

Session 1

Statistics As Problem Solving

Key Terms for This Session

New in This Session

- bias
- census
- data
- population
- qualitative data
- qualitative variables
- quantitative data
- quantitative variables
- random error
- random sample
- representative sample
- sample
- variable
- variation

Introduction

Statistics is a problem-solving process that seeks answers to questions through data. In this session, we begin to explore the problem-solving process of statistics and to investigate how data vary. This process typically has four components: **[See Note 1]**

1. Ask a question.
2. Collect appropriate data.
3. Analyze the data.
4. Interpret the results.

Learning Objectives

In this session, you will learn the following:

- Statistics is a problem-solving process with four components.
- Data consist of measurements of a particular variable.
- There are two types of variables—quantitative and qualitative.
- There are many sources of variation in data, including random error and bias.

Note 1. Many teachers focus solely on the third component of our four-step process for statistical investigations: data analysis. But to properly understand your data, you need to do more than simply examine them. Specifically, there are four things you should consider:

- a. Formulation of a statistical question
- b. The nature of data
- c. Particular ways to examine data
- d. Types of interpretations

Note 1 cont'd. next page

Part A: A Problem-Solving Process (15 min.)

The word *statistics* may bring to mind polls and surveys, or facts and figures in a newspaper article. But statistics is more than just a bunch of numbers: Statistics is a problem-solving process that seeks answers to questions through data.

By asking and answering statistical questions, we can learn more about the world around us. Statistics is used every day to help us gain insight into questions that affect our lives: Is our population growing or shrinking? What is the safest way to invest money? Will eating more fruits and vegetables really make us live longer?

Write and Reflect

Problem A1. What do you think of when you hear the word *statistics*?

Four things make a problem statistical: the way in which you ask the question, the role and nature of the data, the particular ways in which you examine the data, and the types of interpretations you make from the investigation. A statistics problem typically contains four components:

1. Ask a question.

Asking a question gets the process started. It's important to ask a question carefully, with an understanding of the data you will use to find your answer.

2. Collect appropriate data.

Collecting data to help answer the question is an important step in the process. You obtain data by measuring something, so your measurement methods must be chosen with care. Sampling is one way to collect data; experimentation is another.

3. Analyze the data.

Data must be organized, summarized, and represented properly in order to provide good answers to statistical questions. Also, the data you collect usually *vary* (i.e., they are not all the same), and you will need to account for the sources of this variation.

4. Interpret the results.

After you analyze your data, you must interpret it in order to provide an answer—or answers—to the original question.

This four-step process for solving statistical problems is the foundation of all the activities in this course. You will become increasingly familiar with this process as you investigate different statistical problems.

Write and Reflect

Problem A2. Think of a general question that could be answered with statistics. Now think carefully about the four steps listed above. How would you carry out each step in order to answer your question? Be as specific as possible. [See Tip A2, page 25]

Note 1, cont'd.

These four elements serve as the foundation of all the activities in this course. The activities in Part B of this session begin with a question (or questions) and then focus on the nature of data. Each activity emphasizes three points:

- Data consist of measurements of a particular variable.
- There is variation in data.
- There are many potential sources of this variation.

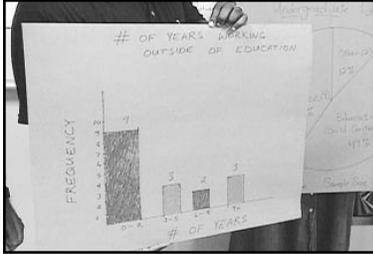
Two questions recur throughout this session: Why are there differences (i.e., variation) in our measurements? What is the source of this variation?

Parts C and D look at two kinds of “bias” in data. Part C examines how measurement bias might arise. Part D explores how bias can occur in sample selection by looking at the difference between human selection and random selection.

Materials Needed: foot-long rulers, yardsticks, tape measures, meter sticks, metric rulers

The following materials are needed for those choosing to do hands-on activities: up to 32 pennies, metric scales that are accurate to 1/100 of a gram, a stopwatch or watch with a second hand, five boxes of raisins.

Part A, cont'd.



Video Segment (approximate times: 8:11-9:42): You can find this segment on the session video approximately 8 minutes and 11 seconds after the Annenberg/CPB logo. Use the video image to locate where to begin viewing.

In this video segment, group members use the data they collected to answer their statistical question. The entire class then discusses related questions suggested by the collected data. Watch the segment after you have completed Part A.

What additional questions would you ask based on this group's data?

Part B: Data Measurement and Variation (65 min.)

