PART I REVIEW

The first and most important question to ask about any statistical study is "Where do the data come from?" Chapter 1 addressed this question. The distinction between observational and experimental data is a key part of the answer. Good statistics starts with good designs for producing data. Chapters 2, 3, and 4 discussed sampling, the art of choosing part of a population to represent the whole. Figure I.1 summarizes the big idea of a simple random sample. Chapters 5 and 6 dealt with the statistical aspects of designing experiments, studies that impose some treatment in order to learn about the response. The big idea is the randomized comparative experiment. Figure I.2 outlines the simplest design.

Random sampling and randomized comparative experiments are perhaps the most important statistical inventions of the twentieth century. Both were slow to gain acceptance, and you will still see many voluntary response samples and uncontrolled experiments. Both random samples and randomized experiments involve the deliberate use of chance to eliminate bias and produce a regular pattern of outcomes. The regular pattern allows us to give margins of error, make confidence statements, and assess the statistical significance of conclusions based on samples or experiments.

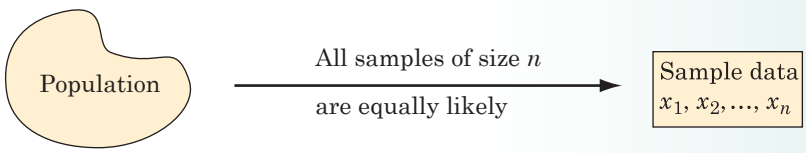

Figure I.1 The idea of a simple random sample.

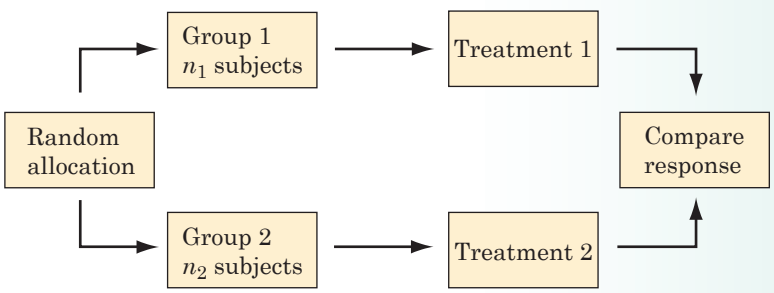

Figure I.2 The idea of a randomized comparative experiment.

When we collect data about people, ethical issues can be important. Chapter 7 discussed these issues and introduced three principles that apply to any study with human subjects. The last step in producing data is to measure the characteristics of interest to produce numbers we can work with. Measurement was the subject of Chapter 8. "Where do the data come from?" is the first question we should ask about a study, and "Do the numbers make sense?" is the second. Chapter 9 encouraged the valuable habit of looking skeptically at numbers before accepting what they seem to say.

PART I SUMMARY

Here are the most important skills you should have acquired after reading Chapters 1 to 9.

A. DATA

1. Recognize the individuals and variables in a statistical study.

2. Distinguish observational from experimental studies.

3. Identify sample surveys, censuses, and experiments.

B. SAMPLING

1. Identify the population in a sampling situation.

2. Recognize bias due to voluntary response samples and other inferior sampling methods.

3. Use Table A of random digits to select a simple random sample (SRS) from a population.

4. Explain how sample surveys deal with bias and variability in their conclusions. Explain in simple language what the margin of error for a sample survey result tells us and what "95% confidence" means.

5. Use the quick method to get an approximate margin of error for 95% confidence.

6. Understand the distinction between sampling errors and nonsampling errors. Recognize the presence of undercoverage and nonresponse as sources of error in a sample survey. Recognize the effect of the wording of questions on the responses.

7. Use random digits to select a stratified random sample from a population when the strata are identified.

C. EXPERIMENTS

1. Explain the differences between observational studies and experiments.

2. Identify the explanatory variables, treatments, response variables, and subjects in an experiment.

3. Recognize bias due to confounding of explanatory variables with lurking variables in either an observational study or an experiment.

