Objective: Classifying Studies

Steps in a Statistical Study:

1. A U.S. organization wants to know what percent of teen drivers regularly wear seatbelts. Find the population and parameter.

   Population: ___________________________ Parameter: ___________________________

2. A telephone company wants to know how long of a wait time a customer can expect when calling for assistance. Find the population and parameter.

   Population: ___________________________ Parameter: ___________________________

3. A company selling baby products is considering advertising on a news program. What could be the population and parameter of interest to this company?

   Population: ___________________________ Parameter: ___________________________
4. A high school softball team is planning on raising money by selling t-shirts. They plan on selling three different colors, so they need to know which colors to order. What is the population and parameter of interest for this question?

Population:_____________________________ Parameter:_____________________________

Study Types

Observation studies are used in place of an experiment when it would be unethical to do an experiment
For example: studying the effects of smoking

Determine a type of Study

5. Determine whether each situation describes a survey, an experiment or an observational study. Then identify the sample and suggest a population from which is may have been selected.

a. A record label wants to test three designs for an album cover. They randomly select 50 teenagers from local high schools to view the covers while they watch and record their reactions.

____________________   ____________________  ____________________

b. The city council wants to start a recycling program. They send out a questionnaire to 200 random citizens asking what items they would recycle.

____________________   ____________________  ____________________

c. Scientists study the behavior of one group of dogs given a new heartworm treatment and another group of dogs given a false treatment or placebo.

____________________   ____________________  ____________________

d. The yearbook committee conducts a study to determine whether students would prefer to have a print yearbook or both print and digital yearbooks.

____________________   ____________________  ____________________

Created by Washington High School Math Department
Choose a Study Method

6. Determine whether each situation describes a survey, an experiment or an observational study. Explain your reasoning.
   a. A pharmaceutical company wants to test whether a new medicine is effective.
   b. A news organization wants to randomly call citizens to gauge opinions on a presidential election.
   c. A research company wants to study smokers and nonsmokers to determine whether 1-year of smoking affects lung capacity.
   d. A national pet chain wants to know whether customers would pay a small annual fee to participate in a rewards program. They randomly select 200 customers and send them questionnaires.

Objective: Designing a study and Recognize Bias

Bias:

Experimental Group:

Control Group:

Looking for Bias

Avoid questions that are:

- [ ]
- [ ]
- [ ]
- [ ]

** Also look for: __________________________________________________________

Recognize Bias

7. Determine whether each survey question is biased or unbiased. If biased, explain your reasoning.
   a. Don’t you agree that the cafeteria should serve healthier food?
   b. How often do you exercise?
   c. How many glasses of water do you drink a day?
   d. Do you prefer watching exciting action movies or boring documentaries?
Designing a Study:
  1. Clearly state the objective
  2. Identify the population
  3. Carefully choose UNBIASED survey questions

Design a Study
Jim is writing an article for his school newspaper about online courses. He wants to conduct a survey to determine how many students at his school would be interested in taking an online course from home.

1.) State the objective of the survey: __________________________________________________

2.) Suggest a population: __________________________________________________________________________________________________________

3.) Write two unbiased questions: __________________________________________________________
   ___________________________________________________________________________________

Example:
8. In a follow-up article, Jim decides to conduct a survey to determine how many teachers from his school with at least five years of experience would be interested in teaching an online course. 1.) State the objective of the survey, 2.) suggest a population, and 3.) write two unbiased survey questions.

1.) ___________________________________________________________________________________

2.) ___________________________________________________________________________________

3.) ___________________________________________________________________________________
   ___________________________________________________________________________________

Identify Experiment Bias

Experiment: An electronics company wants to test whether using a new graphing calculator increases students’ test scores. A random sample is taken. Calculus students in the experimental group are given the new calculator to use and Algebra 2 students in the control group are asked to use their own calculator.

Results: When given the same test, the experimental group scored higher than the control group. The company concludes that the use of this calculator increases test scores.

9. Identify any bias in the design of the experiment and describe how they could be corrected.

Experiment: A research firm tests the effectiveness of a de-icer on car locks. They use a random sample of drivers in California and Minnesota for the control and experimental groups.

