

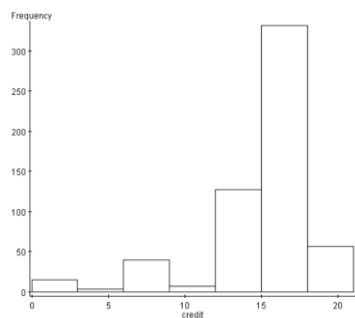
Exam 1 Review Problems

The length of this practice does not reflect the length of the exam. The problems below are to provide you with practice for exam.

Multiple Choice:

1. Undercoverage occurs when the sampling frame
 - a. matches the population of interest.
 - b. does not match the population of interest.
 - c. matches the sample of interest.
 - d. does not match the sample of interest.
2. Adding 10 to every data point will
 - a. not increase the standard deviation but will increase the mean.
 - b. increase the standard deviation but will not increase the mean.
 - c. will increase both the mean and standard deviation.
 - d. will not increase either the mean or standard deviation.
3. The 'box' in a boxplot
 - a. is always one standard deviation wide.
 - b. indicates the start and end of the middle 50% of the data.
 - c. is always 2 standard deviations wide.
 - d. describes the minimum and maximum values from the data.
4. A normal distribution is
 - a. Skewed right.
 - b. Symmetric.
 - c. Skewed left.
 - d. Centered at 0.
5. The state legislature of Kansas is interested in finding out about the opinion of registered voters throughout the state. The legislative staff has created a plan to randomly select voters from throughout the state. The staff combined the voter registration lists from all the counties in the state to get a single list of all registered voters. Using the list of all voters in the state the staff divided the list into rural and urban voters. The staff then used a random number generator to select 500 rural and 500 urban voters. This sample would best be described as:
 - a. A simple random sample of 500 registered voters.
 - b. A simple random sample of 1000 registered voters.
 - c. A cluster sample of registered voters using counties as clusters.
 - d. A stratified sample of registered voters using rural/urban as strata.

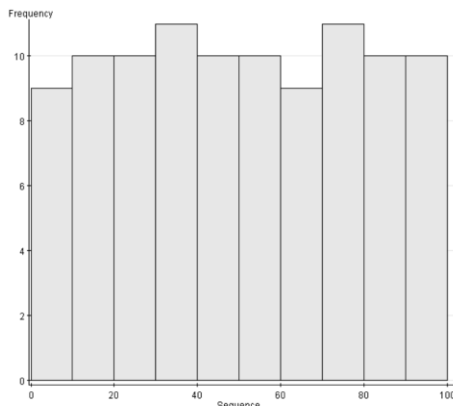
6. As part of a class project, a college student wanted to estimate the proportion of students at his college that owned a red car. To estimate this proportion he went to a large student parking lot during the middle of a class day. He found that 325 out of 1037 cars (31%) were red. What type of sampling method was used here?
- A volunteer response sample.
 - A stratified random sample with color of car as the strata.
 - A simple random sample.
 - A convenience sample.
7. Given that a variable X is normally distributed with mean μ and standard deviation σ , why might we convert to a z-score?
- to find the distribution of X .
 - to easily compare it to other observations.
 - to tell whether the data is skewed.
 - to determine the mean and variance of the distribution.
8. Suppose we want to know the average age of all students taking ST 311 this semester. We take a random sample of 15 students and get a sample mean of 21.4. What is the parameter of interest?
- the average age of the 15 students measured.
 - the sample mean of 21.4.
 - the average age of all students who have ever taken ST 311 at NCSU.
 - the average age of all students taking ST 311 during this semester.
9. The histogram below gives information about the number of credits taken by a sample of students. Which statement is correct about this histogram?



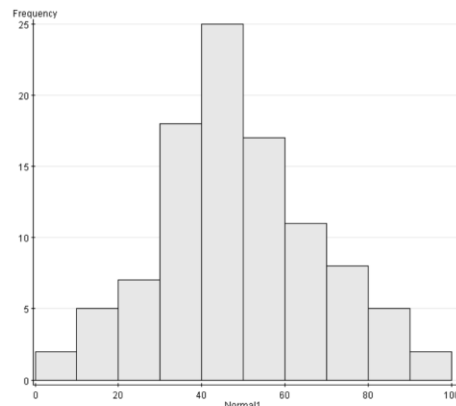
- The histogram is skewed left, so the sample mean is less than the sample median.
- The histogram is skewed left, so the sample mean is greater than the sample median.
- The histogram is skewed right, so the sample mean is less than the sample median.
- The histogram is skewed right, so the sample mean is greater than the sample median.

