

**Mind on Statistics****Chapter 5****Section 5.1**

**Questions 1 to 4:** Decide if the sample is representative (or not) of the population for the question of interest.

1. Question: What proportion of people intends to vote in the next presidential election?  
Sample: 100 baseball fans at a baseball game.  
Population: All voters in the next presidential election.  
A. Representative  
B. Not representative

KEY: B

2. Question: What is the average age of people in Salt Lake City?  
Sample: 10 people picked randomly from all people living in Salt Lake City.  
Population: All people living in Salt Lake City.  
A. Representative  
B. Not representative

KEY: A

3. Question: What is the average number of letters on a page in Webster's unabridged dictionary?  
Sample: The first page for each of the 26 letters in the alphabet.  
Population: All pages in the dictionary.  
A. Representative  
B. Not representative

KEY: B

4. Question: What is the average daily number of hours of sleep over a one year period?  
Sample: Hours of sleep for 100 randomly selected students the night before an exam.  
Population: All students in the university.  
A. Representative  
B. Not representative

KEY: B

5. A national polling organization wishes to estimate the percentage of all teenagers who believe social security will 'be there' for them. The organization surveys a random sample of 1500 teenagers and 37% of this sample says that they believe social security will 'be there' for them. In this survey, what is the population of interest?  
A. The 1500 teenagers who were surveyed.  
B. Teenagers who believe social security will 'be there' for them.  
C. All teenagers.  
D. The people in the sample who believe social security will 'be there' for them.

KEY: C

Chapter 5

6. For which of the following situations would a census be more useful than a sample survey?
- A. To determine what percent of the adults in a state are unemployed.
  - B. To estimate how many of the 200 physicians in a large city accept patients on welfare.
  - C. To estimate how many homeless people there are in a city.
  - D. To estimate how many of the students at a large university are in favor of a proposed fee increase to fund more parking.

KEY: C

7. A census differs from a sample survey because
- A. a sample survey is only done when measuring the units will destroy them.
  - B. a census is only done when measuring the units will destroy them.
  - C. in a sample survey the whole population is measured.
  - D. in a census the whole population is measured.

KEY: D

8. Which of the following is not an advantage of a sample survey over a census?
- A. A sample survey can be used when a census isn't possible.
  - B. It is much faster to collect a sample than a census.
  - C. A sample survey allows the researcher to devote the resources to getting the most accurate information possible.
  - D. Survey results have a good chance to be biased.

KEY: D

9. For a survey of American diets a random sample of 1000 people were contacted. Of the 1000 people, 340 people completed the questionnaire. The results of this study, if applied to all Americans, are questionable because of
- A. a large margin of error.
  - B. selection bias.
  - C. nonresponse bias.
  - D. response bias.

KEY: C

10. A survey-taker randomly selected 1000 students who were studying in the library and found that 90% of these students were in favor of longer library hours. The results of this study, if applied to all students in the university, are questionable because of
- A. lack of accuracy.
  - B. selection bias.
  - C. nonresponse bias.
  - D. response bias.

KEY: B

11. A survey based on a random sample of 1000 students found that 1% of these students, when asked by the professor doing the survey, admitted to having cheated on at least one exam. The results of this study, if applied to all students in the university, are questionable because of
- A. lack of accuracy.
  - B. selection bias.
  - C. nonresponse bias.
  - D. response bias.

KEY: D

12. When the method for selecting subjects produces a sample that does not represent the population of interest, the problem is called
- A. lack of accuracy.
  - B. selection bias.
  - C. nonresponse bias.
  - D. response bias.

KEY: B

13. When a representative sample is selected but only a small proportion are actually contacted, the problem is called
- A. lack of accuracy.
  - B. selection bias.
  - C. nonresponse bias.
  - D. response bias.

KEY: C

14. When a representative sample is selected but respondents give answers that are different from their true opinions, the problem is called
- A. lack of accuracy.
  - B. selection bias.
  - C. nonresponse bias.
  - D. response bias.

KEY: D

15. A weight-loss company is trying to promote its product. Based on a sample of 10 of their clients they model the relationship between starting weight and weight after 4 weeks using their product. It turns out that these 10 clients were selected because they had lost a decent amount of weight in 4 weeks. Clients who had lost hardly any weight or even gained weight were not included for this study. The results of this study, if applied to all clients, are questionable because of
- A. lack of accuracy.
  - B. selection bias.
  - C. nonresponse bias.
  - D. response bias.

