Practice and Examples.

1) About 20% of the engineering students at a large university are women. The school plans to poll a sample of 200 engineering students about the quality of student life.

a) If an SRS of 200 is selected, about how many women do you expect to find in the sample?

$200(.2) = 40$ 40 women

b) If the polls wants to be able to report separately the opinions of male and female students, what type of sampling design would you suggest? Why?

a stratified sample - this would allow us to draw a conclusion about male/female students seperately as well as students as a whole

2) Each of the following is a source of error in a sampling survey. Label each as a sampling error or a nonsampling error and explain your answers.

a) The telephone directory is used as a sampling frame.

sampling error - this is an example of undercoverage since not all of the population is included in sampling frame

b) The person cannot be contacted in five calls.

nonsampling error - this is an example of nonresponse

c) Interviewers choose people walking by on the sidewalk to interview.

sampling error - there is likely some bias as they select people to interview & could be a convenience sample

3) Suppose you want to know the average amount of money spent by fans attending opening day for the Cleveland Indians baseball season. You get permission from the team's management to conduct a survey at the stadium but they will not allow you to bother the fans in the club seating or box seat areas (the most expensive seating). Using a computer, you randomly select 500 seats from the rest of the stadium. During the game, you ask the fans in those seats how much they spent that day.

a) Provide a reason why this survey might yield a biased result.

the most expensive seats are excluded from the sampling frame

b) Explain whether the reason you provided in (a) is a sampling error or nonsampling error.

Sampling error (undercoverage) since part of the population is excluded when choosing the sample

4) A poll of 586 adults who used the Internet in the past week were asked whether "the Internet has made your life much better, somewhat better, somewhat worse, much worse, or has it not affected your life either way." In all, 152 of the 586 said "much better."

a) What is the population for this sample survey?

all adults who use the internet

b) Use the quick method to find a margin of error. Then give a complete confidence statement for a conclusion about the population.

$\hat{p} = \frac{152}{586} = .259$

$moe = \frac{1}{\sqrt{586}} = .041$

$.259 \pm .041 = (.218, .3)$

We are 95% confident that the true proportion of adults who use the internet and have had life made "much better" is between .218 and .3.

5) Determine which type of sampling method was used for each of the following.

a) Student organization looking to get signatures for a petition camp out in front of Class of 1950 Lecture Hall.

convenience sample

b) Select three students from a class to receive ice cream by putting all the students' names in a hat and picking out three names randomly.

simple random sample

c) Select three female students and three male students to receive ice cream by putting all the men's names in one hat and all the women's names in a different hat and picking out three names from each hat.

stratified sample

d) In Fall 1995, the BBC in Britain requested viewers to call the network and indicate their favorite poem.

voluntary response sample

e) Divide the class into four groups (freshman, sophomore, junior and senior) and take a random sample of two students from each group.

stratified sample

f) Priceline.com randomly e-mails a Customer Satisfaction Survey for certain transactions done on its site in which customers choose to either respond or not.

voluntary response sample

g) A researcher has a population of 100 third grade children from a local school district from which a sample of 25 children is to be selected. Each child's name is put on a list, and each child is assigned a number from 1 to 100. Then the numbers 1 to 100 are written on separate pieces of paper and shuffled. Finally, the researcher picks 25 slips of paper and the numbers on the paper determine the 25 participants.

simple random sample

h) A sociologist conducts an opinion survey in a major city. Part of the research plan calls for describing and comparing the opinions of four different ethnic groups: African Americans, Asian Americans, European Americans, and Native Americans. For a total sample of 300, the researcher selects 75 participants from each of the four predetermined subgroups.

stratified sample

i) Every fifth person boarding a plane is searched thoroughly.

systematic sample

6) Which of the following are sources of sampling error and which are sources of nonsampling error? Explain your answer.

a) The subject lies about past drug use.

nonsampling error → response error

b) The typing error is made in recording the data.

nonsampling error → processing error

c) Data are gathered by asking people to mail in a coupon printed in a newspaper.

sampling error → undercoverage (people who don't get the newspaper are excluded)

d) A hot dog vendor is collecting data on where is the best location to sell hot dogs. The city licenses 4 locations. Jack starts at the site in the park that is the most packed.

sampling error → results from sample in most packed location will differ from population

7) A police officer is interested in the drug and alcohol use of all high school students. He does to his local high school to question some of the students.

a) Suppose the officer initially decides to question students by asking them face to face. If a student does not want to tell the officer the truth and simply declares himself drug free, what error is introduced into the study?

Nonresponse  Undercoverage  **Response Error**

b) .Suppose the officers grabs a random sample but only does so of the freshman students. What type of bias does this study suffer from?