10. A pharmaceutical company did research on candidates for a drug trial. They recorded each patients weight (in lbs), blood pressure (in mmHg), race (White, Black, Asian, Other), zip code (27511 or 27606), and level of breathing difficulty (None, Some, or Extreme). Which of the following is true about the variables recorded?
- Race and zip code are the only categorical variables.
 - Race, zip code and level of breathing difficulty are the only categorical variables.
 - Race and level of breathing difficulty are the only categorical variables.
 - Zip code and level of breathing difficulty are the only categorical variables.

11. Consider the two distributions below. What is true about the standard deviation of the distributions?



(a) Distribution A



(b) Distribution B

- Distribution A has a smaller standard deviation than distribution B.
 - Distribution B has a smaller standard deviation than distribution A.
 - The standard deviations of the two distributions are approximately the same.
 - There is not enough information given to answer the question.
12. In a simple random sample, the problem of non-response bias occurs when:
- The sampling frame does not contain all individuals in the population.
 - Some individuals that are randomly selected refuse to participate.
 - Some individuals who answer the questions do not answer the question truthfully.
 - Both B and C but not A.
13. A certain firm has 500 employees. An auditor wishes to estimate the average number of days of sick leave taken per employee over the past quarter. The firm has eight divisions, with varying numbers of employees per division. The auditor decides to randomly sample 3 divisions and finds the number of sick days each employee within these divisions take. This sample is best described as:
- A stratified sample using the divisions as strata.
 - A cluster sample using the divisions as clusters.
 - A simple random sample of 500 employees.
 - A simple random sample of an unknown number of employees.

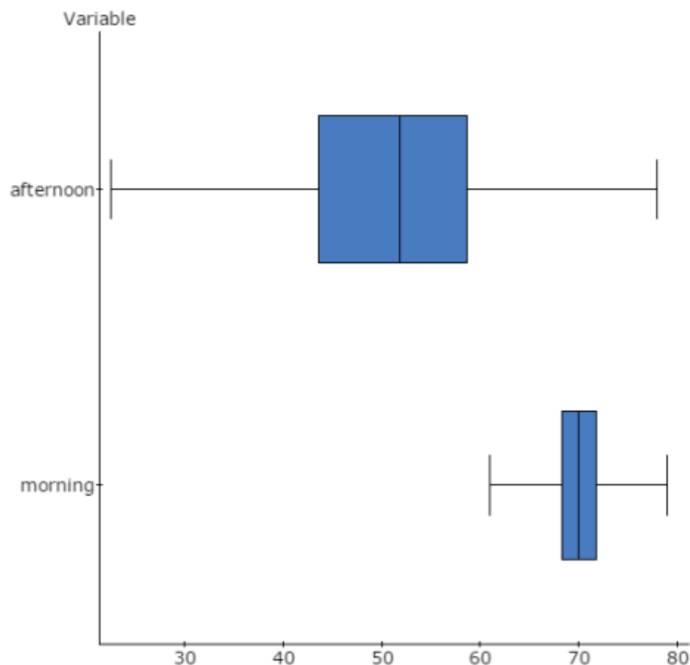
Use the information below for questions 14 through 16.

A local corporation was interested in the purchasing habits of Wake County residents. One of the things they were interested in was the proportion of residents that did a major home improvement project in the last year. The corporation used a list of all households in the county to take a simple random sample. The researcher then visited each of the sampled homes and found that 21 of 100 homes had completed a major home improvement project.

14. What is the population of interest in this situation?
 - a. All Americans.
 - b. All Wake County residents.
 - c. All Wake County residents who have completed a major home improvement project in the last year.
 - d. The 100 homes that were sampled.
15. What is the parameter of interest in this situation?
 - a. The proportion of all Wake County residents who did a major home improvement project in the last year.
 - b. 21%.
 - c. The 100 homes that were sampled.
 - d. All Wake County residents.
16. What is the sampling frame in this situation?
 - a. The 100 homes that were sampled.
 - b. Major home improvement projects.
 - c. The local corporation that conducted the sample.
 - d. The list of all households in the county.
17. A local politician is interested in the public opinion regarding a tax increase to fund improvements to the local public transportation services in Wake County. She decided to take a random sample of 15 voting precincts and polls all the registered voters within those precincts. (Assume that every voter in each of the 15 precincts responds to the poll.)
This is an example of a:
 - a. Systematic random sample.
 - b. Multi-stage random sample.
 - c. Stratified random sample.
 - d. Cluster random sample.
18. Which of the following is not a characteristic of the standard normal distribution?
 - a. It is centered at 0.
 - b. The standard deviation is 1.
 - c. It is symmetric.
 - d. It is skewed to the right.