KEY: B

16. In the College of Engineering exams aren't proctored. Students are required to write out the University honor code. They sign their name to pledge that they didn't cheat. But the engineering professors cannot *observe* if students cheat or not. One way to collect data on this topic would be to use an "exit poll": Have a Graduate Student Instructor ask the students if they cheated on the exam after the student has left the exam room. The results of this study, if applied to all students in the College of Engineering, are questionable because of
- A. lack of accuracy.
  - B. selection bias.
  - C. nonresponse bias.
  - D. response bias.

KEY: D

17. People claim that only 5% of the students who take the introductory statistics course will continue their education in statistics. The Statistics Department believes that this value is too low and decides to take a sample of 20 students out of the 315 students registered that semester. Of the 20 students, 7 (or 35%) say they will continue to study more statistics. It turns out that the 20 students selected for this study were those that received the highest scores on the first quiz. The results of this study, if applied to all students who take the introductory statistics course, are questionable because of
- A. lack of accuracy.
  - B. selection bias.
  - C. nonresponse bias.
  - D. response bias.

KEY: B

18. A researcher is conducting a survey among students to determine their mean age. Data is collected by asking the age of a simple random sample of 150 students. The researcher simply asked the ages of these students. The results of this study, if applied to all students in the university, are questionable because of
- A. lack of accuracy.
  - B. selection bias.
  - C. nonresponse bias.
  - D. response bias.

KEY: D

19. Define the purpose of descriptive statistics.

KEY: Descriptive statistics are graphical or numeric methods whose purpose is to provide summaries of data.

20. Define the purpose of inferential statistics.

KEY: Inferential statistics uses sample information to make conclusions about a broader range of individuals than just those who are observed.

21. Explain what the difference between descriptive statistics and inferential statistics.

KEY: The difference between the two methods is that descriptive statistics is limited to only the data being summarized, whereas inferential statistics applies to a larger population than those observed

22. Explain the difference between a survey and a census.

KEY: A survey is based on results from a (representative) sample, or subset, of the population, while a census is based on results from the entire population.

**Questions 23 to 25:** The next three questions are based on a survey of 500 students living in campus dormitories. Questionnaires were mailed to each of the 500 students. The questionnaire asked if they had ever stolen a bicycle or not. The sample proportion who said they had stolen a bicycle was used to estimate the proportion of all students on campus who had ever stolen a bicycle.

23. Describe why selection bias is a problem for this survey.

KEY: Students living in dormitories are typically freshmen and sophomores, and would not be representative of the population of all students on campus: a sample of dormitory students would exclude juniors and seniors.

24. Describe why response bias could be a problem for this survey.

KEY: Bicycle theft, being a sensitive question, could result in students who had stolen a bike giving a false answer.

25. Describe why non-response bias could be a problem for this survey.

KEY: The questionnaire was mailed, and students could choose to not respond to the questionnaire.

26. A survey is planned to be done by contacting a random sample of people who have bought something over the Internet; the question would ask whether sales made over the Internet should be taxed or not. The purpose is to estimate the proportion of all Americans who would support an Internet sales tax. Explain why selection bias would be a problem for this survey.

KEY: A sample of people would have bought something over the Internet is not representative of the population of all Americans, at least for this question.

27. A random sample of registered 500 voters was taken, and 80% responded that they had voted in the last presidential election. Explain why response bias may be a problem for this survey.

KEY: The people who did not vote may be reluctant to admit that they didn't vote, resulting in a response bias.

28. A telephone survey was planned to ask people whether they planned to watch the next Super Bowl. Explain why non-response bias may be a problem for this survey.

KEY: Many people may refuse to talk to a telephone interviewer. If the people who don't answer the phone differ from people who do in whether they plan to watch the next Super Bowl, non-response bias may result. Also, women are more likely to answer the phone and they may be less likely to watch the Super Bowl.

29. Studies have shown that growing numbers of young girls are beginning to use steroids. At a committee hearing, a survey by the Centers for Disease Control and Prevention of high school students was cited, in which 7.3 percent of ninth-grade girls said they had used illegal steroids. One critic of the results objected to the methodology of the survey, which used an anonymous questionnaire and did not ask what substances were used. He suggested that some of the teenagers might have thought incorrectly that were using anabolic steroids. What type of bias is being raised by this issue?

KEY: The type of bias being raised is response bias; the participants may have provided incorrect information because of what was asked and how questions were worded.

## Section 5.2

30. Which one of the following statements is true about sample size and margin of error?
- A. Increasing the sample size of a survey decreases the margin of error.
  - B. Increasing the sample size of a survey increases the margin of error.
  - C. Increasing the sample size of a survey does not change the margin of error.
  - D. Increasing the sample size of a survey only changes the margin of error if there is an error (a mistake) in the way the survey is conducted.
31. A randomly selected sample of 1,000 college students was asked whether they had ever used the drug Ecstasy. Sixteen percent (16% or 0.16) of the 1,000 students surveyed said they had. Which one of the following statements about the number, 0.16 is correct?
- A. It is a sample proportion.
  - B. It is a population proportion.
  - C. It is a margin of error.
  - D. It is a randomly chosen number.