Results: They concluded that he de-icer is effective.
Objective: Understand Randomization

Random:

Random Sample:

State whether the sample is random. If it is not random, explain why.

10. You survey customers at a mall. You want to know which stores they shop at the most. You walk around a computer shop and choose 20 customers there for your survey.

11. A country radio station wants to know what the most popular type of music is, so they ask their listeners to call in to say their favorite type.

12. You want to survey the students in your school about their exercise habits. At lunchtime you stand by a vending machine. You survey every student who buys something from the vending machine.

13. You want to know what 7th graders think of their science class. You poll 100 random 7th graders.

14. Angela is conducting research about the most common pet owned by residents in her town. To collect a random sample, which method should she choose?

   **Method A:** Telephone every fifth person whose name appears in the town telephone directory.

   **Method B:** Ask individuals who walk their dog in a local park.

Creating a Random Sample

15. A worker at a daycare center wants to select 5 of the 42 children at random to take on a nature walk. Explain how you could generate a random sampling. (Nspire) List the first set of 5 random numbers.

16. A teacher wants to select 4 of the 28 students in her class at random to present their projects on Monday. Explain how you could generate a random sampling. (Nspire)
Objective: Analyze Distribution of Data

The prices for the random sample of personal computers are shown.

<table>
<thead>
<tr>
<th>Price (dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td>723 605 847 410 440 386 572 523</td>
</tr>
<tr>
<td>374 915 734 472 420 508 613 659</td>
</tr>
<tr>
<td>706 463 470 752 671 618 538 425</td>
</tr>
<tr>
<td>811 502 490 552 390 512 389 621</td>
</tr>
</tbody>
</table>

Use your calculator to create a box plot of the data.

Sketch:

Is the box plot symmetric?_____________________

Describe the shape of the box plot.______________________________

Explain what the shape of the box plot says about the data.

__________________________________________________________________________

Estimate the mean of the data based on the shape of the box plot._________________

Change the box plot to a histogram.

How does the shape of the histogram compare to the box plot?_____________________

The points scored per game by a professional football team for the 2008 and 2009 football season are shown.

17. 2008:____________________________________ 18. 2009:____________________________________

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>51</td>
<td>24</td>
<td>27</td>
<td>17</td>
<td>35</td>
<td>27</td>
</tr>
<tr>
<td>28</td>
<td>30</td>
<td>27</td>
<td>21</td>
<td>24</td>
<td>30</td>
<td>14</td>
</tr>
<tr>
<td>20</td>
<td>9</td>
<td>3</td>
<td>10</td>
<td>6</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>37</td>
<td>7</td>
<td>21</td>
<td>13</td>
<td>41</td>
<td>20</td>
</tr>
</tbody>
</table>

Create a histogram for each year. Sketch and describe the shape of each distribution.
19. Describe the shape of the distribution of the data.

a. 

b. 

c. 

d. 

Negatively Skewed

Symmetric

Positively Skewed

Skewed Left

Normally Distributed

Skewed Right

mean median

tail

median mean

tail
Objective: Choose Appropriate Statistics to Represent a Set of Data

Symmetric Data

Ms. Foxworth’s
Mean: 75%
Range: 65%

35, 63, 82, 95, 100

Ms. Ruiz’s
Mean: 75%
Range: 65%

35, 55, 88, 97, 100

Calculate Standard Deviation:

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n-1}}$$

<table>
<thead>
<tr>
<th>Data Point</th>
<th>Deviation ((x - \bar{x}))</th>
<th>Squared Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>35 - 75 = -40</td>
<td>1600</td>
</tr>
<tr>
<td>63</td>
<td>63 - 75 = -12</td>
<td>144</td>
</tr>
<tr>
<td>82</td>
<td>82 - 75 = 7</td>
<td>49</td>
</tr>
<tr>
<td>95</td>
<td>95 - 75 = 20</td>
<td>400</td>
</tr>
<tr>
<td>100</td>
<td>100 - 75 = 25</td>
<td>625</td>
</tr>
<tr>
<td>SUM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

20. What does the standard deviation say about the data?

_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________

21. Compare and contrast Ms. Foxworth’s test results to Ms. Ruiz’s test results.

_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
Skewed Data vs Symmetric Data

22. Students each gave a presentation as part of their class project. The length of each presentation is shown in the table.
   a. Create a histogram and describe the shape of the distribution.

   b. Describe the center and spread using either the mean and standard deviation or the five-number summary. Justify your choice.