Use the information and figure below for questions 19 and 20.

A college professor has two large lectures (morning and afternoon) each with 250 students. She recorded the scores for her midterm exam from both sections. The resulting scores are presented in these boxplots.



19. From the boxplot we can see that the shape of the histogram of the afternoon section scores has which of the following features:
- It is right skewed.
 - It is bimodal.
 - It will have several outliers.
 - None of the above.
20. What is the interquartile range for the morning section scores?
- 20.
 - 68 to 72.
 - 60 to 80.
 - 4.
21. Adam is conducting a study of the height of entering freshmen at NCSU. The data he collects will be
- qualitative/categorical
 - quantitative and discrete
 - quantitative and continuous
 - not enough information was given

22. Adult Great Basin rattlesnakes have a mean length of 36 inches and a standard deviation of 5 inches; the length of adult Southern Pacific rattlesnakes is also 36 inches on average, but with a standard deviation of 7 inches. Assume it is known that the length of both subspecies obeys a normal distribution.

Suppose Jafar is a Great Basin rattlesnake and Nagini is a Southern Pacific rattlesnake. Jafar is longer than exactly 75% of all Great Basin rattlesnakes; Nagini is longer than exactly 75% of all Southern Pacific rattlesnakes. Which snake is longer?

- a. Jafar.
- b. Nagini.
- c. They are of equal length.
- d. Not enough information is provided; it depends.

Use the following information to answer questions 23 and 24.

A graduate student is interested in estimating the percentage of North Carolina State University students that have an Amazon Prime account. The graduate student believes that this percentage is related to the age of the students.

23. In this situation, what is the population of interest?

- a. The list of all students at North Carolina State University.
- b. All students at North Carolina State University.
- c. All students at North Carolina State University with an Amazon Prime account.
- d. The proportion of all students at North Carolina State University that have an Amazon Prime account.

24. In this situation, what is the parameter of interest?

- a. The list of all students at North Carolina State University.
- b. All students at North Carolina State University.
- c. All students at North Carolina State University with an Amazon Prime account.
- d. The proportion of all students at North Carolina State University that have an Amazon Prime account.

25. The residents of Wake County are the subjects of a study on commuter time. The researchers have divided the county residents into groups by imposing a grid over a map of the county. Fifty households from each square in this grid are randomly selected and surveyed. Identify the sampling technique.

- a. Simple random
- b. Stratified
- c. Systematic
- d. Cluster

26. Victoria is taking the same statistics class as Sam. She scored an 85 on the test. The mean score for the statistics test was 63 with a standard deviation of 7.0. What is the z-score of Victoria's test grade (rounded to 2 decimal places)?
- 3.00
 - 3.14
 - 3.14
 - 3.15
27. An Internet site compares the strokes per round of two professional golfers. Which golfer is more consistent: Player A with $\mu = 71.5$ strokes and $\sigma = 2.3$ strokes, or Player B with $\mu = 70.1$ strokes and $\sigma = 1.2$ strokes?
- Player A because his mean and standard deviation are both higher than Player B's.
 - Player B because his mean is lower than Player A's.
 - Player B because his standard deviation is lower than Player A's.
 - Player A because his mean is higher than Player B's, regardless of what the standard deviation is.
28. Sam is taking biology and statistics. He recently had a test in both classes. He scored a 73 on the statistics test and a 28 on the biology test. The mean score for the statistics test was 63 with a standard deviation of 7.0 and the mean score on the biology test was 23 with a standard deviation of 3.9. On which test did Sam score relatively better?
- The statistics test
 - The biology test
 - Neither, his relative score was the same on the 2 tests.
 - Not enough information.
29. Draw a picture to represent $P(Z < 1.49)$, then choose the from the probabilities below the one that best matches. Do not use StatCrunch to find the answer.
- 0.93
 - 0.08
 - 0.68
 - 0.31
30. Draw a picture to represent $P(-2.14 < Z < 0.95)$, then choose the from the probabilities below the one that best matches. Do not use StatCrunch to find the answer.
- 0.12
 - 0.07
 - 0.19
 - 0.81

