

## 1.8: Homework

### 1.1 Definitions of Statistics, Probability, and Key Terms

For each of the following eight exercises, identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.

1. A fitness center is interested in the mean amount of time a client exercises in the center each week.
2. Ski resorts are interested in the mean age that children take their first ski and snowboard lessons. They need this information to plan their ski classes optimally.
3. A cardiologist is interested in the mean recovery period of their patients who have had heart attacks.
4. Insurance companies are interested in the mean health costs each year of their clients, so that they can determine the costs of health insurance.
5. A politician is interested in the proportion of voters in their district who think the politician is doing a good job.
6. A marriage counselor is interested in the proportion of clients she counsels who stay married.
7. Political pollsters may be interested in the proportion of people who will vote for a particular cause.
8. A marketing company is interested in the proportion of people who will buy a particular product.

Use the following information to answer the next three exercises: A Lake Tahoe Community College instructor is interested in the mean number of days Lake Tahoe Community College math students are absent from class during a quarter.

9. What is the population she is interested in?
- a. all Lake Tahoe Community College students
  - b. all Lake Tahoe Community College English students
  - c. all Lake Tahoe Community College students in the instructor's classes
  - d. all Lake Tahoe Community College math students

10. Consider the following:

$X$  = number of days a Lake Tahoe Community College math student is absent

In this case,  $X$  is an example of a:

- a. variable.
  - b. population.
  - c. statistic.
  - d. data.
11. The instructor's sample produces a mean number of days absent of 3.5 days. This value is an example of a:
- a. parameter.
  - b. data.
  - c. statistic.
  - d. variable.

### 1.2 Data, Sampling, and Variation in Data and Sampling

For the following exercises, identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.

12. number of tickets sold to a concert
13. percent of body fat
14. favorite baseball team
15. time in line to buy groceries
16. number of students enrolled at Evergreen Valley College
17. most-watched television show
18. brand of toothpaste
19. distance to the closest movie theatre
20. age of executives in Fortune 500 companies
21. number of competing computer spreadsheet software packages

Use the following information to answer the next two exercises: A study was done to determine the age, number of times per week, and the duration (amount of time) of resident use of a local park in San Jose. The first house in the neighborhood around the park was selected randomly and then every 8th house in the neighborhood around the park was interviewed.

22. “Number of times per week” is what type of data?

- a. qualitative (categorical)
- b. quantitative discrete
- c. quantitative continuous

23. “Duration (amount of time)” is what type of data?

- a. qualitative (categorical)
- b. quantitative discrete
- c. quantitative continuous

24. Airline companies are interested in the consistency of the number of babies on each flight, so that they have adequate safety equipment. Suppose an airline conducts a survey. Over Thanksgiving weekend, it surveys six flights from Boston to Salt Lake City to determine the number of babies on the flights. It determines the amount of safety equipment needed by the result of that study.

- a. Using complete sentences, list three things wrong with the way the survey was conducted.
- b. Using complete sentences, list three ways that you would improve the survey if it were to be repeated.

25. Suppose you want to determine the mean number of students per statistics class in your state. Describe a possible sampling method in three to five complete sentences. Make the description detailed.

26. Suppose you want to determine the mean number of cans of soda drunk each month by students in their twenties at your school. Describe a possible sampling method in three to five complete sentences. Make the description detailed.

27. List some practical difficulties involved in getting accurate results from a telephone survey.

28. List some practical difficulties involved in getting accurate results from a mailed survey.

29. With your classmates, brainstorm some ways you could overcome these problems if you needed to conduct a phone or mail survey.