Asking Questions and Collecting Data

Let's start our exploration of statistics by focusing on the first two steps of the process: "Ask a question" and "Collect appropriate data." The other steps will be explored in later sessions. We'll start with a simple statistical question. [See Note 2]

Problem B1. Let's say you'd like to find out the length of the room you're in.

Ask a question:

How long is the room?

Collect appropriate data:

Measure the length of the room in inches, using two different measurement devices: (1) a one-foot ruler and (2) a yardstick.

Measure the room length five times with each device, and fill in the table on the following page. Record your measurements to the nearest inch.

Note 2. Each activity presented here begins with a question and then considers appropriate data for answering that question. A set of measurements produced by a previous class is provided for each activity.

You may want to collect your own measurements for some of the activities, particularly if you are working with a group. This would provide group members with additional experience and understanding; it does, however, require more time.

Part B, cont'd.

Problem B1, cont'd.

<u>Measurement Instrument</u>	<u>Room Length (in Inches)</u>
Ruler	_____

Yardstick	_____

- Are the five measurements you obtained with the ruler exactly the same? Can you explain why there may be differences?
- Are the five measurements you obtained with the yardstick exactly the same? Can you explain why there may be differences?
- Did you get similar answers using the different measuring tools? Why or why not? Did you get identical answers using the different measuring tools? Why or why not?
- Which measuring tool do you think gave you more accurate results, the ruler or the yardstick? Why?
- Do you think a tape measure would be more or less accurate than a ruler or a yardstick? Why? If you have a tape measure available, use it to measure the same room five times and see how the results compare with your previous measurements.



Video Segment (approximate times: 11:46-15:15): You can find this segment on the session video approximately 11 minutes and 46 seconds after the Annenberg/CPB logo. Use the video image to locate where to begin viewing.

In this video segment, participants discuss the results of the Room-Measurement Activity. Professor Kader then introduces the concept of variation in data. Watch the segment after you have completed Problem B1 and compare the variation in your measurements with those of the onscreen participants.

Which method(s) used by the participants produced the most variation? Which method(s) produced the least variation? How do your results compare with those of the onscreen participants?

Variation, or differences in measured data, occurs for a number of reasons. Examining variation is a crucial part of data analysis and interpretation. In fact, explaining the variation in your data is as important as measuring the data itself.

Part B, cont'd.

Problem B2. Let's study two more statistical questions. For example, suppose you were curious about the relative heights and arm spans of men and women.

Ask a question:

- Are men typically taller than women?
- Do men typically have longer arm spans than women?

Collect appropriate data:

Using a meter stick, measure the heights (without shoes) and arm spans (fingertip to fingertip) of three men and three women. Record your measurements to the nearest centimeter.

- a. Did you get the same height for all six people? Did you get the same arm span for all six people? Why or why not?
- b. If you measured all six heights and arm spans again, would the results be identical? Why or why not?

Problem B3. Let's look at heights and arm spans again, this time measuring 24 people. Here are their data [heights (without shoes) and arm spans were measured to the nearest centimeter, using a meter stick]:

Sex	Height	Arm Span
Male	185	173
Female	160	161
Male	173	177
Female	170	170
Female	188	188
Male	184	196
Female	162	156
Female	170	162
Male	176	177
Female	166	165
Male	193	194
Male	178	178
Male	180	184
Female	162	159
Male	187	188
Male	186	200
Male	182	188
Female	160	157
Male	181	188
Male	192	188
Female	167	170
Female	176	173
Female	155	160
Female	162	161

- a. Examine the 24 measurements for height and arm span. You'll notice that they are not all the same. What is the source of this variation? Can you explain why there are differences?
- b. Suppose your goal was to *prove* that men are typically taller than women. Does this data prove that conclusion? Why or why not?

Part B, cont'd.

Problem B4.

Ask a question:

How much does a penny weigh?

Collect appropriate data:

We used a metric scale to weigh 32 pennies to the nearest centigram (1/100 of a gram). Here are the resulting weights:

3.08, 2.50, 2.46, 3.05, 2.45, 3.12, 3.05, 3.14, 2.48, 3.10, 3.02, 2.47, 3.10, 3.03, 3.11, 2.52,
3.00, 3.09, 3.15, 3.06, 3.18, 2.42, 2.43, 2.50, 3.07, 3.09, 3.00, 3.09, 2.47, 3.05, 2.52, 3.07

- The 32 measurements are not all the same. What is the source of this variation?
- What do you think would happen if you weighed the same penny 32 times? How would you expect that data to compare to the weights of the 32 different pennies?

Take It Further

Problem B5. Based on the data in Problem B4, how much would you expect the 33rd penny to weigh? Could you be sure of its weight before weighing it? If you can't name an exact weight, could you be confident about a *range* of weights that it falls between? Why?

How Long Is a Minute?

