize to your students because all of the examples are not from their field. Explain to them that breadth is part of education.

There are enough examples in the text for some to be used in class lectures. Sometimes it is useful to build a lecture around one or a sequence of exercises. If computers are used, it is a good idea to run some of the examples or exercises on the system you are using. You

can distribute copies of output, enlarge the output on a transparency, or ask students to do specific computer work in advance and bring the results to class. Many classrooms now allow you to do the work in real time with projection for the class to follow. If you teach in a computer lab, we recommend doing examples along with the class so that students follow on their machines and can see your projected screen if they get lost.

Teaching Suggestions

Statistics has a well-earned reputation as the dullest of subjects. Tell someone that you teach statistics and a typical reaction is something like, "I took a course in statistics once. . . . I just barely made it through." We teachers have the responsibility of overcoming that preconception by demonstrating to students that our subject is both intellectually stimulating and useful. We believe that almost any statistics course can be improved by including more data and more emphasis on reasoning, at the expense of fewer recipes and less theory. **"Data" means not just numbers, but problems set in a context that require a conclusion or discussion rather than just a calculation or graph.** In *IPS* we have tried to emphasize data and reasoning, limited of course by our own ability and by the need to present an exposition accessible to beginning students. We hope the text will at least not stand in the way of teachers who want their students to come away with more than a list of recipes. Here are some suggestions from our experience.

Involve the class Mixing blackboard work with direct address (get out from behind the lectern) helps keep students awake. Even better is discussion with the class. Use some exercises as bases for discussion, assigning them "for discussion" in the next class rather than requiring written answers. Ask a student to present an analysis and then give a nontechnical conclusion, as if reporting to a boss. Then, ask the other students to respond with questions or comments.

Leading discussions is interesting but difficult. You must resist the temptation to leap in with the "correct" view. Try rather to guide the class by questions and to get other students to offer alternatives to weak answers. Be patient when asking questions on the spot. An educational psychologist of our acquaintance notes that teachers rarely wait long enough for students to assimilate a question and produce a response. He suggests 30 seconds, which seems an eternity while passing in silence but does produce more response in the end. Above all, don't put down a student who is incorrect or confused. Students notice even the silent disappointment that contrasts so clearly with your response to an intelligent answer. Try to give a positive response to every student who is brave enough to speak up—they are helping you. When you know the class better, you can direct easier questions at the weaker students, thus preventing monopoly by the bright and aggressive. Remember that you are building confidence, not simply conveying information. Discussion isn't easy for teachers oriented toward problem-solution or theorem-proof presentation. But the attempt is essential if your students are afraid of statistics and perceive it as remote from their experience.

Problem assignments The practice of statistics is learned by doing statistics rather than by reading about it. To master the material in *IPS*, students need to work exercises—many of them. Note that brief answers for the odd-numbered exercises appear in the back of the text. These allow students to check their work, but they are often not complete as responses to the problems because they lack graphics or detailed conclusions.

Ask during the session previous to the session in which homework is due if students are having difficulty with any problems. Most students will not have started the homework yet—you are trying to encourage them to start early. Always ask about special difficulties when the assignment is turned in and spend time clarifying them. If necessary, comment on any additional difficulties when you return the assignments. Don't be overly concerned about "giving away" homework answers. Homework is to help learning; the exams will detect those who didn't do their own work.

Remember that your students don't think like mathematicians Good teaching requires an awareness of who your students are. Statistics courses are not mathematics classes, so don't emphasize mathematical fine points or mathematical explanations and derivations. You think mathematically, but your students don't. Instead, teach by examples and add comments on statistics in practice when appropriate. Try sometimes to begin with a motivating example rather than always presenting a technique, then illustrate it by an example. Don't just repeat the text. Outline the material (students like an outline style), explain how each point fits into the overall course, then interpret the text. Point out the tough parts, use different examples, say things a different way, try to explain the intuition behind the formal material. Mathematical understanding is not the only kind of conceptual understanding. Don't always do the same thing: vary your style over the class period and from day to day.

Using Video

One of the most effective ways to convince your students that statistics is useful is to show them real people (not professors) using statistics in a variety of settings. Video allows you to do this in the classroom. Two related video series (available on VHS tapes as well as DVDs) that contain many short documentaries of statistics in use on location are:

- *Against All Odds: Inside Statistics* (*AAO*). This telecourse, consisting of 26 half-hour programs, was prepared by COMAP for the Annenberg/Corporation for Public Broadcasting Project. More information can be found at `www.learner.org` (click on "View Programs"), or call 1-800-LEARNER.
- *Statistics: Decisions Through Data* (*DTD*). This set of 21 shorter modules (5 hours total) is intended for use as a classroom supplement in secondary schools. It was prepared by COMAP for the National Science Foundation and draws on the location segments of *AAO*. It is available from COMAP. Visit `www.comap.com/` or call 1-800-77-COMAP for information.

If you are outside the United States, you can obtain information about both video series by contacting:

COMAP Inc.
Suite 210
57 Bedford Street
Lexington, MA 02173 USA
Fax 1-617-863-1202

Because David Moore was the content developer for these video series, they fit the style and sequence of *IPS* well. In several cases, data from the videos appear in the text. We do not recommend showing complete programs from *AAO* in the classroom. The shorter modules from *DTD* are more suitable for classroom use. Video is a poor medium for exposition, and it leaves viewers passive. It is therefore generally not a good substitute for a live teacher. We suggest regular showing of selected on-location stories from *AAO* or *DTD* in most classrooms, rather than full programs. If you have a large lecture (several hundred students), however, full *DTD* video modules along with computer demonstrations will help hold an audience too large for personal interaction.

Video has several strengths that make short segments an ideal supplement to your own teaching. Television can bring real users of statistics and their settings into the classroom. And psychologists find that television communicates emotionally rather than rationally, so it is a vehicle for changing attitudes. One of our goals in teaching basic statistics is to change students' attitudes about the subject. Because video helps do this, consider showing video segments regularly even if you do not think they help students learn the specific topic of that class period.

Here are some specific suggestions for excerpts from *AAO* and *DTD* that work effectively in class:

- The 14-minute video *What Is Statistics?* is a good way to start a course. This collage of examples from *AAO* forms part of the first unit of *AAO* and is the first module of *DTD*. It is available separately (and inexpensively) from:

  The American Statistical Association
  1429 Duke Street
  Alexandria, VA 22314 USA
  (703) 684-1221 or `www.amstat.org`

- *Lightning research* from Program 2 of *AAO*, Module 3 of *DTD*. A study of lightning in Colorado discovers interesting facts from a histogram.

- *Calories in hot dogs* from Program 3 of *AAO*, Module 5 of *DTD*. The five-number summary and box plots compare beef, meat, and poultry hot dogs.

- *The Boston Beanstalk Club* from Program 4 of *AAO*, Module 7 of *DTD*. This social club for tall people leads to discussion of the 68–95–99.7 rule for Normal distributions.

- *Saving the manatees* from Program 8 of *AAO*, Module 11 of *DTD*. There is a strong linear relation between the number of power boats registered in Florida and the number of manatees killed by boats.
- *Obesity and metabolism* from Program 8 of *AAO*, Module 12 of *DTD*, looks at the linear relationship between lean body mass and metabolic rate in the context of a study of obesity.
- *Sampling at Frito-Lay* from Program 13 of *AAO*, Module 17 of *DTD*, illustrates the many uses of sampling in the context of making and selling potato chips. A student favorite.
- *The Physicians' Health Study* from Program 12 of *AAO*, Module 15 in *DTD*, is a major clinical trial (aspirin and heart attacks) that introduces design of experiments.
- *Sampling distributions* are perhaps the single most important idea for student understanding of inference. Module 19 of *DTD* presents the general idea, the basic facts about the sampling distribution of the sample mean $\bar{x}$, and the application of these ideas to an $\bar{x}$ control chart. The setting is a highly automated AT&T electronics factory.
- *Battery lifetimes* from Program 19 of *AAO* lead to an animated graphic that illustrates the behavior of confidence intervals in repeated sampling. Module 20 in *DTD* is a presentation of the reasoning of confidence intervals using the same setting that can be shown in its entirety.
- *Taste testing of colas* is the setting for an exposition of the reasoning of significance tests in Module 21 of *DTD*. This treatment is preferable to that in *AAO*.
- *Welfare reform in Baltimore,* from Program 22 of *AAO*, is a comparative study of new and existing welfare systems that leads to a two-sample comparison of means.
- *The Salem witchcraft trials,* revisited in Program 23 of *AAO*, show social and economic differences between accused and accusers via comparison of proportions.
- *Medical practice:* Does the treatment women receive from doctors vary with age? This story in Program 24 of *AAO* produces a two-way table of counts.
- *The Hubble constant* relates velocity to distance among extragalactic objects and is a key to assessing the age of the expanding universe. A story in Program 25 of *AAO* uses the attempt to estimate the Hubble constant to introduce inference about the slope of a regression line.

Here is a complete list of the documentary segments in *AAO*, with timings for use if your VCR measures real time, along with ratings from one to four stars. This handy guide was prepared by Professor Edward R. Mansfield of the University of Alabama. We are grateful to him for permission to reproduce it here. Start your VCR timer when the first signal on the tape appears. (If you have DVD copies of these programs, the ratings will be more useful thatn the timings.) Remember that *AAO* programs are packaged two to a tape; the timings for the even-numbered programs may need some adjustment because the gap between programs seems to vary a bit.

Program 1: What Is Statistics?

- 4:48 Domino's Pizza ***
- 13:15 The "What is statistics?" collage of later examples

Program 2: Picturing Distributions

- 33:55 When does lightning strike? ***
- 44:02 TV programming and demographics ***
- 52:17 Diagnostic-related groups **

Program 3: Describing Distributions

- 5:55 Comparable worth in Colorado Springs *
- 16:07 Calories in hot dogs **
- 21:00 Musical analysis of urine data **

Program 4: Normal Distributions

- 33:50 Age distributions and Social Security *
- 46:07 Boston Beanstalk social club for tall people *
- 50:38 Why don't baseball players hit .400 anymore? ***

Program 5: Normal Calculations

- 7:07 Auto emissions at GM Proving Ground *
- 14:10 Cholesterol values **
- 19:50 Sizes of military uniforms **

Program 6: Time Series

- 34:50 The body's internal clock *
- 43:48 Psychology: reaction time study *

Program 7: Models for Growth

- 3:00 Children's growth rates and hormone treatment ***
- 14:00 Gypsy moth infestations **

Program 8: Describing Relationships

- 32:25 Manatees versus motorboats in Florida ***
- 37:55 Cavities versus fluoride levels
- 39:31 1970 draft lottery ***
- 44:04 Obesity: metabolic rate versus lean body mass *

Program 9: Correlation

5:42 Identical twins raised apart ***
16:22 Baseball players' salaries **
20:53 The Coleman Report (education in the 1960s) *

Program 10: Multidimensional Data Analysis

32:28 Chesapeake Bay pollution **
47:42 Bellcore graphics **

Program 11: The Question of Causation

5:42 Simpson's paradox ****
12:47 Smoking and cancer (historical survey) ***

Program 12: Experimental Design

32:46 Observational study of lobster behavior *
36:14 Physicians' Health Study: aspirin and heart attacks ****
43:39 Is Ribavirin too good to be true? ***
47:22 Police response to domestic violence *

Program 13: Blocking and Sampling

4:45 Strawberry field research *
13:28 Undercounting in the Census ***
20:48 Sampling potato chips at Frito-Lay ****

Program 14: Samples and Surveys

41:21 National Opinion Research Center ****

Program 15: What Is Probability?

