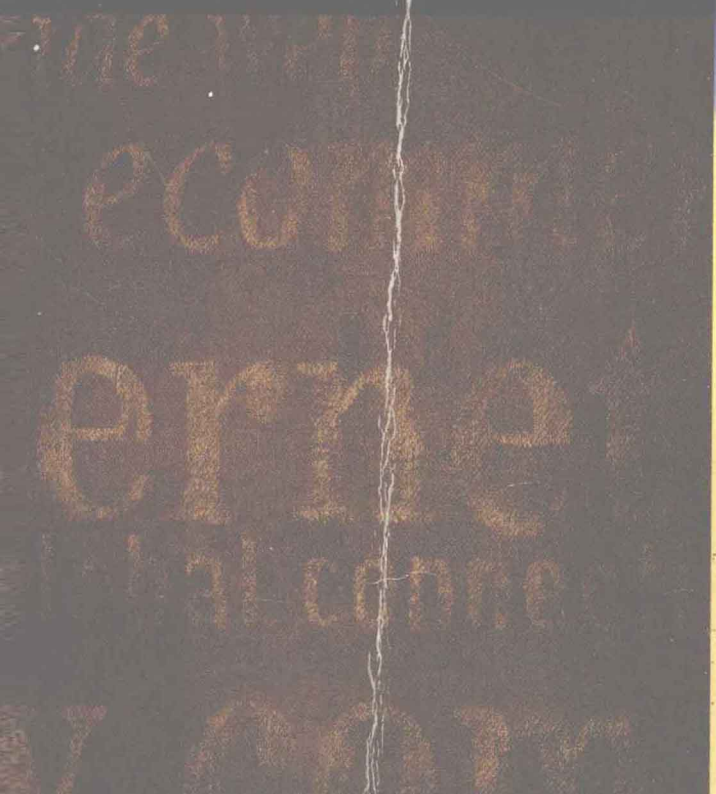
STAT 111

University of Massachusetts

Kuby/Johnson

THOMSON LEARNING

CUSTOM PUBLISHING



STAT 111

University of Massachusetts

Patricia Kuby
Monroe Community College

Robert Johnson
Monroe Community College

BROOKS/COLE
THOMSON LEARNING

Printed in the United States of America

Brooks.Cole
511 Forest Lodge Road
Pacific Grove, CA 93950
USA

For information about our products, contact us:
Thomson Learning Academic Resource Center
1-800-423-0563
http://www.brookcole.com

International Headquarters
Thomson Learning
International Division
290 Harbor Drive, 2nd Floor
Stamford, CT 06902-7477
USA

UK/Europe/Middle East/South Africa
Thomson Learning
Berkshire House
168-173 High Holborn
London WCIV 7AA

Asia
Thomson Learning
60 Albert Street, #15-01
Albert Complex
Singapore 189969

Canada
Nelson Thomson Learning
1120 Birchmount Road
Toronto, Ontario MIK 5G4
Canada
United Kingdom

ISBN 0-534-14013-0

The Adaptable Courseware Program consists of products and additions to existing Brooks/Cole products that are produced from camera-ready copy. Peer review, class testing, and accuracy are primarily the responsibility of the author(s).

Custom Contents

STAT 111

University of Massachusetts

Patricia Kuby
Monroe Community College

Robert Johnson
Monroe Community College

BROOKS/COLE
THOMSON LEARNING

1 STATISTICS

CHAPTER OUTLINE

CHAPTER OBJECTIVES

The purpose of this introductory chapter is to (1) create an initial image of the field of statistics, an image that will grow and develop; (2) introduce several basic vocabulary words used in studying statistics, such as *population, variable*, and *statistic*; and (3) present some initial ideas and concerns about the processes used to obtain sample data.

1.1 WHAT IS STATISTICS?

Statistics is the universal language of the sciences. As potential users of statistics, we need to master both the "science" and the "art" of using statistical methodology correctly. Careful use of statistical methods will enable us to obtain accurate information from data. These methods include (1) carefully defining the situation, (2) gathering data, (3) accurately summarizing the data, and (4) deriving and communicating meaningful conclusions.

Statistics

Statistics involves information, numbers to summarize this information, and their interpretation. The word ***statistics*** has different meanings to people of varied backgrounds and interests. To some people it is a field of "hocus-pocus" whereby a person in the know overwhelms the rest of us. To others it is a way of collecting and displaying large amounts of information. And to still another group it is a way of "making decisions in the face of uncertainty." In the proper perspective, each of these points of view is correct.