Nonresponse  **Undercoverage**  Response Error

c )Suppose the officer distributes an anonymous survey to 100 random high school students but only receives 12 of the surveys back. What types of bias does this study suffer from?

**Nonresponse**  Undercoverage  Response Error

8) A survey of users of the Internet found that males outnumbered females by nearly 2 to 1. This was a surprise, because earlier surveys had put the ratio of men to women closer to 9 to 1. Later in the article, we find this information: "Detailed surveys were sent to more than 13,000 organizations on the Internet; 1,468 usable responses were received. According to Mr. Quarterman, the margin of error is 2.8 percent, with a confidence of 95 percent.

a) What was the response rate for this survey? (The response rate is the percent of the planned sample that responded.) $\frac{1468}{13{,}000} = 0.113$ (11.3%)

b) Use the quick method to estimate the margin of error of this survey. Is your result close to the 2.8% claimed? $\frac{1}{\sqrt{13000}} = .009$ no

c) Use your answers from (a) and (b) to construct the 95% confidence interval.

(.104, .122)

9) Does increasing the sample size increase or decrease the margin of error?

as sample size increases, margin of error decreases

10) Does increasing the sample size narrow or widen the confidence interval?

as sample size increases, margin of error decreases and so confidence interval narrows

12) A nationwide poll was taken of 1432 teenagers (ages 13-18). 630 of them said they have a TV in their room.

a) What is the sample proportion ($\hat{p}$)?

$$\hat{p} = \frac{630}{1432} = 0.44$$

b) What is the margin of error?

$$\frac{1}{\sqrt{1432}} = .026$$

c) What is the 95% confidence interval?

$.44 \pm .026$

$(.414, .466)$

d) Interpret your 95% confidence interval.

We are 95% confident that the true proportion of teenagers with a TV in their room is between .414 and .466.

13. A poll of 1070 teens aged 13-17 finds that 742 have received personal messages online from people they don't know. The announced margin of error for this result is plus or minus three percentage points. A parents group is satisfied with 95% confidence of the survey but wants a smaller margin of error than ± 3%. How can we get a smaller margin of error, still with 95% confidence?

increase the sample size

14) Suppose a 95% confidence interval is stated as $(0.315, 0.365)$. *(Think about working backwards and where the proportion is in relation to the two numbers in the interval and how that is related to the margin of error)*

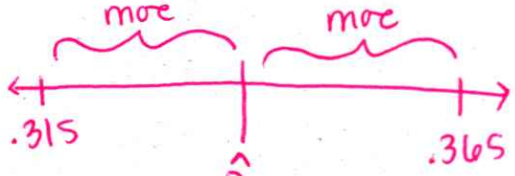
a) What is the sample proportion?

$\hat{p} = .34$

b) What is the margin of error?

$moe = .365 - .34 = .025$

or $.34 - .315 = .025$




$$\hat{p} = \frac{.365 + .315}{2}$$

$\hat{p} = .34$

15) You are interested in the proportion of Stat 113 students that will end up with an A this semester. You survey 20 students enrolled in Stat 113. Identify the

a) population: all students in stat 113

b) sample: 20 students sampled enrolled in stat 113

c) variable *(what is being studied)*: proportion of students who got A's

16) The mean income of all subscribers to a particular magazine is **$26000**. We draw a random sample of 100 subscribers and find that their mean income is **$27300**. Identify the

a) population: all subscribers

b) sample: 100 subscribers

c) parameter: mean income of $26,000

d) statistic: mean income of $27,300

17) The average GPA for all female volleyball players in a particular college is **2.8**, and their mean height is **182cm**. Identify the parameter or statistic if there is any.

parameters: avg. GPA of 2.8 & mean height of 182 cm

statistics: none

18) To reduce bias, you should: use random sampling or use a larger sample *(choose 1)*

19) To recue the variability, you should: use random sampling or use a larger sample *(choose 1)*

20) Out of 54 randomly selected patients of a local hospital who were surveyed, 49 reported that they were satisfied with the care they received. Construct and interpret a 95% confidence interval for the percentage of all patients satisfied with their care at the hospital.