23. The annual rainfall for a region over a 24-year period is shown in the table.
   a. Create a histogram and describe the shape of the distribution.

   b. Describe the center and spread using either the mean and standard deviation or the five-number summary. Justify your choice.

Additional Examples:
Describe the center and spread of the following data using either the mean and standard deviation or the five-number summary. Justify your choice.

24. The number of minutes Janet used each month employees on her cell phone for the last two years are shown in the table.

25. The hourly wages for a random sample of employees in a restaurant are shown in the table.
Objective: Understand the normal distribution and use the Empirical Rule to estimate population predictions.

**Normal Distribution**: a continuous, symmetric, bell-shaped distribution of a random variable.

**Standard Deviation** ($\sigma$ or $s$): measures how spread out the data is compared to the mean.

**Mean** ($\mu$ or $\bar{x}$): the average.

Use the Empirical Rule to fill in the percentages for each section.
Estimate Population Percentages
Sketch the normal curve for birth weights of babies if the weights are normally distributed with a mean of 7.6 lbs and a standard deviation of 1.3 lbs.

a. What percentage of babies weigh under 5 lbs.?

b. What percentage of babies weigh between 7.6 lbs. and 11.5 lbs.?

c. What percentage of babies weigh over 10.2 lbs.?

26. What percentage of adult American females are taller than 5’10”?

27. What percentage of adult American females are between 60” and 67.5” tall?

28. What percentage of adult American females are shorter than 5’2.5”?

29. What percentage of adult American females are between 65” and 70”?

30. What percentage of adult American females are taller than 5’2.5”?

31. What percentage of adult American females are shorter than 70”?

Students counted the number of candies in 100 small packages. They found that the number of candies per package was normally distributed with a mean of 23 candies per package and a standard deviation of 1 piece of candy.

32. Draw the graph of the normal distribution.

33. About how many packages have between 21 and 24 candies?

34. What is the probability that a package selected at random has more than 25 candies?
Additional Examples.
The heights of adult American males are approximately normally distributed with mean 69.5 in. and standard deviation 2.5 in.

35. What percent of adult American males are between 67 in and 74.5 in tall?

36. In a group of 2000 adult American males, about how many would you expect to be taller than 6ft?

The scores on the Algebra 2 final are approximately normally distributed with a mean of 150 and a standard deviation of 15.

37. What percentage of the students who took the test scored above 180?

38. If 250 students took the final exam, approximately how many scored above 135?

39. If 13.5% of the students received a B on the final, how can you describe their scores? Explain.
For an English class, the average score on a research project was 82 and the standard deviation of the normally distributed scores was 5.

40. Sketch a graph of the normal curve showing three standard deviations from the mean.

41. What percentage of students scored between 72 and 82 points?

Objectives: Estimate population percentages that do not fall within the empirical rule. Use z-values to determine how far a data value is from the mean.

Vocabulary:

z-value (score): the number of standard deviations above or below the mean.

Methods for Finding Population Percentages:
- The Empirical Rule – used when the values in question fall exactly 1, 2, or 3 standard deviations from the mean.
- Z-value (score) – used in conjunction with statistics tables, and can be used with any value
- Technology – calculators can be used with any value

Calculating a z-value.

A method is needed to determine a method for finding a score relative to the mean.
Determine the z-value for a data value of 25.

Determine a method for finding the z-value for a data value of 13.

Determine a z-value for a data value of 57.

Explain what the z-value represents.

Explain the difference between a positive and negative z-value.

42. Which is better: scoring a 670 on the math portion of the SAT, or scoring a 29 on the math portion of the ACT?

a. z-value: 

b. z-value:

c. Which score is better? Explain.

