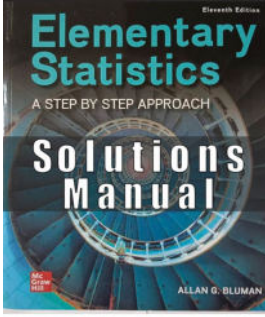

Chapter 1 - The Nature of Probability and Statistics

EXERCISE SET 1-1

1. Statistics is the science of conducting studies to collect, organize, summarize, analyze, and draw conclusions from data.

2. A variable is a characteristic or attribute that can assume different values.

3. In a census, the researchers collect data from all subjects in the population.

4. A population consists of all subjects under study while a sample is a subgroup of the population.

5. Descriptive statistics consists of the collection, organization, summarization, and presentation of data while inferential statistics consists of generalizing from samples to populations, performing estimations and hypothesis testing, determining relationships among variables, and making predictions.

6. Probability is used in gambling, insurance, and weather forecasting.

7. Samples are used more than populations both because populations are usually large and because researchers are unable to use every subject in the population.

8. A biased sample is a sample that is not representative of the population.

9. This is inferential because a generalization is being made about the population.

10. This is inferential since a prediction is being made about the future.

11. This is a descriptive statistic since it describes the weight loss for a specific group of subjects, i.e. , the teenagers at Boston University.

12. This is a descriptive statistic since it is based on the results of a survey of 2739 individuals. However, if an inference were made to all pet owners, it would be an inferential statistic.

13. This is an inferential statistic since a generalization has been made about the population.

14. This is an inferential statistic since a generalization was made about the population of all kindergartens in the state of Oregon.

15. This is an inferential statistic since a generalization was made about the population.

16. This is a descriptive statistic since it describes the results of a specific survey of 1,507 subjects.

17. This is an inferential statistic since it is a generalization made from data obtained from a sample.

18. Answers will vary.

19. Answers will vary.

EXERCISE SET 1-2

1. Qualitative variables are variables that can be placed in distinct categories according to some characteristic or attribute and cannot be ranked; while quantitative variables are numerical in nature and can be ordered or counted.

2. Discrete variables assume values that can be counted while continuous variables are variables obtained by being measured. Theoretically, continuous variables can assume all values between any given two values.

3. Continuous variables need to be rounded because of the limits of the measuring device.

4. There are four types of measurement levels used in statistics: nominal, ordinal, interval, and ratio levels of measurement.

5. Qualitative 6. Qualitative

7. Quantitative 8. Quantitative

9. Quantitative 10. Qualitative

11. Discrete 12. Continuous

13. Continuous 14. Discrete

15. Discrete 16. Continuous

17. 23.5-24.5 feet

18. 6.25-6.35 millimeters

19. 142.5-143.5 miles

20. 19.625-19.635 tons

21. 200.65-200.75 miles

22. 18.5-19.5 quarts

23. Nominal 24. Interval

25. Ratio 26. Ratio

27. Ordinal 28. Nominal

29. Ratio 30. Ratio

EXERCISE SET 1-3

1. Data can be collected by using telephone surveys, mail questionnaire surveys, personal interview surveys, by taking a look at records, or by direct observation methods.

2. Sampling error is the difference between a sample measure and a population measure. Nonsampling error is the result of collecting data incorrectly or selecting a biased sample.

3. Random numbers are used in sampling so that every subject in the population has an equal chance of being selected for a sample. Random numbers can be generated by computers or calculators; however, there are other ways of generating random numbers such as using a random number table or rolling dice.

4. The four basic sampling methods are random, systematic, stratified, and cluster methods.

5. The population could be all people in the United States who earn over $200,000 per annum. A sample could have been created by selecting 500 people randomly from an accounting firm that prepares income taxes. Answers will vary.

6. The population could be all prisons in the free world. A sample of a specific number could be selected by randomly choosing prisons in the United States and prisons in other countries. Answers will vary.

7. The population could be all households in the United States. A sample could be selected using 1000 households in the United States. Answers will vary.

8. The population could be all adults in the United States. A sample could be obtained by checking the blood levels of calcium. Of course, the researcher must be aware of the privacy rights of the respondents. Answers will vary.

9. The population could be all adults in the United States who develop diabetes. The sample could be surveying patient records of these people to see if they have been taking statins. Again, the privacy rights must be considered. Answers will vary.

10. The population could consist of all daily temperatures in Boston since records were kept, and a sample could be obtained by looking at weather report statistics.

11. Systematic 12. Stratified

13. Random 14. Cluster

15. Cluster 16. Random

EXERCISE SET 1-4

1. In an observational study, the researcher observes what is happening and tries to draw conclusions based on the observations. In an experimental study, the researcher manipulates one of the variables and tries to determine how this influences the variables.