31. IQ scores are normally distributed with a mean of 100 and a standard deviation of 15. Estimate the probability a randomly selected person has an IQ score greater than 120.
- 0.90
 - 0.09
 - 0.63
 - 0.37
32. Forty-three percent of marriages end in divorce. You randomly select 15 married couples. Find the mean number of marriages that will end in divorce.
- 2.15
 - 8.55
 - 6.45
 - 2.85
33. Forty-three percent of marriages end in divorce. You randomly select 15 married couples. Find the standard deviation of the number of marriages that will end in divorce.
- 0.245
 - 0.495
 - 1.92
 - 3.68
34. Forty-three percent of marriages end in divorce. You randomly select 15 married couples. Would it be unusual for 5 of these marriages to end in divorce?
- Yes – because 5 is unusually low.
 - Yes – because 5 is below the mean.
 - No – because $\frac{5}{15} = 0.33$, which is greater than 5%.
 - No – because 5 is not far enough from the mean.
35. At a large university it is known that 25% of the students live on campus. The director of student life is going to take a random sample of 500 students and see the proportion of students who live on campus. Which of the following about the distribution of the sample proportion is NOT TRUE?
- The mean of the sample proportion is 0.25.
 - The distribution of the sample proportion is shaped like a normal distribution.
 - The standard deviation of the sample proportion depends on the population proportion.
 - As sample size increases the standard deviation of the sample proportion increases.
36. At NC State University, 16.4% of the undergraduate classes have more than 50 students. If a random sample of 200 undergraduate classes was taken, which of the following would accurately describe the sampling distribution?
- The sampling distribution will be skewed right.
 - The mean of the sampling distribution will be 40%.
 - The standard deviation of the sampling distribution will be 0.0262.
 - The standard deviation of the sampling distribution will be 0.0007.

Use the following information to answer questions 37 through 39.

A marketing director wants increase sales by using one of three new TV advertisements. The director decides to test each new advertisement in a variety of residential areas. Six rural areas, six suburban areas, and six urban areas are randomly chosen to be in the study. Each advertisement is broadcast in two of each type of residential area. Sales figures are measured for each area.

37. What type of study was performed?
- Experiment.
 - Observational Study
38. What is the response variable?
- Residential Area Type (rural, suburban, urban)
 - Advertisement (1, 2, or 3)
 - Sales
 - None of the above
39. What type of design was used?
- Completely Randomized
 - Matched Pairs
 - Stratified Sampling
 - Randomized Block
 - Cluster Sampling

Use this description to answer questions 40 through 42:

Does the type of movie children are watching make a difference in the amount of snacks they eat? A group of 50 children were randomly assigned to watch either a cartoon or a live action musical (25 to each). Crackers were available in a bowl, and the investigators compared the number of crackers eaten by children while watching the different kinds of movies.

40. In this study the response variable is:
- The amount of crackers eaten.
 - The children.
 - Does the type of movie make a difference in the amount of snacks eaten?
 - The type of movie watched.
41. This study is best described as:
- A placebo controlled experiment.
 - A matched pairs experiment.
 - A randomized observational study.
 - A completely randomized design.

42. In the study described above, the cartoon movie was randomly chosen to be shown at 8 AM (right after the children had breakfast) and the live action movie at 11 AM (right before the children had lunch). It was found that during the movie shown at 11 AM, more crackers were eaten than during the movie shown at 8 AM. The investigators concluded that the different types of movies had an effect on appetite. The results cannot be trusted because
- the study was not double blind. Neither the investigators nor the children should have been aware of which movie was being shown.
 - the investigators were biased. They knew beforehand what they hoped the study would show.
 - the time the movie was shown is a lurking variable.
 - the investigators should have used several bowls, with crackers randomly placed in each.
43. The most important advantage of experiments over observational studies is that
- experiments are usually easier to carry out.
 - experiments can give better evidence of causation.
 - the placebo effect cannot happen in experiments.
 - an observational study cannot have a response variable.
44. According to nces.ed.gov, 5.3% of American students drop out of high school. A random sample of 10 American students is chosen and 2 dropped out. Find \hat{p} .
- 0.053
 - 5.3
 - 0.53
 - 0.20
 - 2.65
45. According to nces.ed.gov, 5.3% of American students drop out of high school. A random sample of 10 American students is chosen. Find the mean of the sampling distribution of the sample proportion for $n = 10$.
- 0.053
 - 0.53
 - 5.3
 - 0.0167
 - 1.167
46. According to nces.ed.gov, 5.3% of American students drop out of high school. A random sample of 10 American students is chosen. Find the standard deviation of the sampling distribution of the sample proportion for $n = 10$.
- 0.0708
 - 0.7085
 - 0.0502
 - 9.47
 - 5.3