10:50 Persi Diaconis on randomness *
17:49 Traffic control in New York (simulation model) **

Program 16: Random Variables

33:36 Cheating on AP calculus *
34:33 Space shuttle *Challenger* disaster ****
43:02 Points in a professional basketball game
49:10 Earthquakes in California *

Program 17: Binomial Distributions

3:46 The "hot hand": free throws in basketball ***
9:45 A finance class experiment **
17:22 Sickle-cell anemia *
24:25 Quincunx: falling balls **

Program 18: The Sample Mean and Control Charts

- 33:45 Roulette
- 35:04 Interviews with gamblers **
- 40:44 The casino always wins ****
- 47:03 Control charts at Frito-Lay ***
- 53:41 W. Edwards Deming ****

Program 19: Confidence Intervals

- 11:35 Duracell batteries **
- 18.25 Rhesus monkeys in medical studies *
- 21:21 Feeding behavior of marmosets

Program 20: Significance Tests

- 34:18 Is this poem by Shakespeare? **
- 49:06 Discrimination within the FBI ***

Program 21: Inference for One Mean

- 5:55 National Institute of Standards and Technology **
- 13:30 Taste testing of cola ***
- 21:08 Autism *

Program 22: Comparing Two Means

- 33:32 Welfare programs in Baltimore **
- 45:05 Product development at Union Carbide ***
- 51:00 SAT exams: Can coaching help?

Program 23: Inference for Proportions

- 3:03 Measuring unemployment (Bureau of Labor Statistics) *
- 11:58 Safety of drinking water ***
- 20:15 The Salem witch trials

Program 24: Inference for Two-Way Tables

- 34:11 Ancient humans (markings on teeth) **
- 43:30 Does breast cancer treatment vary by age? **
- 52:02 Mendel's peas **

Program 25: Inference for Relationships

- 3:32 How fast is the universe expanding (Edwin Hubble)? ****

Program 26: Case Study

- 35:49 How AZT for treatment of AIDS was tested ***

2 CHAPTER COMMENTS

Along with the text, some instructors might wish to incorporate some of the supplemental material, available for download at `www.whfreeman.com/ips6e`. These topics—transforming relationships and decision analysis—were removed from the main text for reasons of length. They would most naturally fit (respectively) at the end of Chapter 2 and following Exercise 4.100 on page 301 in Chapter 4.

Chapter 1: Looking at Data—Distributions

Students taking a first course in statistics often do not know what to expect. Some may view statistics as a field where the major task is to tabulate large collections of numbers accurately. Others have heard that statistics is more like mathematics, with a lot of complicated formulas that are difficult to use. Few are expecting a course where they need to use their common sense and to *think*.

The presentation of the material in Chapter 1 sets the tone for the entire course. Statistics is a subject where intelligent judgments are needed and common sense plays a large role. To the extent that you can get the students to think deeply about the material at this stage, you will have succeeded in setting the proper tone. Most of the examples presented in the text can serve as the basis for extensive class discussions. Depending on the interests of the students, some of these may be more suitable than others. You can also try to find or collect data that the particular students you are teaching can relate to.

In one class based on this text, students were asked to take their pulse on the first day of class. No particular instructions concerning how to take one's pulse were given. The data that resulted were used in discussions that illustrated several major themes of the course. A stemplot was immediately constructed. There was one very large outlier—180 beats per minute. We discussed whether or not this value was reasonable. Several of the physical education students commented that someone who was very out of shape might elevate their pulse to such a high level by running up to the third floor where the course was held. Someone from psychology mentioned that a person who was very anxious about the course might have an elevated pulse. One conclusion drawn was that, if you want to measure anything on people, you must pay attention to the circumstances present when the measurement is taken.

Further examination of the data revealed an exceptionally large number of readings that ended with the digit zero. Students in aerobics classes typically take their pulse for 6 seconds and multiple by 10. This led to a discussion of the possibility that different methods were used to obtain the reported pulse rates. Since it was impractical to ask all 50 or so students what method each used, a random sample was taken using the class list and the table of random numbers given in the back of the text (Table B). Students had no trouble following the method. Successive pairs of digits were used and numbers outside the range 1 to 50 were discarded. It was not surprising to find a number corresponding to a

student who was not present in the class on that day. A short discussion about missing data and the problems associated with statistics in the *real* world followed.

The methods reported by the sampled students were very revealing. Most were what you would expect. Measure for 60 seconds, measure for 30 seconds and multiply by 2, measure for 15 seconds and multiply by 4, and of course, the method of those who had taken an aerobics course. One said that he did not have a watch so the value given was a guess. Another (a very polite foreign student whose command of spoken English was not very good) explained that she had no idea about what was asked and simply reported the value given by another student.

The lesson was quite clear. It is not easy to gather good data. What is measured and how it is measured are the two most important questions to ask when undertaking any statistical analysis. Understanding the issues related to these questions does not require mathematical sophistication or the ability to calculate $\sqrt{1/n_1 + 1/n_2}$ by hand. If the students can be sufficiently motivated at the beginning of the course and understand that many interesting things can be learned by looking at data, they will be willing to put in the effort required to master the skills needed for the computations required later.

Look for sets of data that will be interesting to your students. They are not very difficult to find. If the students are concentrated in a few majors, ask faculty from those departments for reprints of articles using elementary statistical methods. Ask questions about how things are measured and whom or what is measured. Do the items or people measured represent any larger group or do they simply represent themselves? Usually, we think of statistics as being used only for inference to some larger known or imagined population. However, many interesting sets of data do not fall into this category, and much about data analysis can be learned by studying them. Government statistics often come from a census or samples so large that the sampling variability is negligible. A good job of describing such data is often not a trivial exercise.

The key concepts in this chapter concern ways of looking at and describing one set of data. We start with graphical displays and then move to numerical summaries. The point of the numerical summary is that it quantifies and expresses compactly something that can be seen roughly in a graphical display. The effectiveness of a numerical summary is based on how well it calls to mind what appears in the graphical display. Thus viewed, the mean is not a very good numerical summary for a set of data with two distinct clusters of observations.

From numerical summaries we proceed to models for data, illustrated by the important case of the Normal distributions. The progression from graphical display to numerical measures to a compact mathematical model is one of the strategies that unifies data analysis and makes it more than a collection of clever methods.

1.1 Displaying Distributions with Graphs

The introductory material on measurement is short but important. What is measured and how it is measured are two fundamental issues that lie at the foundation of any statistical

analysis. Many of the judgments made at this level are subjective and require common sense. Make the point that statistics is about data, not just numbers.

Variation is another key concept. Why are all of the observations in a set of data not the same? Good class discussions can be based on sources of variation in particular sets of data.

The first real statistical analysis presented is the stemplot. Stress the idea that we start with a graphical display before proceeding to numerical summaries. We have presented stemplots with the values increasing as you proceed down the stems. Some computer packages put the larger numbers at the top. Similarly, in the text truncation is recommended (this is clearly the easiest method to use if constructing the plot by hand), whereas some programs will round. Keep in mind that the purpose of the plot is to get a quick look at the data and to help in deciding what sorts of numerical summaries are appropriate. Be flexible. Splitting leaves and other sorts of judgments should be based on common sense rather than on adherence to a set of inflexible rules.

Similar comments regarding flexibility apply to histograms. The histogram is presented as a method for displaying a set of data when a stemplot is unsuitable. It also serves an important role pedagogically. In the last section of this chapter, the density curve is introduced as a model that describes the shape of a histogram when the sample size increases and the widths of the intervals decrease.

The major theme of this section is that we should always carefully look at data. In graphical displays we look for an overall pattern and deviations from that pattern. Simon Newcomb was a famous scientist who was measuring well-defined physical quantities. If he had outliers in his data, we should not be surprised to find them in other sets of data that we may encounter. This section introduces the basic strategy for looking at data: Seek first an overall pattern, then deviations from the pattern. This strategy will be used in other settings in Chapters 2 and 3.

1.2 Describing Distributions with Numbers

The mean should be familiar to almost all students. Therefore, the introduction of the summation notation is facilitated by noting that it is a compact way to express something that they already know. Although the idea behind the median is quite intuitive, some students get confused by the manipulations required to find it. Emphasize that it is a numerical measure for something that can be seen in a stemplot or histogram.

Quartiles and the five-number summary build on the idea of the median. Again emphasize that they are numerical summaries for something that can be seen graphically. With the boxplot, we complete a circle of sorts. Starting with a graphical summary, we calculate numerical summaries and then present these graphically. Note that there are many variations on boxplots. If you use statistical software, the rules for constructing them may differ slightly from those presented in the text. This should not pose any serious problems. In fact, it can be used to illustrate the point that caution is needed in using statistical

packages. We need to know how things are computed if we are to interpret the output meaningfully.

Another thing to emphasize is that no single numerical summary is appropriate for all sets of data. We should examine a stemplot, histogram, or other display to make a decision about which summary is appropriate.

The standard deviation is a fundamental quantity, the meaning of which will be made clear in Section 1.3 on Normal distributions. Here, the variance is a necessary intermediate step required for the calculation of the standard deviation. For students who have a great deal of trouble with computations, the calculations can be organized in a table with columns corresponding to x, x^2, $x - \bar{x}$, and $(x - \bar{x})^2$. We recommend that students have a calculator that computes $\bar{x}$ and s from keyed-in data. This inexpensive tool greatly reduces the arithmetical burden for many standard statistical procedures.

Linear functions are used in several places later in the text. Thus, the material on changing the unit of measurement has two purposes. One idea is that when data are transformed in this way, nothing fundamental is changed. The summary statistics in the new scale are easily computed from the original ones. We are simply choosing a different way to express the same thing. The second idea is that we can use an equation of the form $x^* = a + bx$ (or later $y = a + bx$) to express a relationship between two variables.

1.3 Density Curves and Normal Distributions

Note that Normal distributions are introduced here as a common model for the overall pattern of many sets of data, and not in the context of probability theory. Although this ordering of material is unusual, it has several advantages. The Normal distributions appear naturally in the description of large amounts of data, so that the later assumption for inference that *the population has a Normal distribution* becomes clearer. Moreover, mastering Normal calculations at this point makes it easier to teach the material on probability (Chapters 5 and 6). If the students already know how to compute Normal probabilities and have a fair understanding of the relative frequency interpretation from this section, the transition to general ideas about probability is facilitated. Of course, later chapters present additional facts about Normal distributions in the context of probability.

The key idea is that we can use a mathematical model as an approximation to real phenomena. The 68–95–99.7 rule is a useful device for interpreting μ and σ for Normal distributions.