Problem B6.

Ask a question:

How well can people judge the time it takes for a minute to pass?

Collect appropriate data:

Collect data on this question yourself, and try this experiment with a few friends.

- Did each person estimate the same amount of time?
- If not, what do you think accounts for this variation?

Try It Online!

www.learner.org

This problem can be explored online as an Interactive Activity. Go to the *Data Analysis, Statistics, and Probability* Web site at www.learner.org/learningmath and find Session 1, Part B, Problem B6.

For a non-interactive version of this activity, use a stopwatch. Have two or more subjects engage in a conversation, and ask one of them to let you know when he or she thinks a minute has passed. Record how much time actually passed. Repeat this several times with the same or different subjects.

Part B, cont'd.

Problem B7.

Ask a question:

How many raisins are in a half-ounce box of raisins?

Collect appropriate data:

We counted the number of raisins in 17 half-ounce boxes:

29, 27, 27, 28, 31, 26, 28, 28, 30, 29, 26, 27, 29, 29, 25, 28, 28

- The 17 counts are not all the same. What do you think accounts for this variation?
- Some of the 17 counts are the same. Why do you think this is?

Problem B8.

Ask a question:

Should nuclear power be developed as an energy source?

Collect appropriate data:

Twenty-five people completed the following questionnaire:

- | | | |
|---|-----|----|
| 1. Are you concerned about future accidents at nuclear reactors? | YES | NO |
| 2. Is the growing amount of radioactive waste from nuclear reactors posing problems for the future? | YES | NO |
| 3. Do you think we ought to make a more thorough examination of safe energy supplies? | YES | NO |
| 4. Should we seek to reduce the number of nuclear power stations? | YES | NO |

Their responses are on the next page.

The raisin activity is adapted from *Investigations in Number, Data, and Space, Grade 4*. Copyright 1998 by Dale Seymour. Used with permission of Pearson Education, Inc.

Part B, cont'd.

Problem B8, cont'd.

Person #	Question 1	Question 2	Question 3	Question 4
1	No	Yes	Yes	Yes
2	No	Yes	Yes	No
3	No	Yes	No	No
4	Yes	Yes	Yes	No
5	No	Yes	Yes	No
6	Yes	Yes	Yes	Yes
7	No	Yes	Yes	No
8	Yes	Yes	Yes	Yes
9	Yes	Yes	Yes	Yes
10	Yes	Yes	Yes	Yes
11	Yes	Yes	Yes	Yes
12	No	Yes	Yes	No
13	No	Yes	Yes	No
14	Yes	Yes	Yes	No
15	Yes	Yes	Yes	Yes
16	Yes	Yes	Yes	Yes
17	Yes	Yes	Yes	Yes
18	Yes	Yes	Yes	Yes
19	Yes	Yes	Yes	Yes
20	Yes	Yes	Yes	Yes
21	Yes	Yes	Yes	No
22	No	Yes	Yes	Yes
23	Yes	Yes	Yes	Yes
24	No	Yes	Yes	Yes
25	No	Yes	Yes	No

- For Questions 1, 3, and 4, there are differences in the 25 responses to each question. What are the sources of this variation?
- For Question 2, there was no variation. Why? Would you expect the same results from another sample of 25 people?
- Take a closer look at this questionnaire. How are the questions posed, and how might that influence responses?

Part B, cont'd.

Variables

To answer the previous questions, you collected and examined data. Data is defined in terms of variables, or characteristics that may be different from one observation to the next. When we measure these characteristics, we assign a value for each variable. This set of values for a given variable is known as data.

Let's take a closer look at variables. In Problem B4, the variable is the weight of a penny, and the data are the measured weights on the 32 pennies.

Problem B9. Look back at Problems B3, B6, B7, and B8. What were the variables in each of these problems? [See Tip B9, page 25]

Some questions, such as "What is your height?," are answered with a number. Answers to questions like "What is your sex?" do not require a number.

We distinguish between variables that are measured in numbers and those that are not. This distinction becomes useful and important when we get to the analysis phase of statistical problem solving. These two types of variables are called quantitative variables and qualitative variables.

Quantitative Variables

Quantitative variables represent numbers or quantities; in fact, they are sometimes referred to as numerical variables. A test score, the number of votes cast in an election, the measured amount of soda in a two-liter bottle—these are all examples of quantitative variables.

Qualitative Variables

Rather than numbers, qualitative variables represent categories, such as "excellent" or "female," and they are sometimes referred to as categorical variables. The hometown of a college student, the favorite TV show of a politician, and the model of a car seen on a highway—these are all qualitative variables.

We refer to data in the same terms:

- Data are called quantitative if they come from measurements of a quantitative variable.
- Data are called qualitative if they come from measurements of a qualitative variable.