30. The instructor takes a sample by gathering data on five randomly selected students from each Lake Tahoe Community College math class. The type of sampling used is

- a. cluster sampling
- b. stratified sampling
- c. simple random sampling
- d. convenience sampling

31. A study was done to determine the age, number of times per week, and the duration (amount of time) of residents using a local park in San Jose. The first house in the neighborhood around the park was selected randomly and then every eighth house in the neighborhood around the park was interviewed. The sampling method was:

- a. simple random
- b. systematic
- c. stratified
- d. cluster

32. Name the sampling method used in each of the following situations:

- a. A person in the airport is handing out questionnaires to travelers asking them to evaluate the airport’s service. The person does not ask travelers who are hurrying through the airport with their hands full of luggage, but instead asks all travelers who are sitting near gates and not taking naps while they wait.
- b. A teacher wants to know if students are doing homework so they randomly select rows two and five and then call on all students in row two and all students in row five to present the solutions to homework problems to the class.
- c. The marketing manager for an electronics chain store wants information about the ages of its customers. Over the next two weeks, at each store location, 100 randomly selected customers are given questionnaires to fill out asking for information about age, as well as about other variables of interest.
- d. The librarian at a public library wants to determine what proportion of the library users are children. The librarian has a tally sheet on which they mark whether books are checked out by an adult or a child. The librarian records this data for every fourth patron who checks out books.
- e. A political party wants to know the reaction of voters to a debate between the candidates. The day after the debate, the party’s polling staff calls 1,200 randomly selected phone numbers. If a registered voter answers the phone or is available to come to the

phone, that registered voter is asked whom they intend to vote for and whether the debate changed their opinion of the candidates.

33. A “random survey” was conducted of 3,274 people of the “microprocessor generation” (people born since 1971, the year the microprocessor was invented). It was reported that 48% of those individuals surveyed stated that if they had \$2,000 to spend, they would use it for computer equipment. Also, 66% of those surveyed considered themselves relatively savvy computer users.

- Do you consider the sample size large enough for a study of this type? Why or why not?
- Based on your “gut feeling,” do you believe the percents accurately reflect the U.S. population for those individuals born since 1971? If not, do you think the percents of the population are actually higher or lower than the sample statistics? Why?  
Additional information: The survey, reported by Intel Corporation, was filled out by individuals who visited the Los Angeles Convention Center to see the Smithsonian Institute’s road show called “America’s Smithsonian.”
- With this additional information, do you feel that all demographic and ethnic groups were equally represented at the event? Why or why not?
- With the additional information, comment on how accurately you think the sample statistics reflect the population parameters.

34. The Well-Being Index is a survey that follows trends of U.S. residents on a regular basis. There are six areas of health and wellness covered in the survey: Life Evaluation, Emotional Health, Physical Health, Healthy Behavior, Work Environment, and Basic Access. Some of the questions used to measure the Index are listed below.

Identify the type of data obtained from each question used in this survey: qualitative(categorical), quantitative discrete, or quantitative continuous.

- Do you have any health problems that prevent you from doing any of the things people your age can normally do?
- During the past 30 days, for about how many days did poor health keep you from doing your usual activities?
- In the last seven days, on how many days did you exercise for 30 minutes or more?
- Do you have health insurance coverage?

35. In advance of the 1936 Presidential Election, a magazine titled Literary Digest released the results of an opinion poll predicting that the republican candidate Alf Landon would win by a large margin. The magazine sent post cards to approximately 10,000,000 prospective voters. These prospective voters were selected from the subscription list of the magazine, from automobile registration lists, from phone lists, and from club membership lists. Approximately 2,300,000 people returned the postcards.

- Think about the state of the United States in 1936. Explain why a sample chosen from magazine subscription lists, automobile registration lists, phone books, and club membership lists was not representative of the population of the United States at that time.
- What effect does the low response rate have on the reliability of the sample?
- Are these problems examples of sampling error or nonsampling error?
- During the same year, George Gallup conducted his own poll of 30,000 prospective voters. These researchers used a method they called “quota sampling” to obtain survey answers from specific subsets of the population. Quota sampling is an example of which sampling method described in this module?

36. Crime-related and demographic statistics for 47 US states in 1960 were collected from government agencies, including the FBI’s *Uniform Crime Report*. One analysis of this data found a strong connection between education and crime indicating that higher levels of education in a community correspond to higher crime rates.

Which of the potential problems with samples discussed in Data, Sampling, and Variation in Data and Sampling could explain this connection?

37. Imagine you work for a polling company and a member of your team has proposed the following survey question:

“Do you feel happy paying your taxes while some politicians are allowed to use loopholes and avoid paying their fair share of taxes?”