From the viewpoint of statistics, in contrast to that of probability, we always think of our models as approximations rather than the truth. This point can be illustrated by considering Example 1.30. The $N(505, 110)$ distribution is very useful for describing SAT scores of high school seniors. It can give reasonably accurate answers to interesting questions as illustrated in this example. However, it does not work very well in the extreme tails. For example, since the highest score possible is 800, the proportion of students scoring 835 or better is not appropriately calculated by using $Z \geq (835 - 505)/110 = 3$

and Normal calculations. In a population of 100,000 students, the calculation would give $100{,}000(0.00135) = 135$ students scoring above 835!

Normal quantile plots are used frequently in the text. Students should learn to interpret them. If computer software is not available for constructing these plots, we do not recommend that they be drawn by hand. In this case, stemplots should be used to assess Normality on a routine basis for the exercises.

Note that software packages that create quantile plots may differ in what variable is put on each axis. Most of the plots in the book place the variable being studied on the vertical axis, and the z-score on the horizontal axis. One benefit of this arrangement is that we can interpret the plot in a natural way: If (for example) the upper right part of the plot bends down (below the line), we can interpret it by saying, "The high values of the variable are lower than they 'should' be [to be from a Normal distribution]." (This is natural, since the *high* part of the plot is *below* the line.) Likewise, if the lower left part of the plot bends up, we can say, "The low values of the variable are too high [to be from a Normal distribution]."

Chapter 2: Looking at Data—Relationships

This chapter concludes the part of the text dealing with methods for looking at data. In the first chapter, techniques for studying a single variable were explored. Here, the focus is on general methods for describing relationships between pairs of variables.

Scatterplots are used to examine pairs where both variables are quantitative. Details of least-squares regression are presented. Correlation is introduced and its connection with regression is examined. Relations between categorical variables are described using proportions or percentages calculated from two-way tables. Issues related to the question of causation are illustrated by consideration of smoking and lung cancer.

The descriptive methods in this chapter, like those Chapter 1, correspond to the formal inference procedures presented later in the text. By carefully describing data first, we avoid using inference procedures where they clearly do not apply. Fitting a least-squares line is a general procedure, while using such a line to give a 95% prediction interval requires additional assumptions that are not always valid. In addition, students become accustomed to examining data *before* proceeding to formal inference, an important principle of good statistical practice. The data for a regression (Chapter 10) are displayed in a scatterplot. Side-by-side boxplots are useful with two-sample t tests (Chapter 7) and analysis of variance (Chapter 12). The percentages obtained from two-way tables are the basis for the formal inferences on count data (Chapter 9).

2.1 Scatterplots

From the previous chapter the students will be familiar with the idea that graphical displays of data are useful. The extension to a scatterplot should reinforce this idea. The

distinction between explanatory and response variables is a relatively simple distinction that is essential to the least-squares method.

Constructing scatterplots is an easy task, particularly with a computer. Interpreting them, on the other hand, is an art that takes practice. For classroom discussion, you can use the examples given in the text or those presented in the exercises. Stress the idea that common sense and some understanding of the data are necessary to do a good job of description. Computers can make the plots, but people are needed to describe them. Again, the general rule is to look for overall patterns and deviations from them. Dichotomies such as positive and negative association are useful in many cases but can lead to distorted descriptions when imposed in situations where they do not apply.

Scatterplot smoothing is a useful data analytic tool for looking at some kinds of scatterplots. Try to avoid the type of thinking that says: If I have situation a, then use method b. A method is useful if it helps you to learn something about the data. Sometimes you do not know if it is useful until you try it.

Side-by-side boxplots can give a very informative data summary when one variable (usually the response variable) is quantitative and the other is categorical or concentrated on a small number of possible values. Since the students have already encountered boxplots, the idea of putting several of them alongside each other should follow quite easily.

2.2 Correlation

A key point is that a correlation (or the square of a correlation) is a numerical summary of a relationship between two variables that are linearly related. Thus, as an aid to interpreting a scatterplot, it is a potentially useful descriptor. With outliers, influential observations or nonlinear relationships, however, it may give a very distorted impression. Correlations are somewhat complicated for hand computation, and most of the time they should be found with a calculator or software. However, there may be some benefit to having students compute correlations for some small data sets by hand, to get a "feel" for how correlations work.

2.3 Least-Squares Regression

The principle of least squares is introduced and a method of computation (using r, $\bar{x}$, s_x, $\bar{y}$, and s_y) is presented. Try not to burden the students with excessive computation; most least-squares lines should be found using a computer or calculator. Note that the output may contain many pieces of information related to inference that are not relevant at this time. If computing facilities are not available, assign exercises where most of the computations are given and the amount of arithmetic required is minimal.

2.4 Cautions about Correlation and Regression

A crucial part of using regression and correlation procedures is knowing what can go wrong and when they may not be appropriate tools. Plotting the residuals versus x is very

important. Again, construction of the plot is easy and can be done effectively by a computer, but interpreting the results is an art.

Outliers and influential observations can provide interesting class discussion. When a cause can be found or a lurking variable discovered, the value of carefully looking at data is reinforced. Formal rules for dealing with these situations are not advised. The point is that outliers and influential observations are present in many real data sets. A good analysis finds them and assesses their effects on the results.

2.5 Data Analysis for Two-Way Tables

The computation required in this section in minimal. Percents and proportions are the numerical summaries. On the other hand, some very important ideas are presented. Judgment is required to select what percents to calculate.

Simpson's paradox is not easily understood when first encountered. Careful class discussion of the examples in the text is needed if students are to grasp this idea.

These ideas are explored further in Chapter 9.

2.6 The Question of Causation

The final section of this chapter contains no formulas, but there are a lot of important ideas. It brings out the fact that drawing conclusions about relationships may require judgment and expertise beyond the realm of formal statistical inference.

Chapter 3: Producing Data

This is a relatively short chapter with a lot of ideas and little numerical work. The message is that production of good data requires careful planning. Random digits (from Table B or elsewhere) are used to assign units to treatments and to select simple random samples. There are several good examples that can serve as the basis for classroom discussion.

3.1 Design of Experiments

Terminology used in experimental design is introduced and is illustrated with examples. The advantages of comparative experiments are shown pointedly with the gastric freezing experiment. Students seem to particularly enjoy this example and they often express outrage at the blatant misuse of the data described.

Once the need for comparison and randomization is established, methods for randomization of experimental units to treatments is presented. Note that there are many correct ways to do these randomizations. Many exercises using Table B are given. If you use statistical software, some particular randomization routines may be available. In very general terms, most randomizations can be accomplished by a random sort of the data. In any package that has a random number generator and a sort, this can be done by assigning a random uniform to each case and then sorting the resulting file. For example, to assign

10 people to each of two treatments, set up a file with cases corresponding to the integers 1 to 20. Then assign a random number to each case and sort the file on the basis of the random numbers. The first 10 cases receive treatment one and the others receive treatment two. Note that this is not a particularly efficient algorithm, but it is very clear and it helps the students understand the principle of randomization.

3.2 Sampling Design

The simple random sample (SRS) is described. Note that in later chapters the assumption that data or errors are Normally distributed is expressed by saying that they are an SRS from a Normal population.

Other sampling designs are mentioned, including stratified samples, multistage designs, and systematic random samples. The problems of nonresponse and bias are also discussed.

3.3 Toward Statistical Inference

This section introduces some key ideas that will be further explained and used in later chapters. The idea of a parameter in a population contrasted with a statistic in a sample is fundamental. Through an opinion poll example, this distinction is made clear. The sampling distribution of the sample proportion is shown by simulation for samples of size 100 and 2500, thus illustrating the central limit theorem and the fact that the variability decreases as the sample size increases.

Because students often find probability hard and fail to see its relevance, we have tried to motivate probability by first discussing randomization in data collection and its consequences. In teaching this section, try to get the students to think about and discuss the general ideas. This will facilitate their learning of the details to follow in later chapters.

3.4 Ethics

After relegation to "supplemental status" in the previous edition, this material has been restored to the main text. In addition to introducing the fundamental definitions and terminology of data ethics, this section some of the tough "real-life" questions that arise from the sometimes-conflicting goals of extending human knowledge and respecting rights. Many of the exercises here are good starters for class discussion (or reflective papers).

Chapter 4: Probability: The Study of Randomness

In Chapter 3, it is explained that randomization is needed for the proper design of experiments and sampling plans. Therefore, the study of probability as a model for randomness naturally fits here.

An overview of the entire course is given in the following paradigm. In the first three chapters, we studied how to look at and describe data. Thus, in a sense, we concentrated

on the *sample*. Chapter 3 discussed how data are produced and introduced the idea of the *population* for which inference in desired. Now, we study the population in detail, with emphasis on the mechanisms by which samples are generated. In later chapters, we reverse the orientation and study the methods by which samples lead to inferences about populations.

This chapter and the one that follows will be difficult for many of the students. We have tried to emphasize a conceptual approach to the topics presented. Nonetheless, there are several rules and manipulations that must be mastered.

You should carefully consider how much depth and breadth you want to include for this material. By omitting the optional sections and treating some subjects (such as sample spaces) lightly, you can teach the minimum probability required for a proper understanding of statistical inference. On the other hand, the text provides sufficient material for more comprehensive coverage.

4.1 Randomness

Certainly everyone has some concept of what is meant by the word "random"—but many may have the wrong idea, especially in not understanding the underlying order of "random" events. (Many people might list "chaotic" or "unpredictable" as synonyms for "random.") This section gives some background to understanding randomness and probability in terms of "long-run frequency" by repeated actual or simulated experimentation.

4.2 Probability Models

Sample spaces, assignment of probabilities, and basic probability rules are introduced in this section. Try to emphasize that all of these ideas follow from common sense and are not an arbitrary mathematical system devoid of a intuitive foundation. Many of the students will have had unpleasant experiences with some of this material previously.

Many of the topics can be introduced through an example with the formalization coming later. Give a set of probabilities that have a sum greater than one and ask why this does not make sense. Then explain that their conclusion that the sum must be one is a basic rule of probability theory. Similarly, most of the students can apply the addition rule and the complement rule to real problems without ever having encountered these rules formally. Let them figure out the rule first and then present it formally.

4.3 Random Variables

Random variables are the raw material for statistical inference. We therefore introduce them early in the presentation of probability. If an experiment or sample survey is repeated, different results will be obtained. The probability theory for random variables tells us that there is a certain type of regularity or predictability in these results.

4.4 Means and Variances of Random Variables

The mean, variance, and standard deviation are familiar quantities from Chapter 1. Emphasize that they are a little different here because we are dealing with a population. If the students have a lot of difficulty with this transition, you can ask them to think about some very large samples. For example, in Example 4.21, think of an experiment with 1000 trials in which the results are exactly as expected by Benford's Law: 301 ones, 176 twos, 125 threes, and so on. The calculations can then be performed on this hypothetical large sample (in which the difference between n and $n - 1$ is negligible) and the connection between probability and relative frequency can be reinforced.