$\hat{p} = \frac{49}{54} = .907$

$moe = \frac{1}{\sqrt{54}} = .136$

ci: (.771, 1)

cs: We are 95% confident that the true proportion of patients @ a hospital who are satisfied w/ their care is between .771 & 1

21) Determine if each of the following is a population parameter or a sample statistic.

a) A survey of 1000 U.S. adults found that 40% think that the Internet is the best way to get news and information. sample statistic

b) At a college, 90% of the members of the Board of Trustees approved the contract of the new president. population parameter

c) A survey of 733 small business owners found that 17% have a current job opening. sample statistic

d) The 2182 students who accepted admission offers to Northwestern University in 2009 have an average SAT score of 1442. (Source: Northwestern University)

population parameter

Statistical Reasoning
Unit 5 Review: Sampling and Surveys

Name _____
Date _____ Day _____

**Vocabulary (Matching)**

Bias, Convenience Sample, Margin of Error, Non-response Error, Non-sampling Error, Population, Parameter, Probability sample, Processing Error, Question-wording Bias, Random Sampling Error, Response Error, Sample, Statistic, Simple Random Sample (SRS), Stratified sample, Sampling Frame, Undercoverage, Voluntary Response, Variability

Study terms & definitions!

**Multiple Choice/Short Answer Type Questions**

1. A talk radio show host asked the following question: "Should drivers be banned from using cell phones?" Listeners were encouraged to go online and vote. What type of sampling method is used? voluntary response

2. A student did a survey on how much sleep high school students need. To make collection easy, she surveyed the first 100 students to arrive at school on a particular morning.
   a. What type of sampling method is used? convenience sample
   b. Is the sampling method biased? Why?
      *answers may vary*
      The first 100 students at school were likely up earlier so they may have gotten less sleep than others

3. Do adults typically wash their hands after using a public restroom? In a telephone survey of 1000 U.S. adults, 92% said they always wash their hand.
   a. What type of sampling method is used? convenience sample
   b. Is this sampling method biased? Why (what type of errors could occur)?
      Yes, people may not tell the truth (response error)

4. A school newspaper is conducting a survey to find out students favorite fast food restaurant. Students are divided by age group: Age 14-15, 16-17, 18 and older. The school newspaper randomly chooses students at lunch and ask them their favorite fast food restaurant.
   a. What type of sampling method is used? stratified random sample
   b. Is it biased?
      This sampling method is typically not biased

Statistical Reasoning

**Unit 5 Review: Sampling and Surveys**

Name _____

Date _____ Day _____

5. A school newspaper also wants to know how many years of post-graduate study do teachers have. They number the teachers in the school from 1 to 85 and then use a random number digit table to determine the first teacher selected. Then, the newspaper chooses every 10th teacher. On the day they did the survey, the entire fine arts department was absent due to a field trip.

   a. What type of sampling method is used? systematic sampling

   b. Is it biased? Yes, since the fine arts department was all absent; typically systematic samples are not biased

6. The actual proportion of American who have bought a lottery ticket is about 47%. A Gallup Poll of 1,600 adults shows that 57% of Americans have bought a lottery ticket in the last 12 months. Find the following:

   a. Population: adult Americans
   b. Parameter: 47%
   c. Sample: 1,600 adults
   d. Statistic: 57%

7. A national sample survey interviewed 4,000 people age 18-25 by telephone. One question asked was whether they agreed on this statement: "Some people say we should have a third major political party in this country in addition to Democrats and Republicans." Of the people asked, 53% agreed that we should have a third party.

   a. Population: people aged 18-25
   b. Parameter: percent who agree that we should have a third party (unknown)
   c. Sample: 4,000 people
   d. Statistic: 53%

   What is the margin of error for this scenario? What happens to the margin of error if we increase the sample size to 6,000? What if we decrease it to 1,000?

   $moe = \frac{1}{\sqrt{4000}} = 1.6\%$ .016

   If $n = 6{,}000 \rightarrow moe = \frac{1}{\sqrt{6000}} = 1.3\%$ .013

   If $n = 1{,}000 \rightarrow moe = \frac{1}{\sqrt{1000}} = 3.2\%$ .032

   What is the 95% confidence interval for this scenario? *(Using original*

   $53\% \pm 1.6\% = (51.4\%, 54.6\%)$

   or (.514, .546)

   Make a confidence statement about the scenario.

   We are 95% confident that between 51.4% and 54.6% of adults believe we should have a third party.

   We are 95% confident that the true prop. of adults who believe we should have a third party is between .514 and .546

Statistical Reasoning
**Unit 5 Review: Sampling and Surveys**

Name _____
Date _____ Day _____

8. A table of random digits was used to select 22 students out of a class of 38. The 22 students represent the Sample and the 38 students represent the population.