4. Outline the design of a completely randomized experiment using a diagram like that in Figure I.2. Such a diagram should show the sizes of the groups, the specific treatments, and the response variable.

5. Use random digits to carry out the random assignment of subjects to groups in a completely randomized experiment.

6. Make use of matched pairs or block designs when appropriate.

7. Recognize the placebo effect.

8. Recognize when the double-blind technique should be used.

9. Be aware of weaknesses in an experiment, especially in the ability to generalize its conclusions.

10. Explain why a randomized comparative experiment can give good evidence for cause-and-effect relationships.

11. Explain the meaning of statistical significance.

D. OTHER TOPICS

1. Explain the three first principles of data ethics (protect subjects from harm, informed consent, and confidentiality). Discuss how they might apply in specific settings.

2. Explain how measuring leads to clearly defined variables in specific settings.

3. Evaluate the validity of a variable as a measure of a given characteristic, including predictive validity.

4. Explain how to reduce bias and improve reliability in measurement.

5. Recognize inconsistent numbers, implausible numbers, numbers so good they are suspicious, and arithmetic mistakes.

6. Calculate percentage increase or decrease correctly.

PART I REVIEW EXERCISES

Review exercises are short and straightforward exercises that help you solidify the basic ideas and skills in each part of this book. We have provided "hints" that indicate where you can find the relevant material for the odd-numbered problems.

I.1 Know these terms. A friend who knows no statistics has encountered

some statistical terms in reading for her psychology course. Explain each of the following terms in one or two simple sentences.

(a) Simple random sample. (*Hint*: See pages 24–26.)

(b) 95% confidence. (*Hint*: See pages 47–49.)

(c) Statistically significant. (*Hint*: See page 103.)

(d) Informed consent. (*Hint*: See pages 144–146.)

I.2 Know these terms. A friend who knows no statistics has encountered some statistical terms in her biology course. Explain each of the following terms in one or two simple sentences.

(a) Observational study.

(b) Double-blind.

(c) Nonsampling error.

(d) Matched pairs.

I.3 A biased sample. A student plans to collect student opinions for a class assignment. He decides to post on his Facebook page a request for people to text him their opinions. Explain why this sampling method is almost certainly biased. (*Hint*: See pages 21–24.)

I.4 Select an SRS. A student at a large university wants to study the responses that students receive when calling an academic department for information. She selects an SRS of four departments from the following list for her study. Use software or Table A at line 115 to do this.

Accounting
Architecture
Art
Biology
Business Administration
Chemistry
Communication
Computer Science
Dance
Economics
Electrical Engineering
Elementary Education
English
Foreign Languages
History
Horticulture
International Studies
Marketing
Mathematics
Music
Natural Resources
Nursing
Pharmacy
Philosophy
Physics
Political Science
Pre-med
Psychology
Sociology
Veterinary Science

I.5 Select an SRS. The faculty grievance system at a university specifies that a five-member hearing panel shall be drawn at random from the 25-member grievance committee. Use software or Table A at line 128 to draw an SRS of size four from the following committee members.

Anis
Atlas
Bailey
Banks
Edwards
Frazier
Gardner
Guthrie
Kowalski
Kupka
Lehman

Leonard
Mee
Michelson
Morgan
Murphy
Pagolu
Ramsey
Ray
Sall
Utlaut
Valente
Weese
Wendelberger
Woodall

(*Hint*: See pages 26–29.)

I.6 Errors in surveys. Give an example of a source of nonsampling error in a sample survey. Then give an example of a source of sampling error.

I.7 Errors in surveys. An overnight opinion poll calls randomly selected telephone numbers. This polling method misses all people without a phone. Is this a source of nonsampling error or of sampling error? Does the poll's announced margin of error take this source of error into account? (*Hint*: See pages 64–70.)