Part C: Bias in Measurement (20 min.)

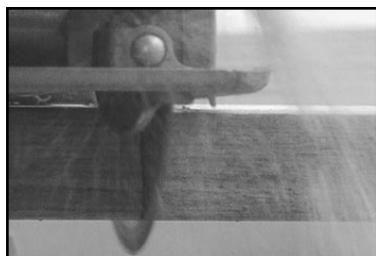
Measurement Error

Measurement is never perfect, and we can always expect measurement errors in our data. Our goal, of course, is to keep these errors to a minimum. For this reason, we need to be aware of the various sources and causes of measurement error. [See Note 3]

Random error is a nonsystematic measurement error that is beyond our control, though its effects average out over a set of measurements. For example, a scale may be properly calibrated but give inconsistent weights (sometimes too high, sometimes too low). Over repeated uses, however, the effects of these random errors average out to zero. The errors are random rather than biased: They neither understate nor overstate the actual measurement.

In contrast, measurement bias, or systematic error, favors a particular result. A measurement process is biased if it systematically overstates or understates the true value of the measurement. Consider our scale example again. If a scale is not properly calibrated, it might consistently understate weight. In this case, the measuring device—the scale—produces the bias. Human observation can also produce bias. The important thing to keep in mind is that biased measurements invariably produce unreliable results.

In any statistical investigation, we can always attribute some of the variation in data to measurement error, part of which can result from the measurement instrument itself. But human mistakes, especially recording errors (e.g., misreading a dial, incorrectly writing a number, not observing an important event, misjudging a particular behavior), can also often contribute to the variability of the measurement and thus to the results of a study.



Video Segment (approximate times: 23:31-24:47): You can find this segment on the session video approximately 23 minutes and 31 seconds after the Annenberg/CPB logo. Use the video image to locate where to begin viewing.

In this video segment, Norm Abram discusses measurement error and bias in carpentry. What examples of bias, or systematic error, were present in this discussion? How could these be avoided?

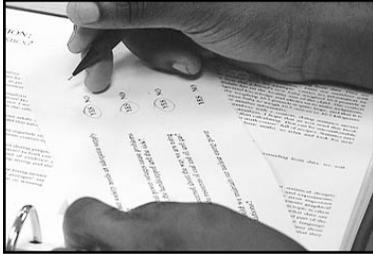
Note 3. Measurement processes may be biased due to human error. For example, error can result from our visual perceptions—the difference between what we think we see and the objective reality of what we've actually observed.

The line-segment activity is based on a visual illusion known as the Muller-Lyer illusion, a phenomenon thoroughly studied by behavioral scientists. In this illusion, two arrows with different arrowheads—one pointing out and one pointing in—are placed next to each other. Although the shafts are of equal lengths, the arrow with the outgoing head looks longer than the other arrow.

If you are working with a group, recording the responses of each group member will make the bias demonstrated in the activity more clear. There are two ways you might record results. The simplest would be to record "+" if a person overstates the length and "-" if he or she understates it. Your results should show an obvious difference between pluses and minuses. With a little more effort you could measure the deviation of perceived length and actual length (e.g., by using a tape measure on the drawn images) and then look at the average deviation as well as pluses and minuses.

In the second part of the line-segment activity, you are asked to judge the length of two lines with vertical bars at each end. Judgements in this case tend to be closer to the actual length; any errors are "random."

Part C, cont'd.



Video Segment (approximate times: 16:51-18:48): You can find this segment on the session video approximately 16 minutes and 51 seconds after the Annenberg/CPB logo. Use the video image to locate where to begin viewing.

In this video segment, Professor Kader and participants discuss the presence of bias in two surveys about nuclear power, including the one presented in Problem B8. How did the surveys' designs bias the studies' findings?

Testing Your Measurement Bias

The difference between what individuals think they see (and how they see it) and the objective reality of what they have observed is called "visual error." These errors in perception, which can significantly misrepresent reality, are a natural consequence of being human.

In this activity, you will have an opportunity to see how well you make two visual judgements. On a piece of paper, try drawing each of the graphics below, but with the length of segment BC equal to the length of AB. Then measure your segments to see how accurate your drawings are.



Try It Online!

www.learner.org

This problem can be explored online as an Interactive Activity. Go to the *Data Analysis, Statistics, and Probability* Web site at www.learner.org/learningmath and find Session 1, Part C.

Problem C1. Would you say that the errors you made in the visual judgements in the activity above were due to random error or to bias? Why or why not?

Part D: Bias in Sampling (20 min.)