As part of preliminary data collection, 11 people responded to this question. Each participant answered “NO!”

Which of the potential problems with samples discussed in this module could explain this connection?

38. A scholarly article about response rates begins with the following quote:

“Declining contact and cooperation rates in random digit dial (RDD) national telephone surveys raise serious concerns about the validity of estimates drawn from such research.” (Scott Keeter et al., “Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey,” *Public Opinion Quarterly* 70 no. 5 (2006), <http://poq.oxfordjournals.org/content/70/5/759.full> (accessed May 1, 2013).)

The Pew Research Center for People and the Press admits:

“The percentage of people we interview – out of all we try to interview – has been declining over the past decade or more.” (Frequently Asked Questions, Pew Research Center for the People & the Press, <http://www.people-press.org/methodol...wer-your-polls> (accessed May 1, 2013).)

- What are some reasons for the decline in response rate over the past decade?
- Explain why researchers are concerned with the impact of the declining response rate on public opinion polls.

### 1.3 Levels of Measurement

39. Fifty part-time students were asked how many courses they were taking this term. The (incomplete) results are shown below:

| # of courses | Frequency | Relative frequency | Cumulative relative frequency |
|--------------|-----------|--------------------|-------------------------------|
| 1            | 30        | 0.6                |                               |
| 2            | 15        |                    |                               |
| 3            |           |                    |                               |

**Table 1.15** Part-time Student Course Loads

- Fill in the blanks in Table 1.15.
- What percent of students take exactly two courses?
- What percent of students take one or two courses?

40. Sixty adults with gum disease were asked the number of times per week they used to floss before their diagnosis. The (incomplete) results are shown in Table 1.16.

| # flossing per week | Frequency | Relative frequency | Cumulative relative frequency |
|---------------------|-----------|--------------------|-------------------------------|
| 0                   | 27        | 0.4500             |                               |
| 1                   | 18        |                    |                               |
| 3                   |           |                    | 0.9333                        |
| 6                   | 3         | 0.0500             |                               |
| 7                   | 1         | 0.0167             |                               |

**Table 1.16** Flossing Frequency for Adults with Gum Disease

- Fill in the blanks in Table 1.16.
- What percent of adults flossed six times per week?
- What percent flossed at most three times per week?

41. Nineteen immigrants to the U.S were asked how many years, to the nearest year, they have lived in the U.S. The data are as follows: 2; 5; 7; 2; 2; 10; 20; 15; 0; 7; 0; 20; 5; 12; 15; 12; 4; 5; 10 .

Table 1.17 was produced.

| Data | Frequency | Relative frequency | Cumulative relative frequency |
|------|-----------|--------------------|-------------------------------|
| 0    | 2         | 219219             | 0.1053                        |
| 2    | 3         | 319319             | 0.2632                        |
| 4    | 1         | 119119             | 0.3158                        |
| 5    | 3         | 319319             | 0.4737                        |
| 7    | 2         | 219219             | 0.5789                        |
| 10   | 2         | 219219             | 0.6842                        |

| Data | Frequency | Relative frequency | Cumulative relative frequency |
|------|-----------|--------------------|-------------------------------|
| 12   | 2         | 219219             | 0.7895                        |
| 15   | 1         | 119119             | 0.8421                        |
| 20   | 1         | 119119             | 1.0000                        |

**Table 1.17** Frequency of Immigrant Survey Responses

- Fix the errors in Table 1.17. Also, explain how someone might have arrived at the incorrect number(s).
- Explain what is wrong with this statement: “47 percent of the people surveyed have lived in the U.S. for 5 years.”
- Fix the statement in **b** to make it correct.
- What fraction of the people surveyed have lived in the U.S. five or seven years?
- What fraction of the people surveyed have lived in the U.S. at most 12 years?
- What fraction of the people surveyed have lived in the U.S. fewer than 12 years?
- What fraction of the people surveyed have lived in the U.S. from five to 20 years, inclusive?