47. According to [nces.ed.gov](https://nces.ed.gov/ipeds/data/ipedsreports/2016/ipeds2016010101.pdf), 5.3% of American students drop out of high school. A random sample of 10 American students is chosen. What do we know about the shape of the distribution of the sample proportion for $n = 10$?
- It is approximately Normal.
 - It is skewed.
 - It is not skewed, but it is also not Normal.
 - We do not know about the shape because we do not know the shape of the parent population.
48. Arjun would like to learn about the average weight of an American male. He plans to find 16 volunteers and ask their weights. Suppose the true average weight of American males is 180 pounds and the standard deviation is 30 pounds. Which of the following about the sampling distribution is true?
- It is 50% likely that 8 volunteers will weigh more than 180 pounds.
 - It is 50% likely that the average weight of the 16 volunteers will be more than 180 pounds.
 - The average weight of the 16 volunteers follows a normal distribution with mean 180 pounds and standard deviation 7.5 pounds.
 - We cannot conduct inference on the sample distribution because the sample size is too small and we don't know whether the population distribution is normal.
49. In a recent study of incomes in Wake county in North Carolina, it was found that the distribution of family incomes is skewed to the right. What can we say about the relationship between mean and median.
- mean and median are same.
 - mean is greater than median.
 - mean is less than median.
 - the median must be thrice the mean.
50. The expected value of a random variable is the
- value that has the highest probability of occurring.
 - mean value over an infinite number of observations of the variable.
 - largest value that will ever occur.
 - most common value over an infinite number of observations of the variable.

Use the following information for questions 51 through 53.

A biologist would like to analyze the nitrogen content of two specific varieties of red clover plants (Kenland and Marathon) exposed to three strains of Rhizobium: 3D0K1, 3DOK5, and 3DOK7. In preparation for this study, the biologist planted 36 samples of one variety of red clover (Kenland) and 42 samples of the other (Marathon). Assume the biologist has asked you, the researcher, to analyze the data. Use this information to answer the following three questions.

51. What are the treatments?

- a. The measured nitrogen content values for all 78 samples.
- b. The three strains of Rhizobium, 3D0K1, 3DOK5, and 3DOK7.
- c. The two red clover varieties, Kenland and Marathon.
- d. The two samples of red clover.

52. What would you use for the explanatory and response variables?

- a. The explanatory variable is the nitrogen content. The response variable is the Rhizobium, with the three levels 3D0K1, 3DOK5, and 3DOK7.
- b. The explanatory variable is the Rhizobium, with the three levels 3D0K1, 3DOK5, and 3DOK7. The response variable is the variety of clover, with the two levels Kenland and Marathon.
- c. The explanatory variable is the variety of clover, with the two levels Kenland and Marathon. The response variable is the Rhizobium, with the three levels 3D0K1, 3DOK5, and 3DOK7.
- d. The explanatory variable is the Rhizobium, with the three levels 3D0K1, 3DOK5, and 3DOK7. The response variable is the nitrogen content.

53. The biologist wants your recommendation on how to randomize the treatment to the red clover samples, but cautions that the two varieties may respond differently to the treatment. What would you suggest?

- a. A completely randomized design.
- b. A randomized block design, blocking by strain of Rhizobium.
- c. A randomized block design, blocking by variety of clover.
- d. A matched pairs design by matching the variety of clover.