Now go back and decide whether you or your friend had a better test score.
Using Technology to Estimate Population Percentages

43. The number of videos uploaded daily to a video sharing site is normally distributed with a mean of 181,099 videos and a standard deviation of 35,644 videos. Estimate each probability using your calculator.

   a. Between 180,000 and 200,000 videos uploaded in a given day.

   Calculate the z score for 180,000:_________________ ____________________________________

   Calculate the z score for 200,000:_________________ ____________________________________

   Calculator (menu/stat/distribution/normal cdf):_________________________________________

   b. Greater than 250,000 videos uploaded in a given day.

   Calculate the z score for 250,000:_________________ _____________________________________

   Calculator (menu/stat/distribution/normal cdf):_________________________________________

44. The life spans of a certain tread of tire are normally distributed with a mean of 31,066 miles and a standard deviation of 1166 miles. Estimate each probability.

   a. Between 30,000 and 32,000 mile life span.

   b. Greater than a 35,000 mile life span.
Additional Example:

45. The cholesterol levels for adult males of a specific racial group are normally distributed with a mean of 158.3 and a standard deviation of 6.6. Estimate each probability.

a. A cholesterol level of greater than 150

b. A cholesterol level between 145 and 165

46. The heights of American males are normally distributed with a mean of 70” and a standard deviation of 3”. The heights of American females are normally distributed with a mean of 65” and a standard deviation of 2.5”.

a. Jennifer is 67” tall and her brother is 72” tall. Which sibling has a taller relative height?

b. Michael is 65” tall and his sister is 61” tall. Which sibling has a shorter relative height?
Objective: To make inferences about a population using a statistic.

Statistical Question:
- Asks about a topic of interest
- Includes a specific population
- May have variability in the responses

Inference: conclusion reached about a population based on data you collect from a sample.

Margin of Error: a statistic expressing the amount of random sampling error in a survey’s results.

Confidence Interval: gives an estimated range of values which is likely to include an unknown population parameter, the estimated range being calculated from a given set of sample

Consider the following statistical questions.
- a. Identify the population and parameter.
- b. Determine if you would use a census to find the population parameter or a sample to find the statistic.

47. A U.S. organization wants to know what percent of teen drivers regularly wear seatbelts.

48. A telephone company wants to know how long of a wait time a customer can expect when calling for assistance.

49. A company selling baby products is considering advertising on a news program.

50. The softball team is planning on raising money by selling t-shirts. They plan on selling three different colors, so they need to know which colors to order.

51. What is the average height of current Miami Heat basketball players?

52. What is the average age of first-time driver’s license applicants in Ohio?
Making Inferences About a Population

Statistical Question: What percentage of students in your high school have a cell phone?

Population: ___________________________  Parameter: ___________________________

Sample: 50 randomly selected students were surveyed and 35 reported having a cell phone.

Inferences: Determine which inference(s) would be valid.
- About 70% of the students in the school have a cell phone.
- Exactly 70% of students in the school has a cell phone.
- Around 70% of students in the state have a cell phone.

53. Suppose you reach in a box of marbles, pull out 20, and find that 8 are red.
- Is this sample proportion likely if 70% of the marbles in the box are red?
- Estimate the percent of red marbles in the box.
- If there are 150 marbles in the box, approximate the number that are red.

54. Joel has nine toy cars in a bin. He pulls out a car, records the type, and returns it to the bin. He repeats the process 30 times. His results are shown in the table.

<table>
<thead>
<tr>
<th>Car</th>
<th>Taxi</th>
<th>Bus</th>
<th>Truck</th>
<th>Ambulance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>4</td>
<td>9</td>
<td>11</td>
<td>6</td>
</tr>
</tbody>
</table>

- Based on these results how many of each car type are in the box?
  - Taxi: __________________
  - Bus: __________________
  - Truck: __________________
  - Ambulance: ____________

Additional Examples:

55. Suppose that 900 American teens were surveyed about their favorite ski category of the 2002 Winter Olympics in Park City, Utah. Ski jumping was the favorite for 20% of those surveyed. Use this result to predict how many of all 31 million American teens favor ski jumping.
56. Suppose you take a deck of playing cards, select 25 cards, and find that 10 are clubs.
   a. Is this sample proportion likely if 25% of the cards in the deck are clubs?
   b. Estimate the percent of cards in the deck that are clubs.
   c. If there are 50 cards in the deck, estimate the number that are clubs.