KEY: A

32. A random sample of Americans at least 65 years old showed that 32% believe their health is extremely important. If the sample size was 100, approximately what was the margin of error?
- A. 1%
  - B. 5%
  - C. 10%
  - D. 15%

KEY: C

33. A random sample of 1001 Americans aged 18-24 years showed that 51% believe their health is extremely important. An approximate 95% confidence interval for the proportion of all 18-24 year olds who believe that their health is extremely important is
- A. between 41% and 61%.
  - B. between 46% and 56%.
  - C. between 48% and 54%.
  - D. between 50% and 52%.

KEY: C

34. A random sample of Americans showed that 4% believe their health care is excellent. If the conservative margin of error was 12.5%, what was the sample size?
- A. 25
  - B. 36
  - C. 49
  - D. 64

KEY: D

35. A Gallup poll reported that 88% of Americans say they have health insurance. If the sample size was 2500, an approximate 95% confidence interval for the proportion of all Americans who say they are insured is
- A. between 80% and 96%.
  - B. between 84% and 92%.
  - C. between 86% and 90%.
  - D. between 87.3% and 88.7%.

KEY: C

36. A Gallup poll reported that 58% of Americans say they are satisfied with the costs of healthcare. If the margin of error was 5%, approximately what was the sample size?
- A. 10
  - B. 200
  - C. 400
  - D. 1600

KEY: C

**Questions 37 and 38:** According to a Gallup poll, about 90% of all American adults owned a cell phone at the time of the poll. The results are based on telephone interviews with a randomly selected national sample of 998 adults, 18 years and older. The margin of error was reported to be 3.5%.

37. What was the population of interest in this Gallup Poll?
- A. All American adults who own cell phones.
  - B. All American adults.
  - C. The 998 adults who participated in the survey.
  - D. The participants in the survey who owned cell phones.

KEY: B

38. Which of the following statements correctly interprets the reported margin of error of 3.5%?
- A. In about 95% of all random samples of this size from the same population, the difference between the sample percent and the population percent will be less than 3.5%.
  - B. In about 3.5% of all random samples of this size from the same population, the sample percent will equal the population percent.
  - C. The probability that a 95% confidence interval based on this poll does not cover the population proportion is 3.5%.
  - D. In about 95% of all random samples of this size from the same population, the difference between the sample percent and the population percent will be more than 3.5%.

KEY: A

39. A recent poll reported that 62% of all college students believe there is extraterrestrial life. The 95% margin of error for the poll was 4%. Which of the following statements is correct?
- A. We can be certain that the percentage of all college students who believe there is extraterrestrial life is in the interval 58% to 66%.
  - B. The chance is 95% that at least 62% of all college students believe there is extraterrestrial life.
  - C. With 95% confidence we can say that the percentage of all college students who believe there is extraterrestrial life is between 58% and 66%.
  - D. The chance is 5% that at least 62% of all college students believe there is extraterrestrial life.

KEY: C

40. A polling agency conducted a survey of 100 doctors on the question: "Are you willing to treat women patients with the recently approved pill RU-486?" The conservative margin of error associated with the 95% confidence interval for the percent who say 'yes' is
- A. 50%.
  - B. 10%.
  - C. 5%.
  - D. 2%.

KEY: B

Chapter 5

41. In a random sample of  $n = 400$  people, 40% say they have allergies. The conservative margin of error is reported to be 5%. What is a 95% confidence interval for the percent of the population that has allergies?
- A. 25% to 55%
  - B. 30% to 50%
  - C. 35% to 45%
  - D. 40% to 45%

KEY: C

42. Which one of the following statements is true?
- A. Increasing the sample size of a survey decreases the margin of error.
  - B. Increasing the sample size of a survey increases the margin of error.
  - C. Increasing the sample size of a survey decreases the impact of response bias.
  - D. Increasing the sample size of a survey increases the impact of response bias.