In a similar way the rules for means and variances build on the parallel rules for sample statistics. This material can be treated briefly and reinforced with homework exercises.

The law of large numbers is the key concept in this section. It expresses a fundamental idea that is part of the foundation for much of statistical inference. Most of the students will have some intuitive idea about this law. Through class discussion, these ideas can be explored and clarified.

4.5 General Probability Rules

This optional section presents some of the traditional material on probability. The addition rules, multiplication rules, and conditional probability are all discussed. The section is optional because these ideas are not essential to an understanding of the statistical ideas that follow in later chapters. Ample exercises are provided for the instructor who wants to cover this material.

Although the basic ideas contained in these probability laws are very intuitive, many students have a great deal of difficulty with the symbolic manipulations involved. By a detailed presentation of examples or exercises, you can emphasize that these calculations are based on common sense and are not simply a collection of abstract rules.

Chapter 5: From Probability to Inference

There are two key ideas in this chapter. First, a statistic is a random variable, the value of which varies from experiment to experiment or sample to sample. Second, as a random variable, it has a sampling distribution with a mean and a standard deviation that can be calculated from the distribution of the basic random variables that are combined to calculate the statistic.

The binomial distribution is treated in detail. The Normal approximation to the binomial is given as an important special case that sets the stage for the central limit theorem presented in the second section. The chapter concludes with an optional section on control charts where probability ideas are shown to be both useful and important.

5.1 Sampling Distributions for Counts and Proportions

The amount of time you choose to spend on this section is quite flexible. The inference for proportions presented in Chapter 8 is all based on the large sample Normal approximation to the binomial. Therefore, the minimum coverage needed consists of the fact that (when n is sufficiently large) the distribution of $\hat{p}$ is approximately Normal with mean p and standard deviation $\sqrt{p(1-p)/n}$.

On the other hand, there is sufficient material for a full treatment of the binomial distribution. Table C gives individual probabilities for a variety of values of n and p and the formula for calculating binomial probabilities is explained in detail. Many computer routines are also available for calculating individual or cumulative binomial probabilities.

Try not to let the computational details interfere with the important pedagogical role played by this section. The students should understand that $\hat{p}$ is a random variable with a sampling distribution that is approximately Normal for large n. In a given experiment or sample, only *one* realization or value of this statistic is obtained. Our inference will be based on this single observation, and probability calculations will be based on an assumed (testing) or estimated (confidence intervals) sampling distribution for this statistic.

5.2 The Sampling Distribution of a Sample Mean

This is a rather short but important section where the mean and standard deviation for the sample mean are explained and the central limit theorem is presented. In terms of statistical applications, there are two key ideas. First, the central limit theorem is widely applicable and is the basis for Normal-based inference. Second, the standard deviation of the mean decreases as $\sqrt{n}$. This section completes the discussion of Normal distributions begun in Chapter 1 by showing the Normal distributions as sampling distributions.

Chapter 6: Introduction to Inference

This is a chapter with a lot of fundamental ideas. Confidence intervals and tests are presented with some cautions concerning the use and abuse of tests. Throughout, the setting is inference about the mean μ of a Normal population with known standard deviation. As a consequence, the z procedures presented are not applicable to most real sets of data, but rather serve to ease the transition to the more useful procedures presented in Chapter 7.

Experience has shown that many students will not master all this material upon seeing it for the first time. Fortunately, they will see all the key ideas again in Chapter 7. By the time they have completed both chapters and worked many exercises, they should be able to grasp the fundamentals. Remember that mastery of the reasoning of inference is far more important than the number of procedures learned.

6.1 Estimating with Confidence

Confidence intervals are straightforward but easily misinterpreted. Try to emphasize the meaning rather than the details of computation. For most students, this is accomplished more effectively by a discussion of Figure 6.3 or similar results than by an algebraic deviation of a formula for the coverage probability.

At this stage, students should be comfortable with the sample mean, population parameters versus sample statistics, the fact that $\bar{x}$ is approximately Normal with mean μ and standard deviation $\sigma/\sqrt{n}$, and the table of Normal probabilities (Table A). Stress the fact that we are simply putting these ideas together in this section to produce a useful statistical procedure. Similar comments hold for Section 6.2.

The subsection on choosing the sample size reinforces the idea that the width of the interval depends upon n. This material is particularly important for students who will design studies in the future.

6.2 Tests of Significance

Although this section is not much longer than the previous one, it contains a few things that students generally find more difficult. The choice of one-sided versus two-sided tests and P-values versus fixed significance level testing are in this category. Try not to let these complications interfere with the presentation of the basic ideas.

Students seem to have the most difficulty when they become sufficiently frustrated that they put their common sense aside and start seeking formulas or rules that cover every possible case. Graphs illustrating the calculation of P-values are very useful here. Note that statistical software universally gives P-values. Therefore, an understanding of this approach to testing is essential for students who will either read or perform statistical studies.

6.3 Use and Abuse of Tests

This section gives some important ideas about significance testing that can serve as the basis for classroom discussion. Students are often reluctant to make judgments about the *importance* of an effect in contrast to its statistical significance. Emphasize that this judgment is an important part of using statistics with real data. Computers are easily programmed to perform calculations and produce P-values. Informed judgments (by humans) are needed to translate and interpret these results.

6.4 Power and Inference as a Decision

The optional material on power is similar to the material on selecting the sample size for confidence intervals. It reinforces what has already been said about the dependence of the behavior of the procedure on the sample size and is very important for students who will be planning studies. The use of software—for example, G•Power (available on the Internet for many platforms) or online power calculators—can make this more approachable.

The last subsection on inference as a decision is also optional. It uses the traditional dichotomy between Type I and Type II errors and serves as a brief introduction to more general applications of decision theory.

Chapter 7: Inference for Distributions

The principles underlying confidence intervals and significance tests were presented in Chapter 6. We now apply these principles to inference problems for one and two samples. Throughout the chapter, we work under the realistic assumption that the population standard deviation is unknown.

The first section deals with inference for one sample. The problem of paired comparisons is treated as a special case. Optional sections on power and a nonparametric alternative to the Normal-based procedure are also given.

Two sample procedures are presented in the second section. The chapter concludes with an optional section on inference for the standard deviation.

From earlier chapters, the students should be familiar with looking at data from one or two samples carefully and with computing sample means and standard deviations. The additional arithmetic required for the construction of confidence intervals and significance tests is not particularly difficult. Many of the exercises provide means and standard deviations so that the students can concentrate on the new ideas presented here.

The transition from the Normal to the t table requires some careful explanation. Because we are dealing with a family of distributions indexed by the degrees of freedom, only selected values can be given in a table. Since most of the exercises are based on real data, no attempt has been made to assure that an entry for the exact degrees of freedom for a particular problem are given in the table. Therefore, approximate values of t^* are to be used for confidence intervals. Similarly, bounds on the P-value can be obtained rather than exact P-values, using Table E. (Nonetheless, since software may be used to give more "exact" answers, the solutions in the latter part of this guide contain these more exact answers, as well as those that come easily from Table E.)

A conservative approach for confidence intervals is to use a value for the degrees of freedom that is less than or equal to the required value. This results in intervals that are at least as wide as the exact interval, thereby ensuring that the coverage probability is at least as large as that specified. (In effect, by taking smaller degrees of freedom than we are "entitled" to, we are pretending that we have less information [a smaller sample] than was actually available to us.) The same approach for tests ensures that the probability of a false rejection of H_0 is less than or equal to the value given. Table E provides sufficient detail so that for all practical purposes, inaccuracies resulting from these difficulties are very small. Students checking their solutions for odd-numbered exercises in the back of the text should be aware of these considerations.

Of course, if a probability function for the t distribution is available in computer software, exact values are easily obtained. Note that it is common practice for many computer packages (and in reporting the results of a statistical analysis) to give P-values in the form $P <$ some number. Using Table E, when an extreme value of the t statistic is found, the P-value will be reported as $P < 0.0005$ for a one sided-test and $P < 0.001$ for a two-sided test.

7.1 Inference for the Mean of a Population

The term *standard error* is introduced for the estimated standard deviation of a statistic. Thus, t statistics are standardized by dividing by a standard error, and confidence intervals are constructed by taking the value of a statistic plus or minus a constant times the standard error.

The usual one-sample confidence intervals and significance tests (which assume Normal distributions) are presented in this section. By first taking differences between pairs of observations, paired comparisons problems are treated in the same way. The idea of robustness is introduced, and practical guidelines for using the t procedures are given.

7.2 Comparing Two Means

Procedures based on Normal assumptions for comparing two means are presented in this chapter. To place the problems in a proper perspective, the situation where the standard deviations are known is discussed first.

When replacing the known standard deviations by sample estimates, difficulties regarding the appropriate degrees of freedom for the t distribution arise. Most computer software packages use an approximation similar to that given just before Example 7.18 (page 460). This formula is given for information only. From an educational point of view, very little is accomplished by having students do a large amount of computation with this formula. In most practical situations, use of the minimum of $n_1 - 1$ and $n_2 - 1$ as the degrees of freedom for the t gives essentially the same results. Rather than having students spend a lot of time with computation, it is better to have them concentrate on the appropriate use of the procedures and the interpretation of the results.

The pooled two-sample t procedures are presented in the optional last part of this section. Occasionally, students will question the usefulness of these procedures. Why assume that the standard deviations are equal when we have already learned procedures that are valid under less restrictive assumptions? From a theoretical point of view, there are answers. We have exact rather than approximate distributions, confidence intervals will be shorter in some stochastic sense, and significance tests have more power. On the other hand, from a practical point of view, there will be very little difference in the results (unless, of course, there are large differences in the standard deviations so that the pooled procedures are invalid). The primary reason for including this material is that it facilitates the introduction of similar ideas in Chapter 12, where analysis of variance is introduced. If you do not plan to cover analysis of variance and time is tight, you could easily omit this material.

7.3 Optional Topics in Comparing Distributions

The usual F test for comparing two variances is presented in this section. It is pointed out that this procedure is not robust with respect to the assumption of Normality and a general discussion of robustness follows. This section can easily be omitted. The very poor performance of F in the two-sample variance setting should not be overlooked merely to teach a procedure for this case.

Note that this is the first place in the text that F distributions are used. They appear later in the chapters on regression and analysis of variance (Chapters 10–13). If you do not cover this section, you will need to spend a little more time explaining the tables when they are needed later in the text.

The power of the t test is treated in a subsection. This material is very similar to that presented in the fourth section of Chapter 6 and makes use of the noncentral t distributions. As was mentioned before, this material reinforces the fundamental idea that the performance characteristics of statistical procedures depend on the sample size. It is important for students who will be planning statistical studies to be familiar with these ideas.

The final optional subsection discusses what can be done in some situations when the data are not Normal. Transformations are mentioned and an explanation of the sign test for paired comparisons is given. The intention here is to point to the two major strategies available, not to discuss them in detail.

Chapter 8: Inference for Proportions

The two sections of this chapter present the standard z procedures for one-sample and two-sample binomial problems. By now, students should be comfortable with the general framework for confidence intervals and significance tests. For those who have not yet completely mastered these concepts, this material affords an additional opportunity to learn these important ideas.