57. Miriam has twenty-five toy animals in a box. She pulls out a toy, records the animal, and returns it to the box. She repeats this process 40 times. Her results are shown in the table below.

<table>
<thead>
<tr>
<th>Animal</th>
<th>Panda</th>
<th>Giraffe</th>
<th>Lion</th>
<th>Elephant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>12</td>
<td>6</td>
<td>5</td>
<td>17</td>
</tr>
</tbody>
</table>

Based on these results, estimate the number of each animal in the box.

Panda: ____________________________  Giraffe: ____________________________
Lion: ____________________________  Elephant: ____________________________

Objective: Understand how to generate sample survey data by using simulations.

Design a Simulation to Model Results

Simulation: a method used to model random events; allows researchers to model real world phenomena

Examples

58. Airline on-time arrival statistics say the Air Nice flight from Philadelphia to Atlanta is on-time 70% of the time. Simulate the arrival time for an Air Nice flight from Philadelphia to Atlanta for 40 flights using a random number generator. Using the digits 0-9, how would you design the simulation?

Possible Outcomes: ____________________________  ____________________________
Data Says: ____________________________  ____________________________
Digit Representation: ____________________________  ____________________________
Results of Simulation: ____________________________  ____________________________

59. Recent polls suggest Emmett Palmer, a candidate for Senate, is leading the race with 55% of votes saying they would vote for him. Simulate the vote for 10 voters. Using two-digit numbers with the numbers 1-20, how would you design this simulation?

Possible Outcomes: ____________________________  ____________________________
Data Says: ____________________________  ____________________________
Digit Representation: ____________________________  ____________________________
Results of Simulation: ____________________________  ____________________________
Additional Example:

60. A researcher has previous research that suggests 60% of birds have built a nest in the same place at least twice. How would you simulate this data using the digits 0-9 for a trial of 50 birds?

Objective: Calculate a margin of error and use it to develop a confidence interval.

Calculating Margin of Error and Confidence Interval

Formulas for Margin of Error:

- Given no Prior Knowledge
  \[ MOE = \frac{1}{\sqrt{n}} \]

- Given a Sample Proportion
  \[ MOE = z^* \sqrt{\frac{\text{proportion} (1 - \text{proportion})}{n}} \]

- Given a Sample Mean
  \[ MOE = z^* \frac{\text{standard deviation}}{\sqrt{n}} \]

<table>
<thead>
<tr>
<th>Percentage Confidence</th>
<th>z*-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>1.28</td>
</tr>
<tr>
<td>90</td>
<td>1.545</td>
</tr>
<tr>
<td><strong>95</strong></td>
<td><strong>1.96</strong></td>
</tr>
<tr>
<td>98</td>
<td>2.33</td>
</tr>
<tr>
<td>99</td>
<td>2.58</td>
</tr>
</tbody>
</table>
Interpreting a Margin of Error

A survey is taken to determine whether the students in your high school would like to have soda served during lunch. A random sample of 200 students was surveyed. The result of the survey indicated 80% were in favor of serving soda during lunch.

- Calculate the margin of error for the following confidence levels.

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>Margin of Error</th>
<th>Confidence Interval</th>
<th>% of sample proportions within the CI</th>
<th>% of sample proportion outside the CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Make a statement about the results of your survey for a 80% Confidence Interval.

What does a 80% Confidence Interval Mean?? If you generate many samples of the same size from a population, then the difference between the sample statistic and the true population parameter (usually mean) will be within the margin of error for about 80% of the sample statistics.

Can you find a margin of error for a population?___________________________________________________

Simulate a population proportion of 70%. Generate a random sample of proportions.

The teacher will generate a list of random number based on the 70% population proportion and assign each student a sample number.

Spreadsheet Page/Formula Box/Menu/Data/Random/Binomial  randbin (200, 0.7, # students)

1. Record Your Assigned Sample:___________________________
2. Calculate Your Percent:_________________________________
3. Calculate your margin of error based on an 80% confidence level:__________________________________________
4. State your Confidence Interval:_____________________________________________________________________
5. Plot your confidence interval on the number line on the white board.