Population and Sample

In data analysis, we use graphs, tables, and numerical summaries to study the variation present in our data. Often, we want to extend our interpretation to a larger group beyond the particular group studied. Such generalizations are only valid, however, if the data we examine are representative of that larger group. If not, our interpretation may misrepresent the larger group! **[See Note 4]**

The entire group that we want information about is called the population. We can gain information about this group by examining just one portion of the population, called a sample.

To gain useful information, the sample must be representative of the population. A representative sample is one in which the relevant characteristics of the sample members are generally the same as the characteristics of the population.

There are several good reasons that we use samples to study populations; chief among them are feasibility and cost. For instance, in a nationwide political survey of the population of all voters in the United States, it would be difficult, if not impossible, to poll every voter. It would also be quite expensive. Statistical theory shows that a survey of 1,000 carefully selected voters suffices to represent the opinions of the millions of people in the population of voters.

Another problem in answering questions about a population arises when we want to inspect or test products. For example, testing an air bag to see if it works properly means that we have to destroy it. We certainly can't test every air bag, but testing a carefully selected sample of air bags will tell us what we need to know about all the air bags in the population.

Problem D1. Think of a statistical question and a population. How could you determine a representative sample of that population? What would be a sample that is not representative? **[See Tip D1, page 25]**

How we select a sample is extremely important. Improper or biased sample selection can produce misleading conclusions. Sample selection is biased if it systematically favors certain outcomes. If we select only Democrats to participate in a political survey, the outcome will reflect Democrats' opinions, but not other political parties'. If we personally select a sample of students we know and like for a school survey, we have just eliminated the differing opinions of those whom we do not know and like. We need to select our sample in an unbiased fashion.

Note 4. A voter poll taken during the 1936 presidential election provides a good example of the danger of biased sampling. The magazine *Literary Digest* sent a survey to 10 million Americans to determine how they would vote in the upcoming election between Democrat Franklin Roosevelt and Republican Alf Landon. More than two million Americans responded to this poll, and 60% supported Landon. The magazine published these findings, suggesting that Landon was guaranteed to win the election.

Despite the findings of the poll, however, Roosevelt defeated Landon in one of the largest landslide presidential elections ever. What happened? The sample used in the *Literary Digest* poll—a sample collected through magazine subscription lists, lists of car owners, and telephone directories—was not representative. Not all Americans at this time owned cars, had telephones, or subscribed to magazines. Moreover, Democrats were much less likely to own a car or have a telephone, and thus were less likely to be included in the sample. As a result, the sample was not representative, and the poll did not predict the outcome of the election.

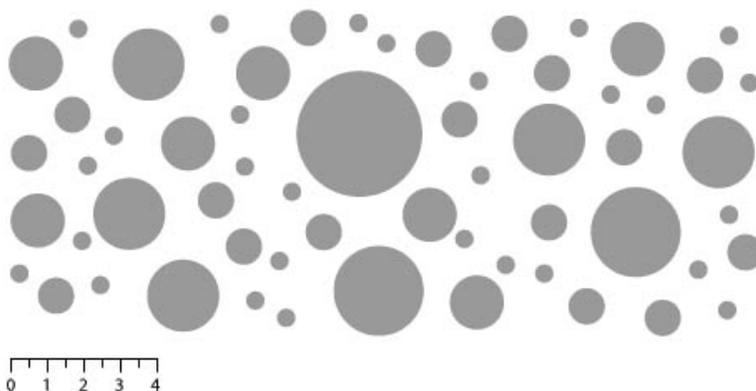
Part D, cont'd.

Random Sampling

Random sampling is a way to remove bias in sample selection. For example, to pick a random sample of 20 people out of a population of 1,000, you might put all 1,000 names in a hat, then draw 20 of them. Random sampling attempts to reduce bias in sample selection, since every member of the population has an equal chance of being selected. [See Note 5]

In this activity, you will have the opportunity to see if you can personally select a sample that is representative of a particular population.

Here are 60 circles. Can you select five circles that best represent the size of all the circles? (The average size of the five circles should equal the average size of all the circles.)



Look at the picture for no longer than 20 seconds. Mark the five circles you choose. Then use the scale on the picture to measure the diameter of those five circles. Find the average diameter of your sample.

The average diameter of all 60 circles is 1 unit. How close to that is your sample?

(Note that a computer, selecting any five of the 60 circles randomly, might generate average diameters ranging from as small as 0.5 units to as large as 2.2 units.)

Try It Online! www.learner.org

This problem can be explored online as an Interactive Activity. Go to the *Data Analysis, Statistics, and Probability* Web site at www.learner.org/learningmath and find Session 1, Part D.