8.1 Inference for a Single Proportion

Confidence intervals and significance tests for a single proportion are presented. A new complication that arises with these problems concerns the standard deviation of $\hat{p}$. For confidence intervals we use the standard error $\sqrt{\hat{p}(1-\hat{p})/n}$, whereas for tests we use $\sqrt{p_0(1-p_0)/n}$. Although this may seem confusing to the students at first, the basic idea is quite reasonable.

The construction of confidence intervals using plus four (Wilson) estimates is introduced here, but it is an optional topic that can be omitted without significant repercussions. (This is a change from the more prominent treatment of this topic in the fourth edition.)

We use all of the information available in a problem for our calculations. For the confidence interval, p is assumed to be unknown and the standard error must therefore be

estimated using the value of $\hat{p}$ obtained from the data. On the other hand, when testing H_0: $p = p_0$, our calculations are based on the assumption that H_0 is true, and therefore we use the value p_0 in the calculations. Note that these choices destroy the exact correspondence between confidence intervals and tests (reject if the hypothesized parameter value is outside of the confidence interval).

The section concludes with a description of the procedure for choosing a sample size to guarantee a given bound on the width of a confidence interval. From a practical point of view, the table given in Example 8.8 is very informative. The widths of intervals vary relatively little for values of $\hat{p}$ between 0.3 and 0.7.

8.2 Comparing Two Proportions

Confidence intervals and significance tests for comparing two proportions are presented in this section. From Chapter 7, students should be familiar with the basics of two-sample problems. As in Section 8.1, we use different standard errors for confidence intervals and tests. The idea of pooling information from two samples to estimate a standard error was first presented in the optional part of the second section of Chapter 7. This principle will be used again in Chapter 11 when we treat analysis of variance.

Once again, plus four confidence intervals are discussed but are not vital to the material.

Chapter 9: Analysis of Two-Way Tables

If desired, this material can be omitted without loss of continuity. The notation here is more complex than in earlier chapters. (There is no easy way to describe the general form of a two-dimensional array of data without two subscripts.)

9.1 Inference for Two-Way Tables

You have some choice regarding depth of coverage for this section. As mentioned, this section can be eliminated entirely if desired. The computations for the X^2 statistic are rather straightforward, although a fair amount of arithmetic is required and care must be taken to ensure accurate results. Computers can be a big help here.

9.2 Formulas and Models for Two-Way Tables

The vague idea of testing for a relationship between two categorical variables is also fairly straightforward. Therefore, it is possible to teach this section spending a minimum amount of time on the two models described.

On the other hand, discussion of the models in detail gives the students a good background for learning about model-based inference for more complex situations. The statistical model gives a clear statement of the assumptions needed for a given procedure and specifies the parameters about which inference is desired. A major part of statistical inference concerns the translation of a vague notion, such as dependence between two

categorical variables, into a testable hypothesis stated in terms of the parameters of a statistical model. To draw meaningful conclusions, the results of the analysis, stated in terms of the model, must then be translated back into the context of the real problem.

9.3 Goodness of Fit

This optional section presents an idea that is a fairly natural extension of the material on comparing proportions from Chapter 8: How can we decide if a set of observed counts (for a categorical variable) is consistent with a particular set of probabilities for those categories? For example, if we roll a fair die 600 times, how much deviation from an "evenly spread" distribution (100 of each number) might we expect to see? Do the colors in a package of M&M's match the stated distribution from `www.mms.com` (Exercise 9.5)?

If you wish to present this material, you might consider introducing it *before* Secions 9.1 and 9.2. Because it concerns analysis of a "one-way table" rather than a two-way table, the necessary chi-square computations can serve as a gentle introduction to the more involved computations of the earlier sections.

Chapter 10: Inference for Regression

Having completed Chapter 2, students should be familiar with the data analytic issues that arise with simple linear regression. We now focus on applying the statistical inference principles of Chapter 7 to this problem. Note that nothing from Chapters 8 and 9 is needed for an understanding of this chapter.

10.1 Simple Linear Regression

There are many interesting problems for which the relationship between two variables can be summarized graphically and numerically with a least-squares line. Not all of these can be analyzed using the methods of this chapter. Inference for linear regression is based on a statistical model that expresses the assumptions underlying the inference procedures.

10.2 More Detail About Simple Linear Regression

Confidence intervals and significance tests should be familiar concepts for the students at this point. Although the computational details are more difficult in this section, the underlying principles are not new. Use of statistical software is very effective here. It allows the students to concentrate on learning the concepts without being unduly concerned with a long sequence of computations leading to a numerically correct answer. Stress that the number is not the *answer;* the interpretation of the number in the context of a real problem is the proper end result of a statistical analysis.

Prediction intervals are a new idea. By contrasting them with confidence intervals, the students can avoid common misinterpretations of confidence intervals.

Most computer packages provide an analysis of variance table as part of the output for a regression. For simple linear regression, the table is not particularly useful. Often, it generates more confusion than useful information for students at this point. On the other hand, the ANOVA table is very important for multiple regression, and it is essential for performing analysis of variance. If you plan to cover these topics, discussion of the ANOVA table for linear regression provides an excellent introduction to this topic. Explain that it is a different way to present some of the information already discussed and that it will be important for topics to be covered later.

The section concludes with optional material on inference for correlation. We see that the test for zero correlation is equivalent to the test for zero slope in the regression. The square of the correlation is expressed in a form that generalizes easily to multiple regression.

Chapter 11: Multiple Regression

The model for multiple regression and a brief overview of the inference procedures are presented. This chapter can be omitted without loss of continuity. For students who will take another statistics course in which this material will be covered in great detail, a discussion of the case study can be very valuable. Many important ideas are presented in the context of the problem of predicting grade point average from high school grades and SAT scores.

Chapter 12: One-Way Analysis of Variance

The idea of using a statistical test to compare population means is familiar from Chapter 7. New issues arise because we may have more than two means to compare. The DATA = FIT + RESIDUAL idea is used in presenting the statistical model underlying the analysis. The model is treated first, thereby emphasizing the principle that inference is based on assumptions and definitions of parameters given by a model.

The analysis of variance table is constructed using some familiar ideas. Pooling of variances was treated in Chapter 8 and again the idea of DATA = FIT + RESIDUAL is used. The ANOVA tables from the regression chapters (Chapter 10 and 11) anticipated many of the ideas presented here.

Some new ideas arise with contrasts and multiple comparisons. Without some discussion of at least one of these topics, one is left with a statistical procedure giving very little in terms of useful interpretable results. Most statistical software packages will perform these computations.

The issue of what multiple comparison procedure to use is rather difficult. We have chosen to present only one procedure—the Bonferroni. However, it is presented in such a way that any other procedure could easily be substituted.

The optional section on power reminds us again that the performance of a statistical procedure depends on the sample sizes and the true values of the parameters. It is

interesting to note that these kinds of calculations are rather easy to perform with any statistical software that has a function for the noncentral F distribution. Tables of power and even graphs of power functions are relatively easy to generate. The use of G•Power (mentioned earlier) also eases these computations.

Note that, in the section on multiple comparisons, the term "minimum significant difference (MSD)" is no longer used.

Chapter 13: Two-Way Analysis of Variance

Two-way analysis of variance is treated conceptually. A major goal is for the students to understand the meaning of interaction. No computational formulas are given; computation is left for the computer. The emphasis is on understanding the assumptions and interpreting the output.

This chapter is really an introduction to more complex analysis of variance designs. In this context, the initial discussion of advantages of two-way ANOVA is particularly important. Again, the description of the model plays a key role.

Interaction is the important new idea. It is illustrated with a variety of examples chosen to demonstrate the different types of interpretations that arise in practice. Stress that the statement that there is interaction between two factors is essentially meaningless in itself. Interpretation requires a careful examination of the means in the context of a given problem.

As in Chapter 12, note that the term "minimum significant difference (MSD)" is no longer used for multiple comparisons procedures.

Chapter 14: Logistic Regression

This chapter might provide useful supplementary material for students who have covered regression, but be aware that even fairly sophisticated statistical software packages may not have the capability to do *multiple* logistic regression. With the right software, and sufficiently prepared students, this short chapter (only 43 exercises) can be covered in a week or two.

Chapter 15: Nonparametric Tests

This chapter presents techniques for use with non-Normal data. It can be used as supplementary material for interested and capable students, building on the sign test introduced in Chapter 7. (In fact, many of the ideas in this chapter are no more difficult than those of earlier chapters; they are arguably much easier than the regression and analysis-of-variance material, even though the Kruskal-Wallis test is an alternative to ANOVA!)

There are only 48 exercises in Chapter 15, but note that one can use these methods on exercises on comparing means (Chapter 7) and analysis of variance (Chapter 12), although in many cases, there may be no reason to question underlying Normality.

Chapter 16: Bootstrap Methods and Permutation Tests

This chapter uses computationally demanding resampling procedures, for which the use of a computer is critical. The solutions (prepared by Tim Hesterberg) were written using S-PLUS. More information about the S-PLUS procedures are given in the solutions to Chapter 16 exercises (page 404 of this guide). Note that a free student version of S-PLUS is available, so your students may use it for this chapter, even if they normally work with other software.

Chapter 17: Statistics for Quality: Control and Capability

This chapter may be particularly useful for business students, as it illustrates applications of statistical ideas to a number of common situations in manufacturing and management. For a light introduction to the topic, the first two sections could stand on their own.

3 SAMPLE EXAMINATIONS

In writing exams, here are several points to keep in mind:

- Use current events and local references.
- Decide how much computation to require, and let the students know in advance. Having students compute standard deviations or correlations on exams may not be the best way to measure their comprehension.
- As an alternative to computation, consider placing computer or calculator output on the test and asking students to interpret it.
- Decide what mix of multiple-choice, problem-solving, and discussion questions to use. If class size makes it necessary, use some multiple-choice questions, but by and large, the subject matter of this course is best tested with questions that require writing ("Give reasons to support your answer."). No multiple-choice questions are given in these samples.

The exams given here can be used as examination items or as class examples. As mentioned earlier, a sample exam distributed in class is useful as review before a scheduled exam.

These questions are taken from the Instructor's Guide to the third edition. Of course, the Test Bank CD-ROM (available separately) and the online quizzes accessible from the text Web site (`www.whfreeman.com/ips6e`) are also excellent resources.

Sample examination I

These questions cover Chapters 1 and 2.

1. A study examined how long aircraft air conditioning units operated after being repaired. Here are the operating times (in hours) for one unit:

97	51	11	4	141	18	142	68	77
80	1	16	106	206	82	54	31	216
46	111	39	63	18	191	18	163	24

   (a) Make a histogram, using classes that are 40 hours wide, beginning with:

$$0 \leq \text{time} < 40$$

$$40 \leq \text{time} < 80$$

   (b) Describe the overall shape of the distribution: Is it roughly symmetric, skewed to the right, or skewed to the left? Are there any outliers?