Questions

1. How many samples captured the proportion of 70%?_______________ What percent is this?_______________
2. How many samples did not capture the proportion of 70%?___________ What percent is this?_______________
3. We would expect with an 80% confidence level approximately 80% of our samples would capture the true proportion. How did our results compare?________________________________________________________
4. Why were our results not exactly 80%?________________________________________________________________
Margin of Error Questions:
How does changing the sample size change the margin of error? Justify your answer.____________________________________

__________________________________________________________________________________________
__________________________________________________________________________________________
How does changing the confidence level (\( z^* \)) change the margin of error? Justify your answer.________________
__________________________________________________________________________________________
__________________________________________________________________________________________
What is the only way to eliminate the margin of error?
__________________________________________________________________________________________
List ways of decreasing the margin of error:
_________________________________ _________________ __________ ________________________

Examples:
61. A magazine article reports that 58% of adults with cellphones have read their email during a business meeting with a margin of error of 3%. What does the margin of error suggest about adult usage of cellphones in business meetings? (Assume a 95% confidence level)
Math: ______________________________________ Interpretation: _________________________________
__________________________________________________________________________________________
Can we determine the approximate sample size? Explain.____________________________________________
__________________________________________________________________________________________
62. A community action agency reports a policeman issues an average of 25 speeding tickets a month with a margin of error of 6 tickets. What does the margin of error suggest about policemen issuing speeding tickets?
Math: ______________________________________ Interpretation: _________________________________
__________________________________________________________________________________________
63. A professional baseball watchdog commission reports between 60% and 68% of players spend their offseason working with youth baseball camps. What is the margin of error? What is the sample proportion?
Proportion:___________________________________ Margin of Error:________________________________

Additional Examples:
64. A newspaper reports that John Q. Public is leading the race with 52% of the vote with a margin of error of \( \pm 3\% \). What does the margin of error suggest about the outcome of the race?
Math: ______________________________________ Interpretation:_________________________________
__________________________________________________________________________________________
65. A psychologist reports that families going to therapy spend between 15 and 19 hours in group sessions. What is the margin of error? What is the sample mean?
Mean:_______________________________________ Margin of Error:________________________________
Calculating Margin of Error for a Sample Mean

Statistical Question: What is the average weight of male students at ABC University?

- 1000 randomly selected male students were surveyed. The sample mean is calculated to be 180 pounds and the standard deviation 30 pounds.
- You want to represent the findings as a 95% confidence interval.
  o Calculate the margin of error:
    \[
    ME = z^* \cdot \frac{\sigma}{\sqrt{n}}
    \]
    
    \(z^*\): confidence coefficient
    \(\sigma\): standard deviation
    \(n\): sample size
  
  o State the confidence interval
    95% Confidence Interval: __________________________

Calculating Margin of Error for a Sample Proportion

Statistical Question: What percent of the population think the president is doing a good job?

- The Gallup Organization’s latest poll sampled 1000 people from the United States, and the results show that 520 people think the president is doing a good job, compared to 480 who don’t think so.
- You want to represent the findings as a 95% confidence interval.

\[
ME = z^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}
\]

\(\hat{p}\): sample proportion
\(z^*\):confidence coefficient
\(n\): sample size

o Calculate the margin of error.
  ME=____________________________

o State the confidence interval
  95% Confidence Interval: __________________________

Question: How does increasing the sample size affect the margin of error? Explain.

____________________________________________________________________________________________________
____________________________________________________________________________________________________
____________________________________________________________________________________________________
**Additional Examples:**

66. A simple random sample of 400 high school seniors gained an average of 22 points on their second attempt at the SAT Mathematics exam. Assuming the change in score has a Normal distribution with a standard deviation of 50.
   a. Calculate the margin of error for 95% confidence.

   b. Find a 95% confidence interval.

   c. If the sample was of 100 high school seniors, what would be the margin of error for 95% confidence?

   d. How does the margin of error for a 400 person sample compare to a 100 person sample?