Note 5. Good sampling practices rely on some form of random selection in order to remove the bias caused by human involvement in the selection process. The sampling activity in Part D is intended to demonstrate how human selection might result in biased results. You are asked to select a sample of five circles from a population of 60 circles in order to estimate the size of the circles in the entire population. You will then compare the accuracy of your sample with the accuracy of a random sample. A bias should appear: Most people tend to pick a sample that greatly overestimates the size of the circles.

Part D, cont'd.

Problem D2. Can you think of any circumstances in which it would be difficult or impossible to select a simple random sample?

You may have noticed that each of the problems you looked at in this session began with a question. Providing answers to questions like these is the goal of statistics. However, the variation present in statistical data makes it difficult to give an exact answer to the question posed.

In order to identify any patterns present in the variation, we must analyze our data by organizing and summarizing it. Once this analysis is complete, we can interpret the results to answer our questions. In later sessions, we will look at the analysis and interpretation components in more detail.

It's also important to remember that when you conduct a statistical investigation, the question you pose is designed to investigate a group ("the population"). The results of an investigation involving a sample are frequently used to draw conclusions about the entire population. If an attempt is made to include every individual from the population in a sample, then the investigation is called a census.

Problem D3. Why is a census still considered a sample?

Homework

Problem H1. We used a metric tape measure to measure the lengths of 20 pine needles to the nearest millimeter:

117, 56, 48, 69, 71, 120, 111, 49, 68, 110, 109, 64, 93, 43, 109, 37, 93, 40, 86, 47

- Think of a question that collecting this data might answer.
- Notice that the 20 pine-needle measurements are not all the same. What is the source of this variation?
- If you have access to some pine needles, measure the length (to the nearest millimeter) of 10 different needles with a metric ruler. If you do not have access to pine needles, measure the length (to the nearest millimeter) of 10 different pieces of spaghetti with a metric ruler. Are the lengths the same?

Homework, cont'd.

Problem H2.

Ask a question:

- Where were my nickels minted?
- When were my nickels minted?

Collect appropriate data:

A nickel's mint location (if present) is located on the head side of the coin below Jefferson's ponytail. We recorded the mint locations and years for 100 nickels ("S" indicates San Francisco, "D" indicates Denver, "P" indicates Philadelphia, and "N" indicates that no mint location was present):

Location	Year	Location	Year	Location	Year
P	1997	P	1986	D	1993
D	1989	N	1976	D	1983
D	1995	P	1985	P	1995
D	1986	D	1997	P	1986
P	1983	P	1984	S	1968
D	1993	P	1995	D	1995
N	1976	P	1985	N	1974
P	1989	P	1988	P	1990
D	1982	N	1978	D	1984
D	1989	D	1988	P	1997
D	1972	D	1998	P	1981
D	1996	P	1990	D	1976
P	1995	D	1982	P	1984
D	1998	P	1998	N	1974
N	1978	P	1983	D	1976
P	1991	D	1973	D	1998
D	1985	D	1986	D	1985
D	1987	D	1987	D	1971
D	1988	D	1991	D	1996
D	1983	P	1990	D	1991
D	1972	N	1973	P	1986
P	1995	D	1982	D	1995
D	1998	P	1996	P	1982
D	1995	S	1970	S	1970
P	1991	P	1996	D	1975
N	1972	P	1991	D	1982
P	1996	D	1994	D	1991
P	1990	P	1989	P	1990
P	1989	P	1997	D	1973
D	1989	P	1993	D	1974
D	1971	D	1986	D	1997
D	1980	N	1975	D	1976
P	1996	D	1990	D	1981
P	1986				

Homework, cont'd.

Problem H2, cont'd.

- There are differences in the mint locations. What is the source of this variation?
- There are differences in the mint years. What is the source of this variation?
- What observations can you make about this data? For example, which location appears most frequently? Which decade?
- When were the coins with no mint location minted? Does this suggest any new statistical questions?
- Record the mint locations and years for 10 of your own nickels. How do your data compare to the data recorded above?
- (Optional) Suppose you knew that the coins with no given mint location were in fact minted in either Philadelphia, Denver, or San Francisco. Can you pinpoint where these coins were minted? Defend your answer, using the data in this problem.

[See Tip H2, page 25]

Problem H3.

Ask a question:

What is your pulse rate?

Collect appropriate data:

Take your pulse rate, in beats per minute, seven times during the same day.

- Are the measurements of your pulse rate the same? If not, what is the source of this variation?
- Suppose you wanted to estimate your typical pulse rate. How could you answer this statistical question?

Try this experiment on yourself. You might take your pulse rate under several different conditions, for example, early in the morning, late in the day, and after exercising. Are your measurements the same?