(c) Is the five-number summary or the mean and standard deviation a better brief summary for this distribution? Explain your choice. Calculate the one of these summaries that you choose.

2. Biologists and ecologists record the distributions of measurements made on animal species to help study the distribution and evolution of the animals. The African finch *Pyrenestes ostrinus* is interesting because the distribution of its bill size has two peaks even though other body measurements follow Normal distributions. For example, a study in Cameroon found that the wing length of male finches varies according to a Normal distribution with mean 61.2 mm and standard deviation 1.8 mm.

   (a) What proportion of male finches have wings longer than 65 mm?

   (b) What is the wing length that only 2% of male finches exceed?

3. The drug AZT was the first effective treatment for AIDS. An important medical experiment demonstrated that regular doses of AZT delay the onset of symptoms in people in whom the AIDS virus is present. The researchers who carried out this experiment wanted to know:

   - Does taking either 500 mg of AZT or 1500 mg of AZT per day delay the development of AIDS?
   - Is there any difference between the effects of these two doses?

   The subjects were 1200 volunteers already infected with the AIDS virus but with no symptoms of AIDS when the study started.

   (a) Outline the design of the experiment.

   (b) Describe briefly how you would use a table of random digits to do the randomization required by your design. Then, use Table B beginning at line 110 to choose *the first five* subjects for one of your groups.

4. A long-term study of changing environmental conditions in Chesapeake Bay found the following annual average salinity readings in one location in the bay:

Year	1971	1972	1973	1974	1975	1976	1977
Salinity (%)	13.2	9.3	14.9	13.9	14.8	13.3	15.0

Year	1978	1979	1980	1981	1982	1983	1984
Salinity (%)	15.3	15.1	13.1	17.0	19.3	15.6	15.3

   (a) Make a plot of salinity against time. Was salinity generally increasing or decreasing over these years? Is there an overall straight-line trend over time?

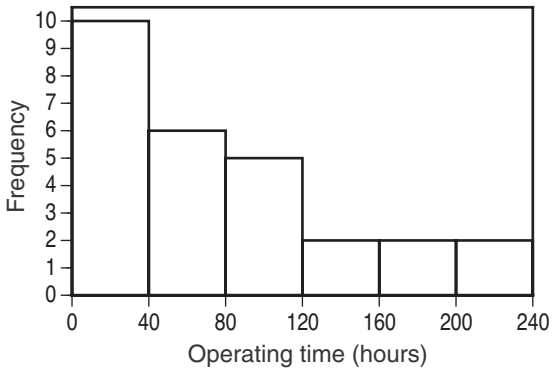
   (b) What is the correlation between salinity and year? What percent of the observed variation in salinity is accounted for by straight line change over time?

(c) Find the least-squares regression line for predicting salinity from year. Explain in simple language what the slope of this line tells you about Chesapeake Bay.

(d) If the trend in these past data had continued, what would be the average salinity at this point in the bay in 1988?

Sample examination I solutions

1. (a) At right.



(b) The distribution of operating times is strongly skewed to the right. There are no outliers.

(c) The five-number summary is preferable for this strongly skewed distribution. First, arrange the observations in increasing order:

1	4	11	16	18	18	18	24	31	39	46	51	54	63
68	77	80	82	97	106	111	141	142	163	191	206	216	

The five-number summary of these $n = 27$ observations is:

$$1 \quad 18 \quad 63 \quad 111 \quad 216$$

2. (a) Wing length x has the $N(61.2, 1.8)$ distribution. So we want the area under a Normal curve such that:

$$x > 65$$

$$\frac{x - 61.2}{1.8} > \frac{65 - 61.2}{1.8}$$

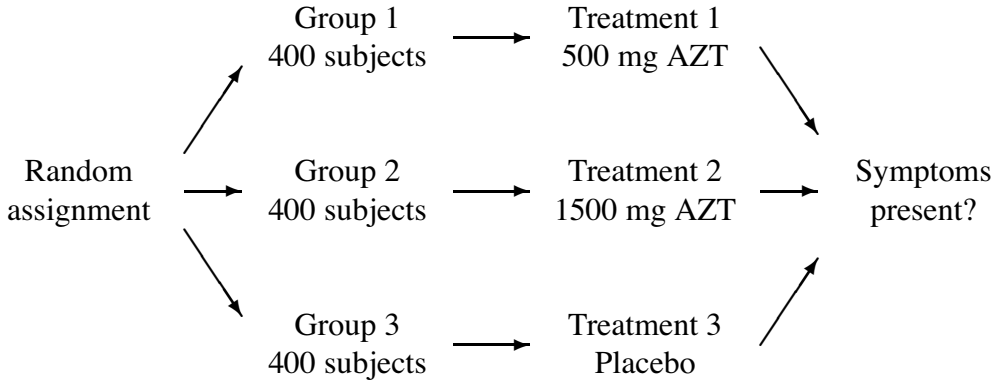
$$z > 2.11$$

Table A gives this area as $1 - 0.9826 = 0.0174$. About 17.4% of male finches have wing lengths exceeding 64.9 mm.

(b) We want the x with area 0.02 to its right, or area 0.98 to its left. In the body of Table A, find $z = 2.06$ as the entry with left tail area closest to 0.98. So:

$$x = 61.2 + (1.8)(2.06) = 64.9 \text{ mm}$$

3. (a) The goals of the experiment require *three* treatment groups, one of which receives a placebo. (Because AZT was the first AIDS drug, it was considered ethical to give a placebo to test its effectiveness. Later drugs were tested against AZT.) Here is the design:



   (b) First, assign labels. We use labels 0001 to 1200. Then, read four-digit groups from line 110, continuing to the following lines. The first five subjects chosen are:

   0676 0041 0404 1197 0640

4. (a) There is an increasing linear trend over time.

   (b) The correlation (use a calculator) is $r = 0.6386$. Because $r^2 = 0.4079$, linear change over time explains about 41% of the observed variation in salinity over this period.

   (c) The least-squares line (use a calculator) is:

   $$\hat{y} = -659.4385 + 0.340879x$$

   That is, salinity is increasing by 0.34% per year on the average.

   (d) The prediction for $x = 1988$ is:

   $$\hat{y} = -659.4385 + (0.3409)(1988) = 18.23\%$$

Sample examination II

These questions cover Chapters 3 through 5.

1. About 22% of the residents of California were born outside the United States. You choose an SRS of 1000 California residents for a sample survey on immigration issues. You want to find the probability that 250 or more of the people in your sample were born outside the United States.

   (a) You would like to use the Normal approximation for the sampling distribution of a sample proportion to answer this question. Explain carefully why you can use the approximation in this setting.

(b) What is the probability that 250 or more of the people in your sample were born outside the United States?

2. The weights of newborn children in the United States vary according to the Normal distribution with mean 7.5 pounds and standard deviation 1.25 pounds. The government classifies a newborn as having low birth weight if the weight is less than 5.5 pounds.

   (a) What is the probability that a baby chosen at random weighs less than 5.5 pounds at birth?

   (b) You choose three babies at random. What is the probability that their average birthweight is less than 5.5 pounds?

3. Answer each of the following short questions.

   (a) Give the upper 0.025 critical value for the standard Normal distribution.

   (b) An animal scientist is studying factors that affect the level of milk production in dairy cows. He wonders:

   > Is the mean production different for cows who are given forage spread on the ground than for cows whose forage is in bunks?

   State the null and alternative hypotheses that you would use in a statistical test of this question. (We can't test these hypotheses yet.)

   (c) An opinion poll asks 1500 randomly chosen U.S. residents their opinion about relations with Russia. The announced margin of error for 95% confidence is ± 3 points. But some people were not on the list from which respondents were chosen, some could not be contacted, and some refused to answer. Does the announced margin of error include errors from these causes?

   (d) A student organization plans to ask 100 randomly selected students whether they favor a change in the grading system. You argue for a sample of 900 students instead of 100. You know that the standard deviation of the proportion $\hat{p}$ of the sample who say "Yes" will be __ times as large with the larger sample. Should this blank be filled in with nine, one-ninth, three, or one-third? Explain your answer.

   (e) You read in a journal a report of a study that found a statistically significant result at the 5% significance level. What can you say about the significance of this result at the 1% level: Is it certainly significant at the 1% level, certainly not significant at the 1% level, or maybe significant and maybe not significant?

4. A friend who hears that you are taking a statistics course asks for help with a chemistry lab report. She has made four independent measurements of the specific gravity of a compound. The results are:

   3.82, 3.93, 3.67, 3.78

   The lab manual says that repeated measurements will vary according to a Normal distribution with standard deviation $\sigma = 0.15$. (This standard deviation shows how precise the measurement process is.) The mean μ of the distribution of measurements is the true specific gravity.

   (a) The lab manual asks for a 95% confidence interval for the true specific gravity. Your friend doesn't know how to do this. Do it for her.

   (b) Now explain to your friend in simple language what "95% confidence" means.

   (c) What critical value from the Normal table would you use if you wanted 80% confidence instead of 95% confidence? Would the 80% confidence interval be wider or narrower than your 95% confidence from (a)? [Do *not* actually compute the 80% confidence interval.]

   (d) The lab manual also asks whether the data show convincingly that the true specific gravity is less than 3.9. State the null and alternative hypotheses used to answer this question. Then, calculate the test statistic, and find its P-value.

   (e) Explain to your friend in one or two sentences what the specific P-value you found in (c) means.

5. Here are several statistical statements. In each case, explain to someone who knows no statistics what the boldface term means. Use just a sentence or two and avoid technical jargon.

   (a) A spinning penny has **probability 0.6** of coming up tails.

   (b) With **90% confidence,** we can say that the national unemployment rate for June was between 5.3% and 5.5%.

   (c) The **expected winnings** on a $1 lottery ticket are $0.53.

6. If you pay me $1, I will let you play the following gambling game. You drop your pencil point into the middle of a table of random digits and then look at the digit on which it lands.

   If it is a 7, you win $5.
   If it is a 1, you win $2.
   Otherwise you win nothing.

   Find the expected value of the amount you win in one play of this game. Explain in simple language what the expected value tells you about your winnings if you play the game many times. Is it to your advantage to pay $1 to play?

Sample examination II solutions

1. (a) First (rule of thumb 1), the population of California is much larger than 10 times the sample size $n = 1000$. So we can use the usual formula for the standard deviation of $\hat{p}$.

   Second (rule of thumb 2), for $n = 1000$ and $p = 0.22$ we have $np = 220$ and $n(1-p) = 780$. Both are much larger than 10, so the Normal approximation will be quite accurate.