KEY: A

43. Recently a group of adults who swim regularly for exercise was evaluated for depression. Each person's depression status was recorded as either currently depressed, likely to get depressed, or not depressed. It turned out that these swimmers were less likely to be depressed than the general population. What sample size was used if the margin of error was about 6%?
- A.  $n = 17$
  - B.  $n = 36$
  - C.  $n = 278$
  - D.  $n = 400$

KEY: C

44. A student organization is trying to decide whether or not to offer more movies on campus. A random sample of students is about to be asked if they are in favor of more free movies on campus. What sample size would be needed to provide a margin of error of about 3%?
- A.  $n = 33$
  - B.  $n = 1000$
  - C.  $n = 1001$
  - D.  $n = 1112$

KEY: D

45. During the graduation ceremonies, a student who writes for the school newspaper decides to sample some of the graduating seniors for a quick survey. After the students receive their diploma and step off the stage, he will ask them the question: "Have you already been accepted to college at this point?" What sample size would be needed to provide a margin of error of about 10%?
- A.  $n = 1$
  - B.  $n = 10$
  - C.  $n = 100$
  - D.  $n = 1000$

KEY: C

46. A random sample of 2000 students showed that 800 were in favor of more movies on campus. Report the results of this poll using the sample proportion and margin of error.

KEY: 40% of the sample favored more movies on campus. The margin of error was 2.2%.

47. A random sample of 400 students involved in club activities showed that 250 were women. What is an approximate 95% confidence interval for the proportion of all students involved in club activities who are women?

KEY: 57.5% to 67.5%



48. Two researchers are each conducting a study to estimate a certain population proportion. Researcher A proposes to take a random sample of  $n = 1000$  units from his population of 1 million units. Researcher B proposes to take a random sample of  $n = 2000$  units from his population of 2 million units. For which researcher will the margin of error be the smallest? Explain.

KEY: The margin of error is based on sample size (provided the population size is at least 10 times larger than the sample size). So the margin of error for Researcher B will be the smallest.

**Questions 49 and 50:** A claim was made that at a local college only 5% of the students that take the introductory statistics course will continue their education in statistics. The Statistics Department at the college believes that this value is too low and decides to conduct a sample survey of 100 students that completed the introductory course the previous term.

49. Identify the population.

KEY: The population is the larger group of units about which inferences are to be made, in this case, the population would be all students who attended the local college who completed the introductory statistics course. (A particular time frame may also be included in the definition of the population.)

50. Of the 100 selected students, 35 reported that they plan to continue an education in statistics. Report the results of this poll using the sample proportion and margin of error.

KEY: 35% of the sample stated they plan to continue an education in statistics. The margin of error was 10%.

**Questions 51 and 52:** According to a survey conducted by the Association for Dressings and Sauces (this is an actual association!), 85% of American adults eat salad at least once a week, so that 15% of American adults eat no salad at all! A nutritionist suspects that the percentage (of no-salad eaters) is wrong. She conducts a survey among a sample of adults in her area to test this idea.

51. What would be the approximate margin of error if the researcher randomly sampled 2000 adults?

KEY: 0.0223

52. What sample size would be needed to provide a margin of error no larger than 2%?

KEY:  $n = 2500$

### Section 5.3

53. Which of the following is *not* true about drawing a simple random sample from a population?
- A. It protects against hidden or unknown biases.
  - B. It gives each member of the population an equal chance of being selected.
  - C. It requires that the population be represented by a list.
  - D. It requires that the size of the sample be no more than 5% of the size of the population.

KEY: D

54. A company has 500 employees and would like to select a simple random sample of 25 of them for a study. Of the following, only one fits the definition of a simple random sample. Which one is it?
- A. Randomly choose one person whose last name begins with A, one with B, and so on, omitting X because it's least common.
  - B. Randomly choose 25 pages from the employee directory, then choose the first person listed on each of those pages.
  - C. Number the employees from 1 to 500 based on seniority and randomly choose one person from the first 20 names on the list, one from the next 20, and so on.
  - D. Number the employees from 1 to 500 in random order and choose the first 25 names on the list.

KEY: D

55. A class consists of 30 students and the instructor wants to take a simple random sample of 5 students from this class. Which of the following represent the labels that are needed in order to select the sample using a random number table?
- A. 1, 2, 3, 4, 5.
  - B. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, ..., 30.
  - C. 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, ..., 30.
  - D. 01, 02, 03, 04, 05.

KEY: C

56. A class consists of 30 students and the instructor wants to take a simple random sample of 5 students from this class. The students are labeled 01 to 30 for the selection process. Which of the following possible samples of size  $n = 5$  is most likely to be selected?
- A. 11, 18, 10, 08, 25.
  - B. 05, 10, 15, 20, 25.
  - C. 01, 02, 03, 04, 05.
  - D. None of the above, as they are all equally likely samples.

KEY: D

57. A population consists of 500 members and a simple random sample of 100 members is to be selected. Which of the following represents the most efficient method for labeling the members for selecting the sample using a random number table?
- A. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, ..., 99, 100.
  - B. 1 = labels 001 and 501, 2 = labels 002 and 502, ... 500 = labels 500 and 000.
  - C. 001, 002, 003, ..., 499, 500.
  - D. 1 = labels 001, 101, 201, 301, ..., 901; 2 = labels 002, 102, 202, 302, ..., 902; and so on.