67. If 64% of a sample of 550 people leaving a shopping mall claims to have spent over $25, determine a 99% confidence interval estimate for the proportion of shopping mall customers who spend over $25. Interpret your interval.

68. In a random sample of machine parts, 18 out of 225 were found to have been damaged in shipment. Establish a 95% confidence interval estimate for the proportion of machine parts that are damaged in shipment. Interpret your interval.

69. A telephone survey of 1000 adults was taken shortly after the U.S. began bombing Iraq. If 832 voiced their support for this action. Create a 99% confidence interval and interpret the interval.

70. An assembly line does a quality check by sampling 50 of its products. It finds that 16% of the parts are defective.
   a. Create a 95% confidence interval for the percent of defective parts for the company and interpret this interval.

   b. If we decreased the confidence level to 90% what would happen to:
      i. the critical value?
      ii. the margin of error?
      iii. the confidence interval?
c. If the sample size were increased to 200, the same sample proportion were found, and we did a 95% confidence interval; what would happen to:
   i. the critical value?
   ii. the margin of error?
   iii. the confidence interval?

71. A nationwide poll was taken of 1432 teenagers (ages 13-18). 630 of them said they have a TV in their room.
   a. Create a 90% confidence interval for the proportion of all teenagers who have a TV in their room and interpret it.
   
   b. What does “90% confidence” mean in this context?
   
   c. If we increased the confidence level to 99% what would happen to:
      i. the critical value?
      ii. the margin of error?
      iii. the confidence interval?
   
   d. If the sample size were changed to 950, the same sample proportion were found, and we did a 90% confidence interval; what would happen to:
      i. the critical value?
      ii. the margin of error?
      iii. the confidence interval?

72. Suppose a 90% confidence interval is stated as (0.3011, 0.4189).
   a. What is the sample proportion from this sample?
   
   b. What is the margin of error?

73. In a simple random sample of 150 households, the sample mean number of personal computers was 1.32. The population standard deviation is 0.41. Construct a 95% confidence interval for the mean number of personal computers.

74. Joe and Sally are going to construct confidence intervals from the sample simple random sample. Joe’s confidence interval will have level 90% and Sally’s will have level 95%.
   a. Which confidence interval will have the larger margin of error?
   
   b. Which confidence interval will be more likely to cover the population mean?

75. A dean at a certain college looked up the GPA for a random sample of 85 students. The sample mean GPA was 2.82, and a 95% confidence interval for the mean GPA of all students in the college was \( 2.76 < \mu < 2.88 \). True or false, and explain.
   a. We are 95% confident that the mean GPA of all students in the college is between 2.76 and 2.88.

   b. We are 95% confident that the mean GPA of all students **in the sample** is between 2.76 and 2.88.

   c. The probability is 0.95 that the mean GPA of all students in the college is between 2.76 and 2.88.

   d. 95% of the students in the sample had a GPA between 2.76 and 2.88.
Determine if Results are Typical Using Simulation

A state is considering increasing the number of courses needed for high school graduation to 4 math and 4 science course. They take a random sample of 200 students and find that 46% of the students sampled favor the change. The state would like for at least half of all students in the state to support the change. Is it possible that 50% of all state students are in favor of the change?

- Remember: A sample statistic does not give you the exact population parameter!

- Statistical Question: Is it possible that the true population parameter is 50% or greater even though the sample statistic is 46%

When determining if a model is consistent with results:

1. **Hypothesis:**
   
2. **Simulate or Sample:**
   
3. **Repeated Trials:**
   
4. **Analyze Response Variable:**
   
5. **Decision about Hypothesis:**

Hypothesis:______________________________________________________________

Design Simulation:________________________________________________________

Collect data from the simulation:

Analyze:

Calculate the probability of getting 46%:_______________________________________________

Conclusion:__________________________________________________________

(The value used to reject the hypothesis varies. In this case use a value of 5% or smaller)

**Questions:**

- Why would a simulation be used instead of collecting more samples?________________________

Additional Example: ADD
Develop a Margin of Error using Simulation

A sample of 100 male junior students were asked their height. The results were normally distributed with a mean of 65”. The shortest height was 59” and the tallest 77”. Can you use this information to approximate a standard deviation? Explain.