I.8 Errors in surveys. A college chooses an SRS of 100 students from the registrar's list of all undergraduates to interview about student life. If it selected two SRSs of 100 students at the same time, the two samples would give somewhat different results. Is this variation a source of sampling error or of nonsampling error? Does the survey's announced margin of error take this source of error into account?

I.9 Errors in surveys. Exercises I.7 and I.8 each mention a source of error in a sample survey. Would each kind of error be reduced by doubling the size of the sample with no other changes in the survey procedures? Explain your answers. (*Hint*: See pages 43–47 and pages 65–70.)

I.10 Errors in surveys. A Gallup Poll found that 46% of American smartphone users agree with the statement, "I can't imagine my life without my smartphone." The Gallup press release says:

Results of attitudes and behaviors of smartphone usage are based on 15,747 members of the Gallup Panel who have smartphones. . . . For results based on this sample, one can say that the margin of sampling error is ±1 percentage point at the 95% confidence level.

The release also points out that this margin of error is due only to sampling error. Give one example of a source of error in the poll result that is *not* included in this margin of error.

I.11 Find the margin of error. On April 24, 2015, "Bruce Jenner—The Interview" was aired. The interview was between Diane Sawyer and Bruce Jenner and discussed Jenner's experience as a transgender person. Jenner came out to the public as Caitlyn Jenner on June 1, 2015. An NBC News Online Survey asked a national sample of 2153 adults aged 18 and over, "Caitlyn Jenner, the transgender Olympic champion formerly known as Bruce Jenner, recently revealed her transition on the cover of *Vanity Fair*. Do you think Caitlyn Jenner's public transition will help society become more accepting of transgender

people?" Twenty percent of those surveyed replied that Jenner's public transition "will help a lot," while 46% responded "will help a little."

(a) What is the population for this 2015 survey? (*Hint*: See pages 8–10.)

(b) Assuming the 2015 survey used random sampling, use the quick and approximate method to find a margin of error. Then give a complete confidence statement for a conclusion about the population. (*Hint*: See pages 45–47.)

I.12 Find the margin of error. A May 2015 Quinnipiac University Poll asked 2105 American adults (18 years and over), "As you may know, David Letterman has retired. Thinking about other late-night television talk show hosts who are currently on TV, who is your favorite?" Jimmy Fallon came out on top, with 20% of Americans favoring him over others.

(a) What is the population for this sample survey?

(b) Assuming the sample was a random sample, use the quick and approximate method to find a margin of error. Make a confidence statement about the opinion of the population.

I.13 What kind of sample? At a party, there are 30 students over age 21 and 20 students under age 21. You choose at random six of those over 21 and separately choose at random four of those under 21 to interview about attitudes toward alcohol. You have given every student at the party the same chance to be interviewed: what is that chance? Why is your sample not an SRS? What is this kind of sample called? (*Hint*: See pages 24–27 and pages 73–75.)

I.14 Design an experiment. A university's Department of Statistics wants to attract more majors. It prepares two advertising brochures. Brochure A stresses the intellectual excitement of statistics. Brochure B stresses how much money statisticians make. Which will be more attractive to first-year students? You have a questionnaire to measure interest in majoring in statistics, and you have 50 first-year students to work with. Outline the design of an experiment to decide which brochure works better.

I.15 Design an experiment. Gary and Greg share an apartment 10 miles from campus. Gary thinks that the fastest way to get to campus is to take the shortest route, which involves taking several side streets. Greg thinks the fastest way is to take the route with the highest speed limits, which involves taking the highway most of the way but is two miles longer than Gary's route. You recruit 20 friends who are willing to try either method and time how long it takes them to arrive on campus. Outline the design of an experiment to decide which method takes less time. (*Hint*: See pages 98–101.)

Exercises I.16 to I.19 are based on an article in the *Journal of the American Medical Association* that compares the use of antibiotics to treat uncomplicated acute appendicitis instead of surgery with an appendectomy. Here is information from the article's summary:

> **Design** Randomized clinical trial conducted in Finland between 2009 and 2012.