KEY: B

**Questions 58 to 60:** The next three questions are based on the following row of digits from a random number table. Use these digits to select a random sample following steps 1 through 4 in the text.

00157 37071 79553 31062 42411 79371 25506 69135

58. Suppose that a group of 10 students are numbered from 0 to 9. Draw a simple random sample of 5 students by starting with the digit **1** in the first string of digits 00**1**57. Which 5 students would be selected?

KEY: The five students numbered 1, 5, 7, 3, 0 would be selected.

59. Suppose that a group of 50 students are numbered from 01 to 50. Draw a simple random sample of 5 students by starting with the digit **15** from the first string of digits 00**15**7. Which 5 students would be selected?

KEY: The five students numbered 15, 33, 10, 42, and 41 would be selected.

60. Suppose that a group of 20 students are numbered from 01 to 20. Draw a simple random sample of 5 students by starting with the digit **01** from the first string of digits **001**57. Which 5 students would be selected? (Do not use the more efficient method of, for instance, choosing number 1 if 21, 41, 61, or 81 is the random number.)

KEY: The five students numbered 01, 07, 17, 06, and 11 would be selected.

**Questions 61 and 62:** First Avenue is in desperate need of repaving. The county road commission would like to survey the people who live along this road. They have a list available of all 176 household addresses along First Avenue with the addresses listed in numerical order by address.

61. Explain an efficient labeling scheme for taking a simple random sample of 20 household addresses.

KEY: The first address would receive the labels 001, 201, 401, 601, 801; the second address would receive the labels 002, 202, 402, 602, 802, and continue on until the last address which would receive the labels 176, 376, 576, 776, and 976.

62. Using the following stream of random digits, determine the first five selected addresses that would be part of the simple random sample of 20 addresses.

16332 40139 64701 46355 62340 22011 47257 74877

KEY: The first five selected addresses would be 163, 124 (= 324 – 200), 013, 164 (= 964 – 800), and 101 (= 701 – 600).

## Section 5.4

**Questions 50 to 53:** Olivia wants to learn a foreign language. To get an idea of how satisfied other students were after taking a foreign language course, she decides to take a random sample of 20 students.

63. If Olivia randomly selects 5 students from French, 5 from German, 5 from Spanish, and 5 from Chinese, the sampling method is a
- A. simple random sample.
  - B. stratified random sample.
  - C. cluster sample.
  - D. systematic sample.

KEY: B

64. If Olivia randomly selects one class among all the foreign language classes taught that year, and then interviews all students in that class, the sampling method is a
- A. simple random sample.
  - B. stratified random sample.
  - C. cluster sample.
  - D. systematic sample.

KEY: C

65. If Olivia randomly selects 20 students among all students taking a foreign language class that year, the sampling method is a
- A. simple random sample.
  - B. stratified random sample.
  - C. cluster sample.
  - D. systematic sample.

KEY: A

66. If Olivia selects every 100<sup>th</sup> student from a list of the 2000 students who took a foreign language that year, the sampling method is a
- A. simple random sample.
  - B. stratified random sample.
  - C. cluster sample.
  - D. systematic sample.

KEY: D

67. A magazine printed a survey in its monthly issue and asked readers to fill it out and send it in. Over 1000 readers did so. This type of sample is called a
- A. cluster sample.
  - B. stratified sample.
  - C. simple random sample.
  - D. self-selected sample.

KEY: D

68. A population of taxpayers is divided into five income levels and a simple random sample is selected from each one for an audit. This is an example of a \_\_\_\_\_ sample.
- A. cluster
  - B. stratified
  - C. simple random
  - D. systematic

KEY: B

69. Taxpayers in a population are listed in order of increasing income, with people with the lowest incomes listed first. A sample is selected by randomly choosing one of the first 100 names, then choosing every 100<sup>th</sup> name from that point forward. This is an example of a \_\_\_\_\_ sample.
- A. stratified
  - B. cluster
  - C. simple random
  - D. systematic

KEY: D

70. Children at a Kindergarten school are divided into three groups: Those who are the oldest/first child (there are 200 of them, call this Group I), those who are the second child (there are 150, call this Group II), and those who are the third (or even fourth or fifth) child (there are 75, call this Group III). The social worker at the school will take a sample to find out how many sleepovers they have had (without parents) at a friend's house. She suspects that first-borns do not get as many as later children, so she decides to take a simple random sample of 15 children from each of the three groups. What sampling technique will be used to obtain the sample of children?
- A. stratified
  - B. cluster
  - C. simple random
  - D. multi-stage