Consider the same question from before:
A state is considering increasing the number of courses needed for high school graduation to 4 math and 4 science courses. They want to know if 50% of students will support the change. They take a random sample of 200 students and find that 46% of the students sampled favor the change.

What is the 95% confidence interval for the population proportion? Is it plausible that 50% of all students favor the change?

Run a simulation for the hypothesis: 50% of students are in favor of the change.

Calculate the sample proportion: _____________________________________________________________

Estimate the sample standard deviation: ______________________________________________________

State the 95% confidence interval: __________________________________________________________

Conclusion: ________________________________________________________________________________

There is a 95% likelihood that a given sample statistic is within 2 standard deviations of the population parameter.

There is a 95% likelihood that a population parameter is within 2 standard deviations of the sample statistic.

Sample Proportion ± 2(standard deviation)
Additional Example
A state is considering increasing the number of courses needed for high school graduation to 4 math and 4 science courses. They want to know if 60% of students will support the change. They take a random sample of 200 students and find that 46% of the students sampled favor the change.

Hypothesis: ____________________________________________________________

Simulation: _____________________________________________________________

Sample Proportion: ______________________________________________________

Sample Standard Deviation: _______________________________________________

State the 95% Confidence Interval: _________________________________________

Conclusion: _____________________________________________________________

Compare Two Treatments

- Effectiveness of attendance intervention program
- 60 students
- 30 students: program and teacher encouragement
- 30 students: teacher encouragement
- Track attendance and grades for two years

Comparison

Median: ________________________________________________________________

IQR: _________________________________________________________________

Variability: ____________________________________________________________

Shape: _______________________________________________________________
• Effectiveness of a new fertilizer
• 20 peanut fields of approximately the same size
• 10 fields: new fertilizer
• 10 fields: fertilizer used in the past
• Track yield of each field

**Comparison**

Median: _________________________________________________

IQR: ____________________________________________________

Variability: _______________________________________________

Shape: __________________________________________________

• Effectiveness of using verbal and visual commands on training dogs
• 50 dogs
• 25 receive verbal and visual commands
• 25 receive verbal commands only
• Records numbers of days until dog performs command on call

**Comparison**

Median: _________________________________________________

IQR: ____________________________________________________

Variability: _______________________________________________

Shape: __________________________________________________
A researcher is interested in determining if a new salt substitute lowers blood pressure more than the leading salt substitute. Sixty people identified as having high blood pressure are recruited to participate. The researcher randomly assigns 30 people to use the new salt substitute and 30 people to use the leading salt substitute. The box plots show each person’s average systolic blood pressure over the course of the experiment. Can we make a decision as to which salt substitute lowers blood pressure?

**Comparison**

Median: ________________________________________

IQR: __________________________________________

Variability: ______________________________________

Shape: __________________________________________

Use simulation to determine if differences are significant???

**Evaluate Reports**

Have students bring in statistical information from newspapers, magazine articles, internet, etc. They will be expected to analyze the accuracy of the information.
A dean at a certain college looked up the GPA for a random sample of 85 students. The sample mean GPA was 2.82, and a 95% confidence interval for the mean GPA of all students in the college was $2.76 < \mu < 2.88$. True or false, and explain:

a. We are 95% confident that the mean GPA of all students in the college is between 2.76 and 2.88.
   **TRUE.** We are using a method that is accurate about 95% of the time (in repeated sampling) so we are 95% confident in the outcome.

b. We are 95% confident that the mean GPA of all students in the sample is between 2.76 and 2.88.
   **FALSE.** The mean GPA of the students in the sample is ALWAYS in the CI – it gives us the CENTER of the interval!

c. The probability is 0.95 that the mean GPA of all students in the college is between 2.76 and 2.88.
   **FALSE.** This is an incorrect interpretation of what the CI means. Suppose the true mean GPA is 2.6. That is NOT in the interval, is it? So how could I say there is a probability of 0.95 that 2.6 is between 2.76 and 2.88? I can’t!

d. 95% of the students in the sample had a GPA between 2.76 and 2.88.
   **FALSE.** GPAs for the students in the sample could be all over the place! The individual GPAs do not have to fall in the interval!