Participants A total of 530 patients aged 18 to 60 years with uncomplicated appendicitis.

Intervention Participants were randomized to receive either an antibiotic therapy ($n = 257$) or a standard appendectomy ($n = 273$).

Results For all but one of the patients receiving surgery, the appendectomy was successful, resulting in a success rate of 99.6%. In the antibiotic group, 70 patients required an appendectomy within 1 year of initial presentation of appendicitis, resulting in a success rate (no recurrent symptoms after 1 year) of 72.7%.

I.16 Know these terms. Explain in one sentence each what "randomized" and "clinical trial" mean in the description of the design of the study.

I.17 Experiment basics. Identify the subjects, the explanatory variable, and several response variables for this study. (*Hint*: See pages 93–95.)

I.18 Design an experiment. Use a diagram to outline the design of the experiment in this medical study.

I.19 Ethics. What are the three first principles of data ethics? Explain briefly what the medical study must do to apply each of these principles. (*Hint*: See pages 141–143.)

I.20 Measuring. Joni wants to measure the degree to which male college students belong to the political left. She decides simply to measure the length of their hair–longer hair will mean more left-wing.

(a) Is this method likely to be reliable? Why?

(b) This measurement appears to be invalid. Why?

(c) Nevertheless, it is possible that measuring politics by hair length might have some predictive validity. Explain how this could happen.

I.21 Reliability. You are laboring through a chemistry laboratory assignment in which you measure the conductivity of a solution. What does it mean for your measurement to be reliable? How can you improve the reliability of your final result? (*Hint*: See pages 174–178.)

I.22 Observation or experiment? The Nurses' Health Study has queried a sample of more than 100,000 female registered nurses every two years since 1976. Beginning in 1980, the study asked questions about diet, including alcohol consumption. The researchers concluded that "light-to-moderate drinkers had a significantly lower risk of death" than either nondrinkers or heavy drinkers.

(a) Is the Nurses' Health Study an observational study or an experiment? Why?

(b) What does "significant" mean in a statistical report?

(c) Suggest some lurking variables that might explain why moderate drinkers have lower death rates than nondrinkers. (The study did adjust for these variables.)

I.23 Observation or experiment? In a study of the relationship between physical fitness and personality, middle-aged college faculty who have volunteered for an exercise program are divided into low-fitness and high-fitness groups on the basis of a physical examination. All subjects then take

a personality test. The high-fitness group has a higher average score for "self-confidence."

(a) Is this an observational study or an experiment? Why? (*Hint*: See pages 7–8 and 12–24.)

(b) We cannot conclude that higher fitness causes higher self-confidence. Suggest other relationships among these variables and perhaps lurking variables that might explain the higher self-confidence of the high-fitness group. (*Hint*: See pages 12–14 and 95–98.)

I.24 Percents up and down. Between January 14, 2015, and July 14, 2015, the average price of regular gasoline increased from $2.08 per gallon to $2.78 per gallon.

(a) Verify that this is a 34% increase in price.

(b) If the price of gasoline decreases by 34% from its July 14, 2015, level of $2.78 per gallon, what would be the new price? Notice that a 34% increase followed by a 34% decrease does not take us back to the starting point.

I.25 Percentage decrease. On Monday, September 10, 2001 (the day before the September 11 attacks), the NASDAQ stock index closed the day at 1695. By the end of Monday, September 17, 2001 (the first full day of trading after the attacks), the NASDAQ stock index had dropped to 1580. By what percentage did the index drop? (*Hint*: See pages 193–196.)

I.26 An implausible number? *Newsweek* once said in a story that a woman who is not currently married at age 40 has a better chance of being killed by a terrorist than of getting married. Do you think this is plausible? What kind of data would help you check this claim?