Free Response:

1. The probability is $p = 0.80$ that a patient with a certain disease will be successfully treated with a new medical treatment. Suppose that the treatment is used on 40 patients. What is the expected number of patients who are successfully treated?
2. For each description below, determine if the scenario can be modeled with a binomial distribution, a normal distribution, or neither. Write the letter of the best choice in the blank provided.
 - A. Normal
 - B. Binomial
 - C. Neither

___ Shishito peppers are usually considered to be mild, however, there is a 1 in 10 chance that one is a hot pepper. Eddie is throwing a party and buys a total of 156 peppers. Assuming the peppers he buys are random, what is the distribution of the number of these peppers which are hot?

___ A game requires the player to flip a fair coin repeatedly until she gets three tails in a row. The player earns points equal to the total number of coin flips that it takes to get three tails in a row. What is the distribution of the number of points the player gets?

___ A report indicated that 64% of adults in the U.S. own their home. A group of 25 U.S. adults is randomly chosen. What is the distribution of the number of these adults who own their home?

___ A hospital reports that 26% of their emergency department visits occur between 9pm and 11pm. On average, there are 16 visits during this time period, per night. What is the distribution of the number of emergency department visits between 9pm and 11pm?

___ The amount of time it takes to repair brakes on a car follows a distribution that is skewed right. A manager takes a random sample of 40 brake repair jobs and finds the sample mean of the repair times. What is the sampling distribution of the sample mean?
3. A sociology instructor at a local community college is interested in the drinking habits of American college students. The instructor randomly selects five students from the 8am freshmen sociology class and conducts one-on-one interviews with each of them. Name at least two flaws in the design of this study. Explain why they are flaws.

4. Dorothy wants to know if a new brand of nail polish lasts longer without chipping than her old brand. To answer her question, she enlists 5 of her friends to help her. She has each of her friends paint the old brand on the left hand and the new brand on the right hand. At the end of a week, the nail on each finger is compared to the same nail on the opposite hand (e.g. left thumbnail to right thumb nail) to determine which nail has more chipping.
 - a. What data collection method was used in this study (observational study, experiment)?
 - b. What type of design was used (completely randomized design, randomized block design, or matched pairs?) Justify your answer.
 - c. Dorothy determines that the new brand of nail polish chips at a similar rate to the old brand of nail polish. Some of her friends disagree. What is one thing Dorothy could have done differently to make her findings more reliable?
5. Kim has job offers from two schools in Wake County and she is trying to decide which to accept. A private school and a public school each offered her \$40,000. The average salary for a public school teacher in Wake County is \$52,344 with a standard deviation of \$5942. The average salary of a private school teacher in Wake County is \$49,745 with a standard deviation of \$6227.
 - a. Which school is offering her the most money relative to the other teachers in that type of school? Justify your answers mathematically.
 - b. Based on your calculations in part A, would either of these salaries be considered unusual? Explain.
6. American Doberman Pinscher dogs (male) are known to have an average weight of 86lbs. A research group is trying to compare this to the weight of European Doberman Pinscher male dogs. They randomly select 20 dogs from all countries in Europe allowing Dobermans and record the weights.
 - a. Identify the population of interest in the group's study.
 - b. Identify the variable of interest in the group's study.
 - c. Was the data collected qualitative or quantitative? Explain your choice.
 - d. What type of study was performed (observational, experiment, simulation, or survey)? Explain your choice.

7. A survey indicates that the time spent for one supermarket trip is normally distributed with mean of 45 minutes and a standard deviation of 12 minutes.
- Draw the distribution for time spent in a supermarket trip. Label the axis appropriately.
 - Find the z-score that corresponds to a randomly selected shopper who spends 30 minutes at the market.
 - Shade the area under the curve that represents the probability that a randomly selected shopper spends less than 30 minutes at the market.
 - True or False: There is about a 10% chance that a random shopper will spend less than 30 minutes in the supermarket.
 - We want to know what number of minutes shopping separates the 15% of shoppers who take the longest in the store from the quickest 85% of shoppers. Draw the distribution and shade the area representing the 15%.
 - Which z-score below best estimates the distance from the mean to the 15% mark?
 - 2.11
 - 0.95
 - 1.04
 - 2.23
 - Use your chosen z-score to find the number of minutes (x-value).
8. A horticulturist conducts a study to determine the effect of chemical A on the yield (in kilograms per acre) of tomato plants. The horticulturist has 60 plots of land in six different counties. He randomly selects 4 from each county and plants tomatoes on those plots. He treats half of the crops in each county with chemical A and lets the other half of the crops grow naturally. At the end of the season, he compares the yield (in kg per acre) of the treated crops to the yield of the natural crops.
- Is this an experiment, an observational study, or a survey?
 - What method of sampling was used?
 - What randomization technique was used, randomized block design or complete randomization?
 - Determine the type of data collected: quantitative or qualitative?
 - Describe any possible confounding factors.