KEY: A

71. A librarian wants to gather data on the number of times books of fiction are checked out during the year. The library has 2000 books of fiction, of which 500 are murder mysteries, 400 are science fiction, and the rest are "other". She decides to sample 120 books. Suppose the librarian decides to take a simple random sample of 30 books from the murder mystery category, 24 from the science fiction category, and the remaining 66 from the other category. What sampling technique was used?
- A. simple random
  - B. cluster
  - C. stratified
  - D. multi-stage

KEY: C

72. A company that produces electronics (stereos, TV's, DVD players, etc.) is thinking about adding an incentive program to the contracts of its employees. To learn about the opinion of its employees regarding this idea, a survey will be sent out to a sample of the employees. To ensure that employees of all three plants are represented in the survey, the statistician decides to take a sample of employees from each of the three plants. At each plant, a list is available of all employees. A 1-in-20 systematic sample will be taken from each plant. What sampling technique is being used to sample the employees?
- A. simple random
  - B. cluster
  - C. stratified
  - D. multi-stage

KEY: D

**Questions 73 and 74:** The Michigan Department of State Police keeps track of the number of points received for various traffic violations by Michigan drivers. The department is interested in examining the relationship between the number of points received and the insurance premium. The following chart shows the number of drivers in each category of points according to the records from last year.

| Point Category    | High (8 or more pts) | Medium (4-7 pts) | Low (0-3 pts) |
|-------------------|----------------------|------------------|---------------|
| Number of Drivers | 500                  | 1200             | 18,000        |

The department will take a simple random sample of 3% of the drivers in the Low category and 2% of the drivers in the other two categories.

73. How many drivers will be included in the study?

- A. 574
- B. 24
- C. 540
- D. 399

KEY: A

74. What sampling method will be used to obtain the sample?

- A. simple random
- B. cluster
- C. stratified
- D. multi-stage

KEY: C

75. There are five cities in a politician's district but redistricting has been proposed for the state. The politician would like to know which city he should try to remove from his district. He plans to conduct a survey to find out his approval rating. Which one of the following sampling plans would be most useful for his purposes?

- A. Take a stratified sample with political parties as the strata.
- B. Take a stratified sample with the five cities as the strata.
- C. Take a simple random sample across his district.
- D. Take a cluster sample with the five cities as clusters.

KEY: B

76. Which of the following is the best example of a population from which a cluster sample would be easier to obtain than a simple random sample?

- A. Passengers who will be flying in the upcoming week on a certain airline.
- B. Registered drivers in a certain state.
- C. Employees of a large company.
- D. Members of a national organization.

KEY: A

77. One problem with using random digit dialing to get a representative sample of households in a city is that

- A. the method can only be used to get a sample for the whole country.
- B. households with multiple telephones on the same line will be over-represented.
- C. households with multiple phone numbers will be over-represented.
- D. households with unlisted phone numbers will not be represented.

KEY: C

**Questions 78 to 80:** These questions pertain to how to collect a random sample of 800 airline passengers from 100 particular flights leaving San Francisco International Airport. Assume there are 200 passengers aboard each flight.

78. Describe how to draw the sample as a stratified random sample, using the 100 flights as the strata.

KEY: Simple random samples of 8 people could be drawn from the 200 passengers from each flight, resulting in a total sample of 800 passengers.

79. Describe how to draw the sample as a cluster random sample, using flights as the cluster.

KEY: A random sample of 4 of the 100 flights could be chosen. All passengers in the 4 flights would be included in the cluster sample, resulting in a sample of 800 passengers.

80. Describe how to draw the sample as a systematic sample from these flights.

KEY: A numbered list of all  $100 \times 200 = 20,000$  passengers on these flights could be obtained. For a systematic sample, one should be drawn from each set of  $20,000/800 = 25$  passengers. Randomly choose among the first 25 passengers on the list to start, and then choose every 25th passenger until 800 passengers were obtained.

81. Both cluster and stratified sampling rely on sub-groups in the population. If the sub-groups are all very similar (they can all be considered mini-populations), which sampling method would be preferred?

KEY: Cluster sampling is the preferred method.

82. On a chilly spring afternoon, 10 lab sections of a statistics class all have full attendance. The 10 lab sections each have the same number of students attending. A survey is about to be administered to a sample of the students. It has been decided to first randomly select 3 of the 10 lab sections and then give the evaluation to a simple random sample of 25% of the students in those selected lab sections. What sampling technique is being used for this survey?

KEY: This is a multistage sampling design, with two stages of sampling: Students are first clustered by lab sections and a few labs are selected at random, then a simple random sample of students from each lab is taken to administer the survey to.