PART I PROJECTS

Projects are longer exercises that require gathering information or producing data and that emphasize writing a short essay to describe your work. Many are suitable for teams of students.

Project 1. Design your own sample survey. Choose an issue of current interest to students at your school. Prepare a short (no more than five questions) questionnaire to determine opinions on this issue. Choose a sample of about 25 students, administer your questionnaire, and write a brief description of your findings. Also write a short discussion of your experiences in designing and carrying out the survey.

(Although 25 students are too few for you to be statistically confident of your results, this project centers on the practical work of a survey. You must first identify a population; if it is not possible to reach a wider student population, use students enrolled in this course. How did you choose your sample? Did the subjects find your questions clear? Did you write the questions so that it was

easy to tabulate the responses? At the end, did you wish you had asked different questions?)

Project 2. Measuring. Is the right hand generally stronger than the left in right-handed people? You can crudely measure hand strength by placing a bathroom scale on a shelf with the end protruding, then squeezing the scale between the thumb and the four fingers. The reading of the scale shows the force exerted. Using several right-handed subjects, try to determine whether this method of measuring hand strength is reliable. Write an account of your findings. For example, did you find that subjects used different grips, so that careful instructions were needed to get a consistent way of measuring? Prepare written instructions for subjects.

Project 3. Experimenting. After you or other members of your team have refined the measurement of hand strength in the previous project, carry out the matched pairs experiment of Project 2 with at least 10 subjects. Write a report that describes the randomization, gives the data, reports the differences in strength (right hand minus left hand), and says whether your small experiment seems to show that the right hand is stronger, on the average.

Project 4. Describe a medical study. Go to the website of the *Journal of the American Medical Association* (http://jama.ama-assn.org). Unlike the *New England Journal of Medicine* (http://content.nejm.org), *JAMA* makes the full text of some articles freely available online. Select an article from the current issue or from a past issue that describes a study whose topic interests you. Write a newspaper article that summarizes the design and the findings of the study. (Be sure to include statistical aspects, such as observational study versus experiment and any randomization used. News accounts often neglect these facts.)

Project 5. Do you drive an SUV? Are staff at your college more or less likely than students to drive an SUV? Design and carry out a study to find out, and write a report describing your design and your findings. You must first be clear about what "SUV" means so that each car is clearly an SUV or not. Then you must locate a suitable sample of cars–perhaps from a student parking area and a staff parking area. If the areas are large, you will want to sample the cars rather than look at all of them. Consider using a systematic sample (Exercise 4.34, page 90).

Project 6. Data ethics. Locate a news discussion of an ethical issue that concerns statistical studies. Write your own brief summary of the debate and any conclusions you feel you can reach.

Here is an example of one way to approach this project. Testing new drugs on human subjects continues to be an ongoing concern in medical studies. How does one balance the need for knowledge with protecting subjects from possible harm? Searching the archives at the website of the *New York Times* (www.nytimes.com) for "experiments and ethics" (to use

the *New York Times* search engine, you should enter "+experiments +ethics"), one finds many articles, including a promising one in the issue of May 7, 2015. To read the article, you may be able to link to it directly online, but if not, you will have to either pay a fee or go to the library. You could also try searching the Web with Google. We entered "drugs and human guinea pigs" and found some possible leads.

Project 7. Measuring income. What is the "income" of a household? Household income may determine eligibility for government programs that assist "low-income" people. Income statistics also have political effects. Political conservatives often argue that government data overstate how many people are poor because the data include only money income, leaving out the value of food stamps and subsidized housing. Political liberals reply that the government should measure money income so it can see how many people need help.

You are on the staff of a member of Congress who is considering new welfare legislation. Write an exact definition of "income" for the purpose of determining which households are eligible for welfare. A short essay will be needed. Will you include nonmoney income such as the value of food stamps or subsidized housing? Will you allow deductions for the cost of child care needed to permit the parent to work? What about assets that are worth a lot but do not produce income, such as a house?