Descriptive and inferential statistics

The field of statistics can be roughly subdivided into two areas: descriptive statistics and inferential statistics. **Descriptive statistics** is what most people think of when they hear the word *statistics*. It includes the collection, presentation, and description of sample data. The term **inferential statistics** refers to the technique of interpreting the values resulting from the descriptive techniques and making decisions and drawing conclusions about the sampled population.

Statistics is more than just numbers—it is data, what is done to data, what is learned from the data, and the resulting conclusions. Let's use the following definition:

STATISTICS

The science of collecting, describing, and interpreting data.

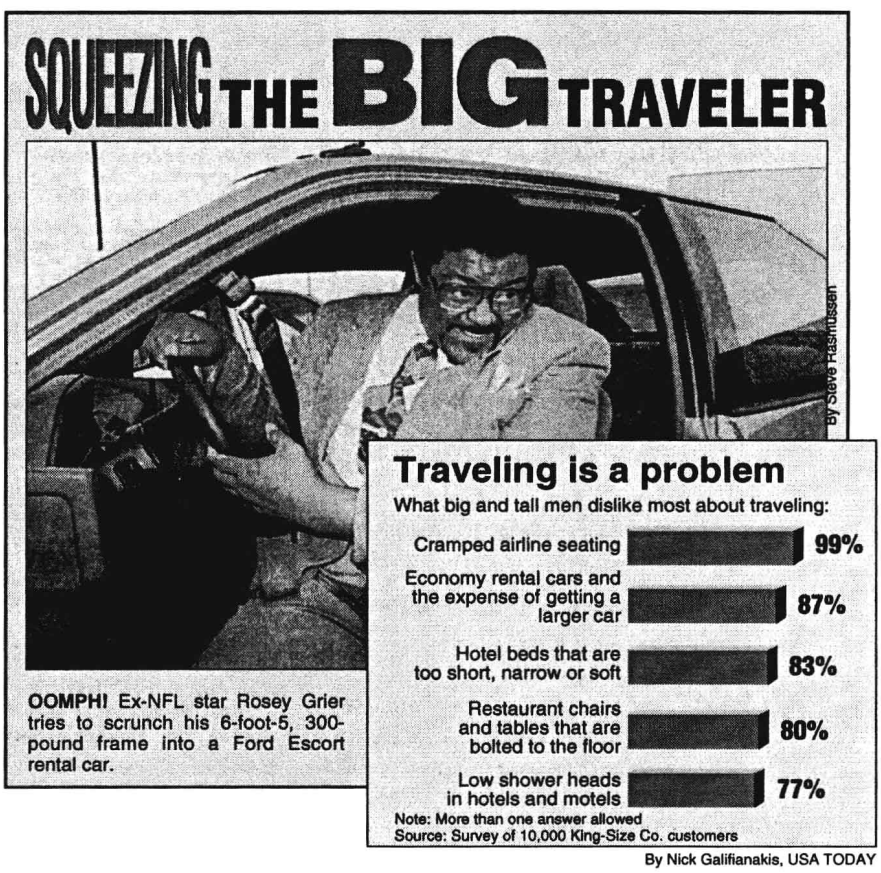
Before we begin our detailed study, let's look at a few illustrations of how and when statistics can be applied.

Case Study 1-1

Measuring Physical Discomfort

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

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



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

EXERCISE 1.1

Refer to "Traveling is a problem."

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

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

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

Airline seats have never been roomy. "They were originally designed to fit a very well-known athlete: (jockey) Willie Shoemaker," jokes Ed Perkins, editor of Consumer Reports Travel Letter, which measures the distance between seats every two years. Over the past 20 years, airlines have gradually squeezed about 10% more seats into their jets. All the while, they have inched the rows on all jets closer together. Rows are 4 inches closer together than they were 20 years ago.

Case Study 1-2

Explaining Our Sexual Behavior

Sex has been said to be "everybody's favorite topic."

Sex in the Snoring '90s

A survey of American men shocks the nation with what it didn't find. They're the sorts of questions any average American male asks: How many times a night should I be having sex? In what grade should I start? ... But there is nowhere to turn for answers. If you wanted to know something the government considered important, like the size of the sardine catch, you could look it up in a minute. However, when you don't pay taxes on it, the government has no reason to keep track of it. Now, in what the authors call the first scientifically valid survey of its kind, 3,321 American men in their 20s and 30s were questioned about their sexual practices, and for the first time in the history of sex research, the results are not astounding.