**42.** How much time does it take to travel to work? Table 1.18 shows the mean commute time by state for workers at least 16 years old who are not working at home. Find the mean travel time, and round off the answer properly.

|      |      |      |      |      |      |      |      |      |      |
|------|------|------|------|------|------|------|------|------|------|
| 24.0 | 24.3 | 25.9 | 18.9 | 27.5 | 17.9 | 21.8 | 20.9 | 16.7 | 27.3 |
| 18.2 | 24.7 | 20.0 | 22.6 | 23.9 | 18.0 | 31.4 | 22.3 | 24.0 | 25.5 |
| 24.7 | 24.6 | 28.1 | 24.9 | 22.6 | 23.6 | 23.4 | 25.7 | 24.8 | 25.5 |
| 21.2 | 25.7 | 23.1 | 23.0 | 23.9 | 26.0 | 16.3 | 23.1 | 21.4 | 21.5 |
| 27.0 | 27.0 | 18.6 | 31.7 | 23.3 | 30.1 | 22.9 | 23.3 | 21.7 | 18.6 |

**Table 1.18**

**43.** A leading business magazine publishes data on small businesses (defined as businesses that have been publicly traded for at least a year, have a stock price of at least \$5 per share, and have reported annual revenue between \$5 million and \$1 billion). Table 1.19 shows the ages of the chief executive officers for the first 60 ranked small businesses.

| Age   | Frequency | Relative frequency | Cumulative relative frequency |
|-------|-----------|--------------------|-------------------------------|
| 40–44 | 3         |                    |                               |
| 45–49 | 11        |                    |                               |
| 50–54 | 13        |                    |                               |
| 55–59 | 16        |                    |                               |
| 60–64 | 10        |                    |                               |
| 65–69 | 6         |                    |                               |
| 70–74 | 1         |                    |                               |

**Table 1.19**

- What is the frequency for CEO ages between 54 and 65?
- What percentage of CEOs are 65 years or older?
- What is the relative frequency of ages under 50?
- What is the cumulative relative frequency for CEOs younger than 55?
- Which graph shows the relative frequency and which shows the cumulative relative frequency?

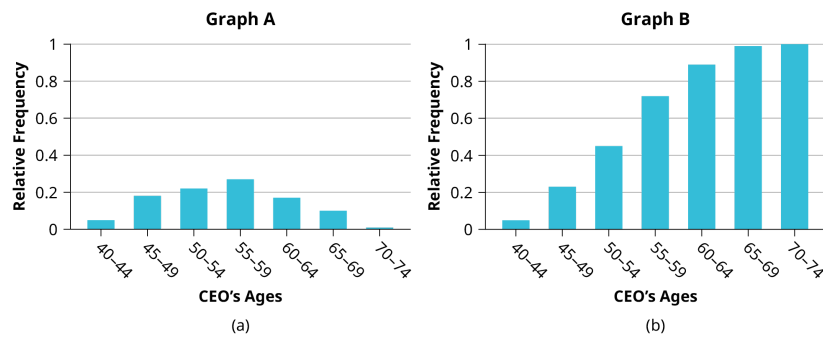


Figure 1.11

Use the following information to answer the next two exercises: Table 1.20 contains data on hurricanes that have made direct hits on the U.S. Between 1851 and 2004. A hurricane is given a strength category rating based on the minimum wind speed generated by the storm.

| Category | Number of direct hits | Relative frequency | Cumulative frequency |
|----------|-----------------------|--------------------|----------------------|
| 1        | 109                   | 0.3993             | 0.3993               |
| 2        | 72                    | 0.2637             | 0.6630               |
| 3        | 71                    | 0.2601             |                      |
| 4        | 18                    |                    | 0.9890               |
| 5        | 3                     | 0.0110             | 1.0000               |
|          | Total = 273           |                    |                      |

**Table 1.20** Frequency of Hurricane Direct Hits

44. What is the relative frequency of direct hits that were category 4 hurricanes?

- a. 0.0768
- b. 0.0659
- c. 0.2601
- d. Not enough information to calculate

45. What is the relative frequency of direct hits that were AT MOST a category 3 storm?

- a. 0.3480
- b. 0.9231
- c. 0.2601
- d. 0.3370

1.8: Homework is shared under a [not declared](#) license and was authored, remixed, and/or curated by LibreTexts.