   (b) The question concerns the *count* of 1000 California residents born outside the United States. Translate into a question about the *proportion:*

$$\text{count} \geq 250 \quad \text{is} \quad \hat{p} = \frac{\text{count}}{1000} \geq \frac{250}{1000} = 0.25$$

   The mean of $\hat{p}$ is $p = 0.22$. The standard deviation is:

$$\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(0.22)(0.78)}{1000}} = 0.0131$$

   Now use the Normal approximation to find the probability:

$$\begin{aligned} P(\hat{p} \geq 0.25) &= P\left(\frac{\hat{p} - 0.22}{0.0131} \geq \frac{0.25 - 0.22}{0.0131}\right) \\ &= P(Z \geq 2.29) \\ &= 1 - 0.9890 = 0.0110 \end{aligned}$$

2. (a) The weight x of a single child has the $N(7.5, 1.25)$ distribution. So:

$$P(x < 5.5) = P\left(\frac{x - 7.5}{1.25} < \frac{5.5 - 7.5}{1.25}\right) = P(Z < -1.60) = 0.0548$$

   (b) The mean birth weight $\bar{x}$ of a sample of three children still has mean 7.5 pounds, but its standard deviation is:

$$\frac{\sigma}{\sqrt{3}} = \frac{1.25}{\sqrt{3}} = 0.7217 \text{ pound}$$

   The probability we want is therefore:

$$P(\bar{x} < 5.5) = P\left(\frac{\bar{x} - 7.5}{0.7217} < \frac{5.5 - 7.5}{0.7217}\right) = P(Z < -2.77) = 0.0028$$

3. (a) Use Table C to see that the upper 0.025 critical value is $z^* = 1.960$. *Or,* use the 68–95–99.7 rule to approximate it by 2.

(b) The key words are "is different," indicating that the alternative hypothesis is two-sided:

$$H_0: \mu_G = \mu_B \qquad \text{vs.} \qquad H_a: \mu_G \neq \mu_B$$

Here μ_G and μ_B are the mean milk production for all cows of this breed with forage spread on the ground and in bunks, respectively.

(c) No. The margin of error in a confidence interval covers only the random sampling error due to chance variation in random sampling.

(d) The standard deviation goes down as the sample size n goes up, at the rate $\sqrt{n}$. (The standard deviation of a sample proportion $\hat{p}$ is $\sqrt{p(1-p)/n}$.) So a sample nine times larger has a standard deviation *one-third* as large.

(e) A result significant at the 5% level is in the extreme 5% of the sampling distribution; so it *may or may not* also be in the extreme 1%.

4. (a) First, $\bar{x} = 3.80$. The 95% confidence interval is:

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}} = 3.80 \pm 1.960 \frac{0.15}{\sqrt{4}} = 3.80 \pm 0.147$$

(b) "95% confidence" means that we got this interval by using a method that in 95% of all samples will produce an interval that covers the true specific gravity.

(c) $z^* = 1.282$, the upper 0.10 critical value. The 80% confidence interval is narrower than the 95% confidence interval because the critical value required is smaller.

(d) The hypotheses are:

$$H_0: \mu = 3.9 \qquad \text{vs.} \qquad H_a: \mu < 3.9$$

The test statistic is:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{3.8 - 3.9}{0.15/\sqrt{4}} = -1.33$$

and its P-value (one-sided on the low side) is:

$$P(Z \leq -1.33) = 0.0918$$

(e) There is probability 0.0918 that the mean of four readings would be as small as 3.8 if the true specific gravity were 3.9. That is, we observed a value in the smallest 9.2% of all results we could get if 3.9 were correct. This is only weak evidence that the specific gravity is less than 3.9 because a value this small would come up more than 9% of the time just by chance.

5. (a) In the long run, the coin comes up tails about 60% of the time.

(b) The interval was derived from a method that gives an interval that includes the true unemployment rate 90% of the time.

(c) If we buy many lottery tickets, the average winnings will be close to \$0.53.

6. The expected winnings are $(0.1)(\$5) + (0.1)(\$2) = \$0.70$. This represents average winnings for many plays of the game; it is not a good idea to spend \$1 on this game.

Sample final examination

This is comprehensive, covering material from (roughly) Chapters 1–12.

1. An historian examining British colonial records for the Gold Coast in Africa suspects that the death rate was higher among African miners than among European miners. In the year 1936, there were 223 deaths among 33,809 African miners and 7 deaths among 1541 European miners in the Gold Coast. (Data courtesy of Raymond Dumett, Department of History, Purdue University.)

   Consider this year as a sample from the prewar era in Africa. Is there good evidence that the proportion of African miners who died during a year was higher than the proportion of European miners who died? [State hypotheses, calculate a test statistic, give a P-value as exact as the tables in the text allow, and state your conclusion in words.]

2. An agricultural researcher reasons as follows: A heavy application of potassium fertilizer to grasslands in the spring seems to cause lush early growth but depletes the potassium before the growing season ends. So spreading the same amount of potassium over the growing season might increase yields. The researcher therefore compares two treatments: 100 pounds per acre of potassium in the spring (Treatment 1) and 50, 25, and 25 pounds per acre applied in the spring, early summer, and late summer (Treatment 2). The experiment is continued over several years because growing conditions may vary from year to year.

   The table below gives the yields, in pounds of dry matter per acre. It is known from long experience that yields vary roughly Normally. (Data from R. R. Robinson, C. L. Rhykerd, and C. F. Gross, "Potassium uptake by orchardgrass as affected by time, frequency and rate of potassium fertilization," *Agronomy Journal*, 54(1962) 351–353.)

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5
1	3902	4281	5135	5350	5746
2	3970	4271	5440	5490	6028

   (a) Do the data give good evidence that Treatment 2 leads to higher average yields? [State hypotheses, carry out a test, give a P-value as exact as the tables in the text allow, and state your conclusions in words.]

   (b) Give a 98% confidence interval for the mean increase in yield due to spreading potassium applications over the growing season.

3. Prior to an intensive TV advertising campaign, the producers of Nike athletic shoes find that 29 of a random sample of 200 upper-income adults are aware of their new leisure shoe line. A second random sample of 300 such adults is taken after the campaign. Now, 96 of the persons sampled can identify the new line.

   Give a 99% confidence interval for the increase in the proportion of upper income adults showing brand awareness.

4. Here are data on the years of schooling completed, x, and annual income, y (in thousands of dollars), for a sample of 18 40-year-old men.

Years	10	16	12	6	12	12	16	16	18
Income	28	38	16	13	25	30	35	27	28
Years	12	10	12	16	14	11	12	19	16
Income	28	26	21	34	30	21	27	29	24

   A scatterplot (don't do it) shows a generally linear relation, but with considerable scatter about the line of best fit. A computer least-squares regression program gives the output shown. (The "Coef" column gives the intercept a and slope b; the "Std Err" column gives the standard errors of these statistics. The "Residual Standard Error" is the observed standard deviation s about the regression line.)

```
                Coef     Std Err    t Value
Intercept    10.84249    5.103363   2.124577
x2            1.186813   0.372311   3.187693

Residual Standard Error = 5.02275     R-Square = 0.3884116
N = 18        F Value = 10.16139 on 1, 16 df
```

   (a) What percent of the observed variation in income is explained by the straight-line relation between income and education?

   (b) Is there strong evidence that there is a straight-line relation between education and income? (State hypotheses, carry out a test, use a table to find values between which the P-value falls, and state your conclusion.)

   (c) Consider 40-year-old men who have 16 years of education. (These are men with four years of college but no further education.) Give a 95% confidence interval for their average income.

5. Answer each of the following questions. (No explanation is needed—just a short answer.)

   (a) You are reading an article in your field that reports several statistical analyses. The article says that the P-value for a significance test is 0.045. Is this result significant at the 5% significance level?

   (b) Is the result with P-value 0.045 significant at the 1% significance level?

(c) For another significance test, the article says only that the result was significant at the 1% level. Are such results always, sometimes, or never significant at the 5% level?

(d) Reaction times of a subject to a stimulus are often strongly skewed to the right because of a few slow reaction times. You wish to test:

$$H_0: \mu_1 = \mu_2$$

where μ_1 is the mean reaction time for Stimulus 1, μ_2 for Stimulus 2. You have two independent samples, eight observations for Stimulus 1 and ten for Stimulus 2. Which, if any, of the tests you have studied can be used to test this?

(e) The article contains a 95% confidence interval. Would the margin of error in a 99% confidence interval computed from the same data be less, the same, or greater?

6. A friend in a political science course asks for your help. She just read that the correlation between the percent of eligible citizens in a city who register to vote and the percent of registered voters in the city who really do vote is 0.88. She read further that the correlation is significant ($P < 0.01$).

(a) Explain to your friend in plain language what $r = 0.88$ means.

(b) Then explain what it means to say that this correlation is statistically significant ($P < 0.01$). Use plain language—no technical terms.

(c) Now explain to you friend what the fact that $r^2 = 0.77$ says about predicting voter turnout from registration data.

7. Here are data on the percent of sugar in some popular breakfast cereals.

Product	% Sugar
All Bran	19
Alpha Bits	38
Cap'n Crunch	40
Cheerios	3
Corn Flakes	5
Golden Grahams	30
Grape Nuts Flakes	13
Post Toasties	5
Product 19	10
Raisin Bran (General Foods)	48
Raisin Bran (Kellogg)	29
Rice Krispies	8
Special K	5
Sugar Smacks	56
Wheaties	8

(a) Make a stemplot of these data. Describe the overall shape of the distribution. Are there any clear outliers?

(b) Based on your findings in part (a), choose a numerical summary for this distribution. Calculate your summary.

Sample final examination solutions

1. This is a two-sample setting with:

   Population 1 = African miners
   Population 2 = European miners

   We want to test:

   $$H_0\colon p_1 = p_2 \qquad \text{vs.} \qquad H_a\colon p_1 > p_2$$

   The two sample proportions are:

   $$\hat{p}_1 = \frac{223}{33,809} = 0.006596 \quad \text{and} \quad \hat{p}_2 = \frac{7}{1541} = 0.004543$$

   The pooled sample proportion is therefore:

   $$\begin{aligned} \hat{p} &= \frac{\text{count of deaths in both samples combined}}{\text{count of miners in both samples combined}} \\ &= \frac{223+7}{33{,}809+1541} \\ &= \frac{230}{35{,}350} = 0.006506 \end{aligned}$$

   and the test statistic is:

   $$\begin{aligned} z &= \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \\ &= \frac{0.006596 - 0.004543}{\sqrt{(0.006506)(0.993494)\left(\frac{1}{33{,}809}+\frac{1}{1541}\right)}} \\ &= \frac{0.002053}{0.0020943} = 0.980 \end{aligned}$$

   Table A gives the P-value as $1 - 0.8365 = 0.1635$. There is, surprisingly, no significant evidence that the African death rate is higher.

2. This is a *matched pairs* setting because the observations are paired by years.

   (a) The hypotheses, expressed in terms of the mean differences, Treatment 2 − Treatment 1, are:

   $$H_0: \mu = 0 \qquad \text{vs.} \qquad H_a: \mu > 0$$

   The differences are:

   $$68 \quad -10 \quad 305 \quad 140 \quad 282$$

   with

   $$\bar{x} = 157 \quad \text{and} \quad s = 135.672$$

   Apply the one-sample t test to these differences. The test statistic is:

   $$t = \frac{\bar{x} - 0}{s/\sqrt{n}} = \frac{157}{135.672/\sqrt{5}} = 2.588$$

   The P-value based on the t distribution with $n - 1 = 4$ degrees of freedom falls between 0.025 and 0.05 (using Table C). This is moderately strong evidence that Treatment 2 produces a higher mean yield.