**Questions 83 and 84:** A population under study is divided into 10 smaller groups (e.g. lab sections for a large class). The opinions of the people within each group vary quite a bit, whereas the variation between the groups is not that great.

83. If you were to take a sample of these people and would like to utilize these groups in some fashion, which sampling technique should you use?

KEY: Cluster sampling can be used, because the groups don't vary much from one to the next. So no matter which group is selected, we will get a good representation of all the opinions.

84. Suppose you decide to take a cluster sample. However, you don't really know exactly how many people make up this population, would you still be able to take your sample?

KEY: Yes, for cluster sampling we only need to know the number of clusters.

## Section 5.5

85. A television news program asks viewers to log onto their website and vote on a different issue each day. The website allows visitors to select one of two choices and submit a vote. The results of the poll are reported the next day on the news program. The population to which the results of these polls can be extended is
- A. all viewers of the news program.
  - B. all visitors to the website.
  - C. all viewers who have voted in any of the polls.
  - D. only viewers who voted in that specific poll.

KEY: D

86. Which of the following would be most likely to produce selection bias in a survey?
- A. Using questions with biased wording.
  - B. Only receiving responses from half of the people in the sample.
  - C. Conducting interviews by telephone instead of in person.
  - D. Using a random sample of students at a university to estimate the proportion of people who think the legal drinking age should be lowered.

KEY: D

87. If the target population is all U.S. adults, then a telephone directory would not make a good sampling frame because
- A. people with unlisted numbers or no phones are not included.
  - B. not everyone can understand English.
  - C. some people have more than one phone number.
  - D. some people will not answer their phone.

KEY: A

**Questions 88 to 91:** A poll is planned to determine what proportion of all students favor an increase in fees to support a new track and field stadium. For each study design, identify the main problem.

88. The questionnaire will be published in the student newspaper and the first 1000 completed questionnaires will be analyzed.
- A. Wrong sampling frame
  - B. Volunteer response
  - C. Response bias
  - D. Nonresponse bias

KEY: B

89. Questionnaires will be left on tables by the student union. When 1000 questionnaires are completed, the results will be reported.
- A. Wrong sampling frame
  - B. Volunteer response
  - C. Response bias
  - D. Nonresponse bias

KEY: B

90. A random sample of 1000 students with email addresses will be drawn from a list of all students with email addresses. Not every student has a campus email address.
- A. Wrong sampling frame
  - B. Volunteer response
  - C. Response bias
  - D. Nonresponse bias

KEY: A



91. A random sample of students will be called on the phone on Friday night and asked whether they favor a fee increase.
- A. Wrong sampling frame
  - B. Volunteer response
  - C. Response bias
  - D. Nonresponse bias

KEY: D

92. For the *Literary Digest* poll of 1936 the magazine sent out questionnaires to 10 million people, of whom 2.3 million answered the questionnaire. What was the main problem with this poll?
- A. Wrong sampling frame
  - B. Volunteer response
  - C. Response bias
  - D. Selection bias

KEY: B

93. A librarian wants to gather data on the number of times books of fiction are checked out during the year. The library has 2000 books of fiction, of which 500 are murder mysteries, 400 are science fiction, and the rest are “other”. She decides to sample 120 books. Suppose the librarian takes the next 120 fiction books that are turned in and calculates the average number of times these books were checked out in the last year. What sampling technique was used?
- A. Self-selected sample
  - B. Volunteer sample
  - C. Convenience sample
  - D. Random sample

KEY: C

94. A short questionnaire is to be given to a sample of people. The questionnaire involves the topic of abortion. The woman who will hand out the questionnaire is a big pro-life supporter and on the day of the survey shows up with pamphlets and picket-signs expressing her opinion. What type of bias may this cause?
- A. Selection bias
  - B. Response bias
  - C. Non-response bias
  - D. Volunteer bias

KEY: B

95. A professor wanted to find out whether students enrolled in her statistics class would like her to increase her office hours. She plans to take a random sample of 20 students among those coming to class on the next lecture day. Explain what the difficulty with her method is.

KEY: Her sampling frame is all students who came to that lecture. Since not all students come to every lecture, she used the wrong sampling frame.

96. A student survey to find out whether more classes should be offered in the evening was mailed to all enrolled students at a university. Explain what a potential difficulty with this method is.

KEY: The sample of students who respond will be based on a volunteer response. The sample of respondents may not be representative of the opinion of all students in the university. Those who want more evening classes are probably more likely to respond.

97. A questionnaire on a table at a shopping mall was completed by 100 shoppers. The results were used to claim that the average amount of money spent by Americans increased from the previous year. Explain why this claim is questionable.