THE REPORT

In the most complete study since the Kinsey report, a new survey charts American sexual activity. The results are—yawn—striking.

1
is the median number of times men said they had sex per week.

7.3
is the median number of sexual partners, over lifetime, per man.
(Guttmacher, 1993)

INFIDELITY

How faithful are the women of America? It depends on whom you ask.Two studies produced very different answers.

26%
Married women who have had extramarital affairs.
(Janus, 1993)

39%
Married women who have cheated on their husbands.
(Cosmopolitan Reader's Survey, 1993)

One thing the Battelle study did reveal was the huge philosophical gap between two schools of sex research. One group conducts personal interviews with subjects chosen on the basis of census data to represent a scientific sample of the population, which is supposed to provide better statistical validity. The other uses anonymous mail-in questionnaires, which seem to elicit juicier responses. The former generally publish papers in scientific journals, while the latter typically write articles in women's magazines and best-selling books.

Janus, who distributed his survey through graduate students and in piles left in doctors' offices, asserts that people are more likely to be truthful on an anonymous questionnaire than in face-to-face interviews. The Battelle group's interviews were all conducted in the respondents' homes, by female interviewers who knocked on the door with no previous introduction. Thirty percent of the men contacted in this way refused to participate. Janus suggests that the missing information is probably in that 30 percent. On the other hand, the people motivated to pick up, fill out and mail in a questionnaire on sex represent a self-selected group—selected, in part, by being interested in sex. Janus' data also show that his subjects were considerably richer, less Protestant and more Jewish than the nation as a whole.

Source: Jerry Adler, *Newsweek*, April 26, 1993.

EXERCISE 1.2

Refer to "Sex in the Snoring '90s."

a. What group was surveyed?
b. Why do seemingly similar surveys produce such different results?

The uses of statistics are unlimited. It is much harder to name a field in which statistics is not used than it is to name one in which statistics plays an integral part. The following are a few examples of how and where statistics are used:

- In education descriptive statistics are frequently used to describe test results.
- In science the data resulting from experiments must be collected and analyzed.
- In government many kinds of statistical data are collected all the time. In fact, the U.S. government is probably the world's greatest collector of statistical data.

A very important part of the statistical process is that of studying the statistical results and formulating appropriate conclusions. These conclusions must then be communicated to others. Nothing is gained from research unless the findings are shared with others.

Statistics are being reported everywhere—newspapers, magazines, television. We read and hear about all kinds of new research results, especially health-related findings.

EXERCISES

1.3 Determine which of the following statements is descriptive in nature and which is inferential. Refer to "Traveling is a problem" in Case Study 1-1 (p. 5).

a. 99% of all big and tall travelers dislike cramped airline seating the most.

b. 99% of the 10,000 King-Size Co. customers dislike cramped airline seating the most.

1.4 Refer to the USA Snapshot® "Reacting to crime" below.

a. What group of people were polled?

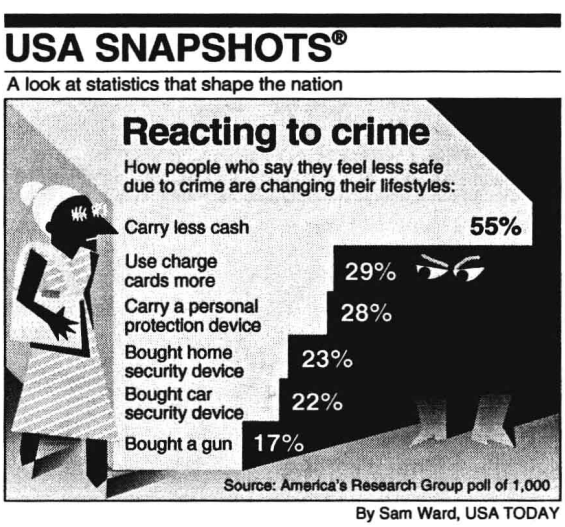
b. How many people were polled?

c. What information was obtained from each person?

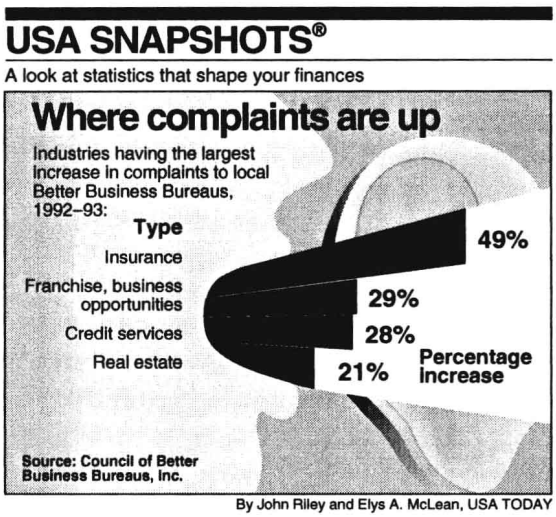
d. Explain the meaning of "55% Carry less cash."

e. How many people answered "Carry less cash"?

f. Why do the reported values (55%, 29%, 28%, ...) add up to more than 100%?