   (b) For 98% confidence and 4 degrees of freedom, use $t^* = 3.747$. The confidence interval is:

   $$\begin{aligned} \bar{x} \pm t^* \frac{s}{\sqrt{n}} &= 157 \pm 3.747 \, \frac{135.672}{\sqrt{5}} \\ &= 157 \pm 227.3 \\ &= (-70.3, \ 384.3) \end{aligned}$$

3. There are two independent samples. We want a confidence interval for a difference between two population proportions. The sample proportions are:

   $$\hat{p}_1 = \frac{29}{200} = 0.145 \quad \text{and} \quad \hat{p}_2 = \frac{96}{300} = 0.320$$

   We can use procedures based on the Normal approximation because the population is large and:

   $$n\hat{p}_1 = 29 \quad n(1 - \hat{p}_1) = 191 \quad n\hat{p}_2 = 96 \quad n(1 - \hat{p}_2) = 204$$

   are all more than 5.

   The standard error for $\hat{p}_2 - \hat{p}_1$ is:

   $$\begin{aligned} \text{SE} &= \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}} \\ &= \sqrt{\frac{(0.145)(0.855)}{200} + \frac{(0.320)(0.680)}{300}} \\ &= \sqrt{0.0013452} = 0.03668 \end{aligned}$$

The 99% confidence interval for $p_2 - p_1$ is:

$$\begin{aligned}(\hat{p}_2 - \hat{p}_1) \pm z^*\text{SE} &= (0.320 - 0.145) \pm (2.576)(0.03668) \\ &= 0.175 \pm 0.0945 \\ &= (0.0805,\ 0.2695)\end{aligned}$$

We are 99% confident that between 8% and 27% of upper-income adults are aware of the new shoe line.

4. (a) The output says `R-Square = 0.3884116`, so the linear relationship explains 38.8% of the observed variation in income.

   (b) The null hypothesis of "no relation" says that the slope of the true regression line is 0. The hypotheses are:

   $$H_0 : \beta = 0 \qquad \text{vs.} \qquad H_a : \beta \neq 0$$

   The computer output shows that the t statistic for the test is $t = 3.187693$. The degrees of freedom are $n - 2 = 16$. From Table C we see that t falls between the 0.0025 and 0.005 upper critical values of $t(16)$. Doubling these values because H_a is two-sided, $0.005 < P < 0.01$. There is strong evidence that a linear relationship exists.

   (c) The predicted mean income for $x = 16$ is:

   $$\hat{y} = 10.84249 + (1.186813)(16) = 29.831$$

   or \$29,831. The rest of this is a bit tedious by hand, so consider your options. (Not all students may be capable of handling this on a test.) Here goes. Using a calculator for $\bar{x}$ and s_x gives $\bar{x} = 13.33$ and:

   $$\sum (x - \bar{x})^2 = (n-1)s_x^2 = (17)(3.27198)^2 = 182$$

   The proper standard error for estimating the mean income is

   $$\begin{aligned}\text{SE}_{\hat{\mu}} &= s\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum (x - \bar{x})^2}} \\ &= 5.02275\sqrt{\frac{1}{18} + \frac{(16 - 13.33)^2}{182}} \\ &= 5.02275\sqrt{0.0946275} = 1.5451\end{aligned}$$

   The 90% confidence interval is therefore:

   $$\begin{aligned}\hat{y} \pm t^*\text{SE}_{\hat{\mu}} &= 29.831 \pm (2.120)(1.5451) \\ &= 29.831 \pm 3.276 \\ &= (26.555,\ 33.107)\end{aligned}$$

   or \$26,555 to \$33,107.

5. (a) Yes. The P-value is less than 0.05.

   (b) No. The P-value is greater than 0.01.

   (c) A result significant at the 1% level lies in the extreme 1% of a sampling distribution. This is certainly in the extreme 5%, so the result is always significant at the 5% level.

   (d) The samples are small and the distributions are strongly skewed. It would be unwise to use the t test in this setting. A nonparametric test (if these have been studied) might be appropriate.

   (e) The margin of error would be greater. Higher confidence is paid for with a greater margin of error.

6. (a) There is a strong tendency for both quantities—percent registered, and the percent who turn out—to both be high or both be low. That is, when one is high, the other is usually also high, and when one is low, the other is usually low.

   (b) If there was no relationship between the two quantities, a correlation this high would be observed less than 1% of the time.

   (c) There will be variation in voter turnout from one city to the next. If we know the percent registered, we can account for (explain) about 77% of that variation.

7. (a) The distribution is right-skewed, with no particular outliers.

0	355588
1	039
2	9
3	08
4	08
5	6

   (b) Because of the skewness, use the five-number summary: 3 5 13 38 56.

4 SOLUTIONS TO EXERCISES

About these solutions

The solutions that follow were prepared by Darryl Nester, with the exception of Chapter 16 (prepared by Tim Hesterberg). In some cases, solutions were based on those prepared for earlier editions of *IPS*. Erin Leatherman provided extensive review (and correction) of the solutions, especially focusing on those which were new (or revised) from previous editions. In spite of the care that went into that process, I might have missed a subtle change in an exercise that should have resulted in a change in the solution. Should you discover any errors or have any comments about these solutions (or the odd answers, in the back of the text), please report them to me:

Darryl Nester
Bluffton University
Bluffton, Ohio 45817
e-mail: `nesterd@bluffton.edu`
WWW: `www.bluffton.edu/~nesterd`

Acknowledgments

I should mention the software I used in putting these solutions together:

- For typesetting: TEX — specifically, Textures, from Blue Sky Software.
- For the graphs: DeltaGraph (SPSS) and Adobe Illustrator.
- For statistical analysis: primarily Minitab and Excel, along with S-PLUS (for Chapter 17), SAS (for Chapter 14), and some freeware and shareware software (G•Power and GLMStat). Additionally, I used the TI-83 calculator from Texas Instruments.

Note: The solutions given to the applet exercises, and the sample output screens, were based on the current versions of the applets at the time the solutions were written. As revisions are made to these applets, the appearance of the output screens (and in some cases, the answers) may change. Screenshots were taken on a computer running Macintosh OS X, but that should have no significant impact on the appearance.

Using the table of random digits

Grading SRSs chosen from Table B is complicated by the fact that students can find some creative ways to (mis)use the table. Some approaches are not mistakes, but may lead to different students having different "right" answers. Correct answers will vary based on

- The line in the table on which they begin (you may want to specify one if the text does not).
- Whether they start with, e.g., 00 or 01.
- Whether they assign multiple labels to each unit.
- Whether they assign labels across the rows or down the columns (nearly all lists in the text are alphabetized down the columns).

Some approaches can potentially lead to wrong answers. Mistakes to watch out for include the following:

- They may forget that all labels must be the same length (e.g., assigning labels such as 0, 1, 2, . . . , 9, 10, . . . rather than 00, 01, 02, . . .).
- In assigning multiple labels, they may not give the same number of labels to all units. For example, if there are 30 units, they may try to use up all the two-digit numbers, thus assigning four labels to the first 10 units and only three to the remaining 20.

As an alternative to using the random digits in Table B, students can pick a random sample by generating (pseudo)-random numbers, using software (like Excel) or a calculator. With many, if not all, calculators, the sequence of random numbers produced is determined by a "seed value" (which can be specified by the user). Rather than pointing students to a particular line of Table B, you could specify a seed value for generating random numbers, so that all students would obtain the same results (if all are using the same model of calculator).

On a TI-83, for example, after executing the command `0→rand`, the `rand` command will produce the sequence (rounded to four decimals) 0.9436, 0.9083, 0.1467, . . ., while `1→rand` initiates the sequence 0.7456, 0.8559, 0.2254, . . . So to choose, say, an SRS of size 10 from 30 subjects, use the command `0→rand` to set the seed, and then type `1+30*rand`, and press ENTER repeatedly. Ignoring the decimal portion of the resulting numbers, this produces the sample

29, 28, 5, 15, 13, 23, 2, 11, 30, 7

(Generally, to generate random numbers from 1 to n, use the command `1+`n`*rand` and ignore the decimal portion of the result.)

Using statistical software

The use of computer software or a calculator is a must for all but the most cursory treatment of the material in this text. Be aware of the following considerations:

- *Standard deviations:* Students may be confused by software that gives both the so-called "sample standard deviation" (the one used in the text) and the "population standard deviation" (dividing by n rather than $n - 1$). Symbolically, the former is usually given as s and the latter as σ (sigma), but the distinction is not always clear. For example, many computer spreadsheets have a command such as "`STDEV(...)`" to

compute standard deviations, but you may need to check the manual to find out which kind it is.
As a quick check: For the numbers $1, 2, 3$, $s = 1$ while $\sigma \doteq 0.8165$. In general, if two values are given, the larger one is s and the smaller is σ. If only one value is given, and it is the wrong one, use the relationship $s = \sigma\sqrt{\frac{n}{n-1}}$.

- *Stemplots:* The various choices one can make in creating a stemplot (e.g., rounding or truncating the data) have already been mentioned. Minitab opts for truncation over rounding, so all of the solutions in this guide show truncated-data stemplots (except for exercises that instructed students to round). This usually makes little difference in the overall appearance of the stemplot.

- *Significant digits in these solutions:* Most numerical answers in these solutions (and in the odd-numbered answers in the back of the text) are reported to four significant figures. In many cases, that is an absurd overrepresentation of the accuracy of those numbers, but those digits are provided to give students a better "check" on their answers.

  This extra accuracy is a double-edged sword, however, as a student might have a correct answer that does not agree with all the digits in the printed answer. This might occur (rarely, I hope) because my answer is wrong, but it may be due to rounding, differences in software accuracy, or use of an approximation. For example, in reporting binomial probabilities for exercises in Chapter 5, I have listed four or five answers for some problems: exact answers and Normal approximations (with or without the continuity correction, and computed with software or using Table A). In the back-of-the-text answers, only one answer is given, so that students may have to be satisfied with being "close."

- *Quartiles and five-number summaries:* Methods of computing quartiles vary between different packages. Some use the approach given in the text (that is, Q_1 is the median of all the numbers below the location of the overall median, etc.), while others use a more complicated approach. For the numbers $1, 2, 3, 4$, for example, we would have $Q_1 = 1.5$ and $Q_3 = 3.5$, but Minitab reports these as 1.25 and 3.75, respectively, while Excel reports 1.75 and 3.25.

  In these solutions (and the odd-numbered answers in the back of the text), I opted to report five-number summaries are they would be found using the text's method. Because I used Minitab for most of the analysis in these solutions, I wrote a Minitab macro to compute quartiles the *IPS* way. This and other macros are available on my Web site.

- *Boxplots:* Some programs that draw boxplots use the convention that the "whiskers" extend to the lower and upper deciles (the 10th and 90th percentiles) rather than to the minimum and maximum. The information conveyed by such a boxplot is essentially the same as by those described and used in the text; if students use such software, point out that they can still draw the same conclusions about the underlying distribution.