9. A department store is interested in finding out the optimal percent discount the store should offer during a sale to maximize profit. They choose 6 of the 24 closest zip codes and send a questionnaire to every home address in those 6 zip codes. The questionnaire asks residents the following question:

Which of the following sales discounts is the smallest amount that would bring you into the store?

- A. 10-15 % off
- B. 20-25% off
- C. 30-35% off

To get more responses, the store offers 50% off any one item for each resident who returns the questionnaire.

- a. Is this an experiment or observational study?
- b. What method of sampling was used?
- c. What group of people represents the population?
- d. What is the sample?
- e. Describe any problems/biases you find with this study.

10. The mean income per household in 2019 in North Carolina is \$57,341 per year with a standard deviation of \$8,008 per year. The mean income per household in 2019 in Alaska is \$75,463 per year with a standard deviation of \$4,632 per year. Assume the distribution of income values for each state is bell-shaped.

- a. A North Carolina household made \$38,322 in 2019. Is this value unusual?
- b. An Alaska resident made \$87,645 in 2019. Is this value unusual?
- c. Which income value is more unusual?

11. A company claims that 15% of its plain candies are orange. A random sample of 20 such candies is selected.

- a. What is the distribution of C = number of candies that are orange out of 20 candies?
- b. How many of the 20 candies do we expect to be orange?

12. Suppose the amount spent on rent (in dollars) for NCSU students per month is normally distributed with a mean of \$535 and a standard deviation of \$75. What rent would be considered unusually high?

13. The Canadian government is interested in studying tourism patterns in their country. Among other things, they wish to know the percentage of all Americans who have visited Canada in the past five years. Because they are researching multiple aspects of Canadian tourism simultaneously, they only have the budget to do a relatively small study. A junior data scientist has proposed two study methods:

- Method A: Conduct a study of Americans living in border states, e.g., Maine, Washington state, North Dakota, etc. It will allow them to contact a random sample of 1,200 Americans living in those states because they can conduct the study themselves.
- Method B: Hire a research agency to conduct a study of randomly selected Americans throughout the country. Because of the cost of hiring the research agency, they will only get a sample size of only 85.

For both methods, assume that they are able to get a list of all Americans and what states they live in. Do not worry about how they got that list or contacted them, and assume everyone responded and responded honestly to the survey. The tourism office only has the budget to conduct one of the two studies described above. Which would you choose, method A or method B? Why?

14. Carbon monoxide concentration in the atmosphere is believed to be symmetric with a mean of 82 parts per million (ppm) and a standard deviation of 7 ppm.

- Can we find the probability that the mean carbon monoxide concentration of 25 random samples is greater than 86 ppm? Explain why or why not.
- Fill in the probability statement below:

$$P(\bar{X} > 86) = P\left(Z > \quad\quad\quad\right) = P(Z > \quad\quad\quad)$$

- Is the probability that the mean carbon monoxide concentration of 25 random samples is greater than 86 ppm more or less than 5%? Explain how you know.

15. In a particular country, it is required that manhole covers be at least 23 inches in diameter so that workers can easily fit down through them. Men in this country are known to have shoulder widths that are normally distributed with a mean of 20.1 inches and a standard deviation of 1.2 inches, so roughly 99% of the population would fit.

The government is considering reducing the required minimum diameter to 21 inches. Is it reasonable to do so? Support your answer mathematically.

16. Can we find these probabilities? If so, set up the probability statement, draw the distribution, and shade the corresponding area under the curve. If not, explain why not.
- a. A study in the journal *Neurology* suggests that 29% of people have sleepwalked. Suppose this is accurate. If a random sample of 800 people is found, find the probability that 258 or more of them have sleepwalked.
 - b. A particular type of root canal therapy has a 95% success rate. If a random sample of 55 patients eligible for such therapy is selected, what is the probability that less than 50 are successful?
 - c. Only a small proportion of people ages 75 years and older carry credit card debt. Of those who do, the average debt was \$8100 in 2022, with a standard deviation of \$2104. If 100 randomly selected people aged 75 years or older with credit card debt are selected, find the probability that their mean debt is less than \$7500.