Suggested Readings

These readings are available as downloadable PDF files on the *Data Analysis, Statistics, and Probability* Web site. Go to:

www.learner.org/learningmath

Krus, David and Webb, James (Autumn, 1997). "Demonstrating Variance Using the Muller-Lyer Illusion," *Teaching Statistics*, 19 (3), 72-76.

Principles and Standards for School Mathematics (Reston, VA: National Council of Teachers of Mathematics, 2000). Standards on Data Analysis and Probability: K-2, 108-115; 3-5, 176-181; and 6-8, 248-255.

Tips

Part A: A Problem-Solving Process

Tip A2. Here are two such general questions: “Are girls smarter than boys?” and “Which airline has the best on-time departure performance?”

Part B: Data Measurement and Variation

Tip B9. At least one of these questions has more than one variable. A variable is any characteristic that may change from one observation to the next.

Part D: Bias in Sampling

Tip D1. A population might be the students at a certain school, the members of the Republican Party, or all the soda cans shipped to the nearest convenience store this year. A representative sample must have all the same characteristics as the population.

Homework

Tip H2. View the Session 2 video to see how the onscreen participants solved this problem.

Solutions

Part B: Data Measurement and Variation

Problem B1. Answers will vary. Here is one sample data set and the solutions to problems a–e.

Measurement Instrument	Room Length (in Inches)
Ruler:	196
	193
	186
	197
	196
Yardstick:	195
	195
	193
	192
	196

- The five measurements obtained using the ruler are not all exactly the same. The most likely cause of these differences is the method of measurement, such as the way the ruler was laid out.
- The five measurements obtained using the yardstick are not all exactly the same. Again, the most likely cause of these differences is the method of measurement.
- Yes, the answers are roughly similar. They should not be significantly different from one another, since the measurements were made in a similar way each time. However, the answers from the two measurement tools are not identical, since the methods of measurement were different.
- We would expect the yardstick to be more accurate than the ruler, since, with fewer measurements to be taken and added, there is less potential for error.
- We would expect a tape measure to be even more accurate than a yardstick, since there are still fewer measurements to take.

Problem B2.

- a. You should find that the heights and arm spans are different for the six people, since people are inherently different and come in all shapes and sizes.
- b. No, they would probably not be identical. There are many potential reasons—most likely, because of measurement errors, such as recording errors.

Problem B3.

- a. The heights are not the same, nor are the arm spans. These measurements vary partly because of the differences between people. You may notice that we also have data on sex; these values vary depending on whether the person measured is male or female. The sex of an individual also has an effect on the variation in the list of heights and arm spans. Also, there is always a possibility of variation due to measurement errors. We cannot expect measurements of height or arm span with a meter stick to be exact every time. And still another type of measurement error may occur: A mistake might be made in recording the person's sex.

Solutions, cont'd.

Problem B3, cont'd.

- b. Although the data suggest that men are typically taller than women, there is not enough data to prove this conclusively.

Problem B4.

- a. Here are some possible sources of the variation in this data:
 - Measurement errors may have occurred.
 - Pennies may be more worn than newer ones and therefore weigh less.
 - Different ingredients may have been used for making pennies in different years.
 - Pennies may have been made at different mints, using different equipment.
 - A penny may have something attached to it, such as a piece of dirt or gum.
- b. You should still expect to find some very slight variation in the data as a result of measurement errors, but the values should be much closer than if you had measured 32 different pennies.

Problem B5. We might expect the 33rd penny to weigh roughly three grams. There is no way to be absolutely sure of its weight beforehand, since there is variation in the data we were given for the first 32 pennies. Judging from the first 32 pennies, it is quite likely that the 33rd will be between 2.42 and 3.18 grams, and less likely that it will be between 3.00 and 3.10 grams.

Problem B6. One possible source of the variation in this data is that some people are better at judging time than others; some may consistently overestimate or underestimate the minute. A second source, as always, is measurement error. Finally, people learn from experience—their own or someone else's. After witnessing the measurement errors in the first estimate, they are likely to adjust their second estimates accordingly.

Problem B7.

- a. Here are some possible sources of the variation in this data:
 - Measurement errors may have occurred.
 - Raisins come in many sizes and boxes are filled by weight. It takes fewer large raisins and more smaller raisins to fill a half-ounce box.
 - The machine that fills the boxes is not perfect; it may include too many or too few raisins.
 - Each box probably doesn't contain exactly one half-ounce of raisins when you account for clumping or air in the box.
- b. Since all of the values were very close, and there are few enough possible values for the number of raisins, we should expect some of them to be exactly the same. This would also be true of a large class taking a test with 20 questions; some students in this class would get the same score, just by chance.

Problem B8.