1.5 Percentages do not always tell the whole story.



a. What impression do you get from the statistics (percentages) reported in "Where complaints are up"?
b. Suppose that in 1992 there were 39 and in 1993 there were 58 complaints. What is the percentage of increase?
c. Suppose that in 1992 there were 490 and in 1993 there were 593 complaints. What is the percentage of increase?
d. Compare the increase in complaints listed in (b) and (c).
e. Do you have the same impression about the information presented in the bar graph after seeing the number of complaints in (b) and (c)? Explain.

1.6 The importance of eating healthful foods and having a healthful lifestyle is a topic in today's newspapers and magazines.

Pass a bagel, hold the coffee

By Nancy Hellmich
USA Today

The USA's breakfast of champions now includes more bananas and bagels — and less coffee.

It studied food consumption reports of 2,000 households during a 14-day period:

• Coffee is still the most popular drink, poured at 40% of all a.m. meals — down from 60% in 1984 largely because it's less popular among those under 35.

• Bananas are now the No. 1 fruit, replacing grapefruit — the favorite in 1984.

• Bagels are the fastest growing food on the menu. People ate an average of 7 bagels per person in 1993, up from 2.6 in 1984.

• Fried eggs are down to 11 per person in 1993, from 23 in 1984. Bacon has also taken a beating, down to 13 servings per person in 1993 from 21 in 1984.

• Doughnuts, sweet rolls and Danish are still popular, with a consumption of 12 per person in '93, from 11 in '84.

a. Who was surveyed?
b. What information was collected from each household?
c. What conclusion do you draw from the five survey summary statements with regard to healthful foods?
d. What conclusion do you draw from the five survey summary statements with regard to convenience?

1.7 If you were an independent researcher, how would you test the accuracy of the statement made in Case Study 1-1 that "airline seats [were] built for 5-foot-9, 170-pounders"? Include who you would survey and what information you would collect.

1.2 | Introduction to Basic Terms

In order to begin our study of statistics we need to first define a few basic terms.

Population

A collection, or set, of individuals or objects or events whose properties are to be analyzed.

The population is the complete collection of individuals or objects that are of interest to the sample collector. The concept of a population is the most fundamental idea in statistics. The population of concern must be carefully defined and is considered fully defined only when its membership list of elements is specified. The set of "all students who have ever attended a U.S. college" is an example of a well-defined population.

Typically, we think of a population as a collection of people. However, in statistics the population could be a collection of animals, or manufactured objects, or whatever. For example, the set of all redwood trees in California could be a population.

Finite
Infinite

There are two kinds of populations, finite and infinite. When the membership of a population can be (or could be) physically listed, the population is said to be *finite*. When the membership is unlimited, the population is *infinite*. The books in your college library are a finite population; the OPAC (Online Public Access Catalog, the computerized card catalog) lists the exact membership. All the registered voters in the United States form a very large finite population; if necessary, a composite of all voter lists from all voting precincts across the United States could be compiled. On the other hand, the population of all people who might use aspirin and the population of all 40-watt light bulbs to be produced by Sylvania are infinite.

Sample

A subset of a population.

A sample consists of the individuals, objects, or measurements selected by the sample collector from the population.

VARIABLE

A characteristic of interest about each individual element of a population or sample.

A student's age at entrance into college, the color of the student's hair, the student's height, and the student's weight are four variables.

DATA (SINGULAR)

The value of the variable associated with one element of a population or sample. This value may be a number, a word, or a symbol.

For example, Bill Jones entered college at age "23," his hair is "brown," he is "71 inches" tall, and he weighs "183 pounds." These four pieces of data are the values for the four variables as applied to Bill Jones.

DATA (PLURAL)

The set of values collected for the variable from each of the elements belonging to the sample.

The set of 25 heights collected from 25 students is an example of a set of data.

EXPERIMENT

A planned activity whose results yield a set of data.

This includes both the activities for selecting the elements and obtaining the data values.

PARAMETER

A numerical value summarizing all the data of an entire population.