2. The three types of observational studies are
a. Cross-sectional studies in which the data are collected at one point in time
b. Retrospective studies in which the data are collected from past records
c. Longitudinal studies in which the data are collected over a period of time

3. One advantage of an observational study is that it can occur in a natural setting. In addition, researchers can look at past instances of statistics and draw conclusions from these situations. Another advantage is that the researcher can use variables, such as drugs, that he or she cannot manipulate.
One disadvantage is that since the variable cannot be manipulated, a definite cause-and-effect situation cannot be shown. Another disadvantage is that these studies can be expensive and time-consuming. These studies can also be influenced by confounding variables. Finally, in these studies, the researcher sometimes needs to rely on data collected by others.

4. One advantage of an experimental study is that the researcher can decide how to select the subjects and assign them to the specific groups. The researcher can also control the independent variable. One disadvantage of an experimental study is that it may not occur in a natural setting. Another one is called the Hawthorne effect, and which is when the subject may change their natural behavior because they realize that they are participating in a research study.

5. In an experimental study, the researcher has control of the assignment of subjects to the groups whereas in a quasi-experimental study, the researcher uses intact groups.

6. An independent variable is the one being manipulated by the researcher whereas the dependent variable is the one being studied to identify the effects of the independent variable.

7. In research studies, a treatment group subject receives a specific treatment while those in the control group do not receive a treatment or are given a placebo.

8. The Hawthorne effect was the effect discovered when researchers realized that the workers at the Hawthorne plant knew that they were participating in a research study and this fact actually changed workers' behavior.

9. A confounding variable is one that can influence the results of the research study when no precautions were taken to eliminate it from the study.

10. Sometimes subjects respond favorably when given a placebo rather than the real treatment.

11. Blinding is used to help eliminate the placebo effect. Here the subjects are given a sugar pill that looks like the real medical pill. The subjects do not know which pill they are getting. When double blinding occurs, neither the subjects nor the researchers are told who gets the real treatment or the placebo.

12. Randomization is confounding variables since randomly assigning subjects to groups tends to "balance out" the inconsistencies
(such as age, social class, etc.) that each of the subjects brings to the study.

13. In a completely randomized design, the subjects are assigned to the groups randomly, whereas in a matched-pair design, subjects are matched on some variable. Then one subject is randomly assigned to one group, and the other subject is assigned to the other group. In both types of studies, the treatments can be randomly assigned to the groups.

14. Replication is done to determine if results apply in different settings. For example, if the experiment was done with college students, the researchers might replicate the experiment with adults or in another part of the country.

15. Observational

16. Observational

17. Experimental

18. Observational

19. Independent variable - minutes exercising Dependent variable - catching a cold

20. Independent variable - hugs and hand holding
Dependent variable - heart rate and blood pressure

21. Independent variable - happy face on the check
Dependent variable - amount of the tip

22. Independent variable - marital status
Dependent variable - cause of death

23. Age, income, socioeconomic status. Answers will vary.

24. Age, intelligence, family ties. Answers will vary.

25. Income, number of hours worked, type of boss. Answers will vary.

26. Climate, diet, medicine. Answers will vary.

27. How is a perfect body defined statistically?

28. What is meant by the "whole truth"?

29. How can 24 hours of pain relief be measured?

30. How can reading a book increase a person's IQ?

31. How much weight, if any, will be lost?

32. The steps in conducting a statistical study are as follows:
Step 1. Formulate the purpose of the study.
Step 2. Identify the variables for the study.
Step 3. Define the population.
Step 4. Decide what sampling method to use to collect the data.
Step 5. Collect the data.
Step 6. Summarize the data and perform any statistical calculations needed.
Step 7. Interpret the results.

33. Only 20 people were used in the study.

34. The only time claims can be proved is when the entire population is used.

35. It is meaningless since there is no definition of "the road less traveled." Also, there is no way to know that for every 100 women, 91 would say that they have taken "the road less traveled."

36. Since the results are not typical, the advertisers selected only a few people for whom the weight loss product worked extremely well.

37. There is no mention of how this conclusion was obtained.

38. "74% more calories" than what? No comparison group is stated.

39. Since the word may is used, there is no guarantee that the product will help fight cancer and heart disease.

40. What is meant by "24 hours of acid control"? Also can provide implies that it might or might not provide control.

41. No. There are many other factors that contribute to criminal behavior.

42. Possible answer: It could be the amount of caffeine in the coffee or tea. It could have been the brewing method.

43. Answers will vary.

44. Answers will vary.

45. Answers will vary.

46. Answers will vary.

REVIEW EXERCISES - CHAPTER 1

1. Inferential 2. Descriptive

3. Descriptive 4. Descriptive

5. Inferential 6. Inferential

7. Descriptive 8. Inferential

9. Ratio 10. Ordinal

11. Interval 12. Ratio

13. Ratio 14. Ratio

15. Ordinal 16. Ratio