- a. Here are some possible sources of the variation in this data:
 - Differences of opinion
 - Measurement errors
 - Misread or misunderstood questions
 - Untruthful responses

Solutions, cont'd.

Problem B8, cont'd.

- b. It may be that so many people would say “Yes” to this question that finding one of 25 people to say “No” is unlikely. Possibly a person may have misunderstood the question or may not have responded truthfully to what he or she perceived as a “loaded” question. Although the data suggest that another 25 people might respond in the same way, there is not enough data to prove this conclusively.
- c. The wording on each of these questions makes one think about the negative aspects of nuclear power. A respondent might be influenced to answer “Yes” to the final question after reading the first three. This questionnaire may be considered to be biased against nuclear power.

Problem B9. Here are the variables for each problem:

- Problem B3: Sex, height, arm span
- Problem B6: Time in seconds (i.e., people’s estimates of 60 seconds)
- Problem B7: Number of raisins in a box
- Problem B8: Answers to Questions 1, 2, 3, and 4

Part C: Bias in Measurement

Problem C1. The errors are mainly due to the bias in the presentation. A random error is just as likely to be an overestimate as an underestimate, whereas bias is a systematic error that consistently overestimates or underestimates the true length. Most people who compare the lengths overestimate the length of BC, with its outward-pointing arrow, because of the optical illusion caused by the direction of the arrowheads. They are less likely to make this error when the arrows are replaced by perpendicular lines.

Part D: Bias in Sampling

Problem D1. One such question is “Are girls better math students than boys?” Consider this question for the population of a certain school. A representative sample would be a selection of girls across grade and ability levels and a selection of boys across grade and ability levels. An unrepresentative sample might select only one grade level or one ability level. Comparing the girls and boys in the most challenging math course at the school would be a very unrepresentative sample.

Problem D2. It is difficult to select a simple random sample if full information about the population is not available. It would be extremely difficult to select a simple random sample of the world’s ant population, for instance, since it would be impractical (if not impossible) to obtain enough information about the population to set up the random sample.

Problem D3. A census is still considered a sample because there is no guarantee that the attempt to include everyone has been successful. For example, every 10 years, the U.S. population census misses between 1% and 3% of the individuals in the population, and accidentally counts some people more than once. A full census for all but the smallest populations would be impossible to complete successfully.

Solutions, cont'd.

Homework

Problem H1.

- a. One question might be, "What is the average length of a pine needle from a fully grown pine tree?"
- b. Here are some possible sources of the variation in this data:
 - Pine needles are inherently different and come in different sizes.
 - Measurement errors may have occurred.
 - The needles may have come from different places on the tree, which could affect their lengths.
 - The needles may have come from different trees, whose different locations could affect the needles' lengths.
 - Some of the needles may have come from one type of pine tree and some from another.
- c. They probably will not all be identical, due to the many possible sources of variation.

Problem H2.

- a. Here are some possible sources of the variation in this data:
 - Coins are minted in different places.
 - Measurement errors may have occurred, including recording errors of the mint location.
 - These coins were obtained from different locations, some with more of a certain type of coin. For example, on the East Coast, you're more likely to see nickels with a P marking.
- b. This variation is due to the fact that the same mints produce many different coins each year, and they all remain in circulation.
- c. The location D appears most frequently; roughly half the data (51 coins out of 100) come from this location. The 1990s is the most frequently appearing decade, with 40 coins; the 1980s is a close second, with 37.
- d. The coins with no mint location were minted between 1972 and 1978. One new question might be, "Where were the coins with no mint location minted?"
- e. The data should be roughly comparable but not identical, due to variation. Some causes of this variation are the location in which the coins were obtained, the condition of the coins that are kept in circulation, and measurement error.
- f. It seems clear that the coins with no mint location were minted in Philadelphia. The range of dates for the N coins is 1972-1978, and the earliest date for a P coin is 1981. In contrast, there are many coins from Denver minted through the 1970s. The implication is that Philadelphia didn't begin to imprint its location until after 1978. (There are not enough coins with the S mint mark to suggest that the larger number of coins with no mint location came from San Francisco.)

Solutions, cont'd.

Problem H3.

- a. The measurements are very unlikely to be the same. Here are some possible sources of the variation in this data:
 - Measurement errors may have occurred.
 - The measurements are taken at different times of the day, which affects pulse rate.
 - The measurements are taken under different conditions, for example, at different levels of physical activity or stress.
- b. This is difficult, since there is a lot of variation in pulse-rate data. You might take a large number of readings, then partition them into activities (working out, sleeping, eating, etc.) to find a typical pulse rate for each activity. Any single number to estimate overall pulse rate would be subject to a lot of variation, and any range that contains all possible pulse rates for one day would be a very wide one.