KEY: The sample of shoppers who responded was a self-selected sample, so the results cannot be generalized to any larger population.

98. Explain what is wrong with the following statement: “When a survey is mailed out to 500 people and only 150 respond, this results in a bias called response bias.”

**KEY:** The resulting bias is non-response. Responding to a survey (or not) is voluntary, and those who respond are likely to have stronger opinions than those who do not respond. This type of non-response bias can lead to systematically overestimating or underestimating the truth about a population.

99. A college department wants to learn about the jobs that its alumni are working in. An online survey is set up on the department website which invites alumni to complete and includes demographic questions and questions about their job (current and history). Give a problem that can arise with this survey.

**KEY:** There are a number of problems that could arise, listed here are a few possible answers. First it is an online survey, thus only those alumni who check this website would potentially be able to respond. It is based on volunteer response and those alum who are not currently working or working in a ‘less prestigious’ job would be less inclined to respond.

## Section 5.6

100. Which of the following types of bias could *not* be caused by improper wording of questions?

- A. Response bias
- B. Selection bias
- C. Intentional bias
- D. Unintentional bias

KEY: B

101. Which of the following would be most likely to produce response bias in a survey?

- A. Using questions with biased wording.
- B. Only receiving responses from half of the people in the sample.
- C. Conducting interviews by telephone instead of in person.
- D. Using a random sample of students at a university to estimate the proportion of people who think the legal drinking age should be lowered.

KEY: A

102. What is the correct definition for an anonymous survey?

- A. The researcher promises not to release identifying information about respondents.
- B. The participants do not know which treatment they are receiving.
- C. The researcher does not know the identity of the respondents.
- D. The researcher does not know which treatment participants are receiving.

KEY: C

103. Sometimes questions are worded in such a way that the meaning is misinterpreted by a large percentage of the respondents. Which of the following types of bias would be most likely to be produced?

- A. Nonresponse bias
- B. Response bias
- C. Intentional bias
- D. Unintentional bias

KEY: D

104. Consider the following survey question: “Do you agree that the alcohol industry should be banned from targeting the teenage girls in their advertising?” Which of the following types of bias is this question an example of?

- A. Nonresponse bias
- B. Selection bias
- C. Intentional bias
- D. Unintentional bias

KEY: C

105. Consider the following survey question: “Shouldn’t federal funding not be reduced for HIV awareness programs, since many in U.S. misunderstand the information about an HIV vaccine?” Which of the following types of bias is this question an example of?

- A. Ordering of questions
- B. Unnecessary complexity
- C. Desire of respondents to please
- D. Unintentional bias

KEY: B

106. Two polls were conducted to determine how much support there was for public funding of private schools. However, the results from the two polls were contradictory. One poll asked whether “students should attend private school at public expense”. The other poll asked “How much do you support providing parents with the option of sending their children to the school of their choice – either public, private or parochial – rather than only to the school to which they are assigned?” Explain why the wording of the questions may have affected the results.
- KEY: The poll with the question “public expense” emphasized the public cost of the program, which may have encouraged a negative response, while the other poll emphasized the freedom of parents to choose the school for their children, which may have encouraged a positive response.
107. Reword the following survey question so that it is less confusing. “Do you disagree with the policy of not allowing first year students parking permits on campus?”
- KEY: Do you support the policy that first-year students are not allowed to buy a parking permit on campus?
108. The following statement that was provided as part of the fine print at the end of the reported results of a survey: “*In addition to sampling error, question wording in conducting surveys can introduce error or bias into the findings of public opinion polls.*” What is the type of *bias* being referred to here?
- KEY: This type of bias is response bias.
109. Explain the difference between an open question and a closed question.
- KEY: An open question is one in which respondents are allowed to answer in their own words; a closed question is one in which respondents are given a list of alternatives from which to choose their answer.
110. The biggest problem with open questions is that the results can be difficult to summarize. If closed-form questions are preferred, explain how a pilot survey might be used to better formulate the closed-form questions.
- KEY: The closed-form questions should first be presented as open questions in a pilot survey to a test sample before the real survey is conducted. Then the most common responses to the open questions can be included in the list of choices for the closed questions in the actual survey.
111. For a short mid-term survey of his students, a Graduate Student Instructor has put four questions on a piece of paper. He is planning to hand these out this afternoon. The last two questions are: (1) “What grade do you think you are getting in the course?” and (2) “What grade would you give the instructor so far?” Both questions can be answered with A, B, C, D, or F. What is a potential problem with these questions?
- KEY: The order of the questions could influence the answers of the students. Students who are not doing well in the course may be affected by that thought and give the instructor a bad grade, even though he is doing well.