---

9. A student at a large university wants to study the responses that students receive when calling an academic department for information. She wants to select an SRS of 6 departments from the following list for her study.

| | | | | |
|---|---|---|---|---|
| 1 Agronomy | 8 Art and Design | 14 Audiology | 20 Biochemistry | 26 Biology |
| 2 Chemistry | 9 Communication | 15 Computer | 21 Science | |
| 3 Consumer Science | 10 Education | 16 Electrical | 22 English | |
| 4 Foreign Languages | 11 History | 17 Horticulture | 23 Industrial Eng | 27 Managemnt |
| 5 Mathematics | 12 Nursing | 18 Pharmacology | 24 Philosophy | 28 Physics |
| 6 Political Science | 13 Psychology | 19 Sociology | 25 Statistics | 29 Veterinary |
| 7 Anatomy | | | | |

27816 78416 18329 21337 35213 37741 04312 68508

08421 44753 77377 28744 75592 08563 79140 92454

Use the partial table of random digits above to select an SRS of 6 departments. Circle your answers.

27 → Management
13 → Psychology
16 → Electrical
21 → Science
18 → Pharmacology
4 → Foreign Languages

10. The school wants to know student feeling about the food in the cafeteria that is served daily. A representative stands in the cafeteria and randomly selects 40 students during A lunch to answer a short survey. What type of sampling method is used? What is the sample for this scenario? What is the population?

Sampling method → simple random sample (SRS)
sample: 40 students
population: all students at school

11. What if we took the above scenario and asked 10 people from each lunch period (A, B, C, and D). Does that change the sampling method? Do students have an equal chance of being chosen?

Sampling method would change to a stratified random sample → each student in each lunch has an equal chance of being chosen

Statistical Reasoning

Unit 5 Review: Sampling and Surveys

Name _____

Date _____ Day _____

15. A national sample survey interviewed 3,800 people age 18 and older nationwide by telephone. One question asked was whether they agreed on this statement: "Some people say we should have a third major political party in this country in addition to Democrats and Republicans." Of the people asked, 53% agreed that we should have a third party.

a. Identify the following:

Population: all people age 18 and older in America

Parameter: the percentage who agree we should have a third party (unknown)

Sample: 3,800 people age 18 and older

Statistic: 53%

b. Find the margin of error for this scenario.

$m.o.e. = \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{3800}} = 0.016 = 1.6\%$

c. What is the 95% confidence interval for this scenario.

$53\% \pm 1.6\% \rightarrow 53 - 1.6 = 51.4\%$, $53 + 1.6 = 54.6\%$ } (51.4%, 54.6%)

or

(.514, .546)

d. Make a confidence statement about the percent of all those age 18 or older in the nation who would believe we should have a third major political party.

We are 95% confident that between 51.4% and 54.6% of American adults believe we should have a third political party.

We are 95% confident that the true proportion of American adults who believe we should have a third political party is between .514 and .546

**STA 2e: Chapter 5 Review** **Name** _____

***Match each word or phrase with its definition.***

| | | | |
|---|---|---|---|
| Q | 1. anecdotal data | A. | error which occurs when an individual chosen for the sample cannot be contacted or refuses to cooperate |
| V | 2. bias | B. | a number that describes the population |
| J | 3. convenience sample | C. | error that occurs when a subject gives an incorrect response |
| G | 4. margin of error | D. | any sample which uses chance to select the sample |
| A | 5. nonresponse error | E. | sampling error not related to the act of selecting a sample from the population |
| E | 6. nonsampling error | F. | A sample obtained by dividing the population into subgroups according to various homogeneous characteristics and then selecting members from each subgroup for the sample |
| B | 7. parameter | G. | measure of the sampling variability which tells how much confidence we have in the results of a survey |
| N | 8. population | H. | bias which occurs when some groups in the population are left out of the process of choosing the sample |
| D | 9. probability sample | I. | A sample obtained by numbering each element in the population and then selecting every third or fifth or tenth, etc., number from the population to be included in the sample |
| R | 10. processing error | J. | sample which chooses the individuals easiest to reach |
| S | 11. question-wording bias | K. | how spread out the values of the sample statistic are when we take many samples |
| T | 12. random sampling error | L. | list of individuals from which a sample is actually selected |
| C | 13. response error | M. | number that describes a sample |
| P | 14. sample | N. | the entire group of individuals about which we want information |
| L | 15. sampling frame | O. | sample consisting of people who choose themselves by responding to a general appeal |
| U | 16. simple random sample | P. | the part of the population from which we actually collect information |
| M | 17. statistic | Q. | information about one individual, but absolutely nothing about the population this individual represents |
| F | 18. stratified sample | R. | error in mechanical tasks such as in doing arithmetic or in entering responses into a computer |
| I | 19. systematic sample | S. | bias created by the wording of the survey question(s) |
| H | 20. undercoverage | T. | deviation between the sample statistic and population parameter caused by chance in selecting a random sample |
| K | 21. variability | U. | a sample of size n individuals from the population chosen in such a way that every set of n individuals has an equal chance to be the sample actually selected |
| O | 22. voluntary response sample | V. | systematic favoring of certain outcomes |