



Probability & Statistics Workbook

Sampling

TYPES OF STUDIES

■ 1. The following table shows the age and shoe size of six children. Does the data have a positive correlation, negative correlation, or no correlation?

| Age | Shoe size |
|-----|-----------|
| 3 | 7 |
| 3 | 6 |
| 5 | 9 |
| 6 | 12 |
| 6 | 11 |
| 7 | 13 |

■ 2. A class conducts a survey and finds that 75 % of the school spends 2 or more hours on social media each day. Would the data fit into a one-way or two-way table? Is the study observational or experimental?

■ 3. The following table shows the number of classes from which students were absent and their final grade in the class. Does the data have a positive correlation, negative correlation, or no correlation?



| | | | | | | | | | | | | |
|---------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Number of absences | 0 | 0 | 1 | 2 | 3 | 3 | 3 | 5 | 5 | 6 | 7 | 10 |
| Final grade | 95% | 97% | 90% | 86% | 80% | 74% | 70% | 65% | 64% | 58% | 55% | 45% |

- 4. The table below shows the favorite winter activity of 50 adults. Is it a one-way data table? Why or why not?

| | Skiing | Snowboarding | Ice Skating |
|--------------|---------------|---------------------|--------------------|
| Men | 9 | 13 | 6 |
| Women | 8 | 7 | 7 |

- 5. Is the following experiment an example of a double-blind experiment? If not, what could be changed to make it a double-blind experiment?

“A soda company has developed a new flavor and wants to know how it compares in taste to competitor sodas. An employee of the soda company conducts a survey where participants are asked which soda tastes the best. The sodas are given to participants in unmarked plastic cups by the employee.”

- 6. A new cancer drug is being used to treat cancer in children and adults. The hospital conducts a study to measure the effectiveness of the new drug. Cancer patients are placed into groups according to their age and each age range is split into two groups. One group is given traditional treatment of the cancer and the other group is given the new drug. Will



the data fit into a one-way or two-way table? Is the study observational or experimental?



SAMPLING AND BIAS

- 1. The zoo conducts a survey on why patrons enjoy coming to the zoo. They ask families with children about why they like to visit the zoo as they're leaving. Give a reason why the sampling method may be biased.

- 2. The owner of a restaurant gives a survey to each customer. Included in the survey is the question "Have you ever not tipped your waiter or waitress?" Give a reason why the sampling method may be biased.

- 3. A health club wants to purchase a new machine and would like to know which machine members would most like to have. It creates a survey where members can rate the different machines that the health club is considering purchasing, and posts it at the reception desk for members to fill out if they choose to do so. Does the sample contain a bias? If so, what kind?

- 4. A biologist wants to study a group of prairie dogs for parasites, but cannot examine the entire population. Which sampling method would be better in this case, a stratified random sample or a clustered random sample?



- 5. A hospital is studying the health effects of obesity. They group patients into different groups according to a specific weight range and study a variety of biometrics. What type of sampling is this?

- 6. A museum wants to find out the demographics of its patrons. They set up a survey and ask every 5th customer about their age, ethnicity, and gender. What type of sampling is this?



SAMPLING DISTRIBUTION OF THE SAMPLE MEAN

- 1. The population of 32 year-old women in the United States have an average salary of \$42,000, but the distribution of their salaries is not normally distributed. A random sample of 24 women is taken. Does the sample meet the criteria to use the central limit theorem?
- 2. There are 130 dogs at a dog show who weigh an average of 11 pounds with a standard deviation of 3 pounds. A sample of 9 dogs is taken. What is the standard deviation of the sampling distribution of the sample mean?
- 3. A large university population has an average student age of 30 years old with a standard deviation of 5 years, and student age is normally distributed. A sample of 80 students is randomly taken. What is the probability that the mean of their ages will be less than 29?
- 4. A cereal company packages cereal in 12.5-ounce boxes with a standard deviation of 0.5 ounces. The amount of cereal put into each box is normally distributed. The company randomly selects 100 boxes to check their weight. What is the probability that the mean weight will be greater than 12.6 ounces?



- 5. A large hospital finds that the average body temperature of their patients is 98.4° , with a standard deviation of 0.6° , and we'll assume that body temperature is normally distributed. The hospital randomly selects 30 patients to check their temperature. What is the probability that the mean temperature of these patients \bar{x} is within 0.2° of the population mean?
- 6. A company produces volleyballs in a factory. Individual volleyballs are filled to an approximate pressure of 7.9 PSI (pounds per square inch), with a standard deviation of 0.2 PSI. Air pressure in the volleyballs is normally distributed. The company randomly selects 50 volleyballs to check their pressure. What is the probability that the mean amount of pressure in these balls \bar{x} is within 0.05 PSI of the population mean?



CONDITIONS FOR INFERENCE WITH THE SDSM

- 1. There are 1,000 students at our school, and we ask 150 of them to tell us their height as they exit school at the end of the day. Have we met the conditions for inference?
- 2. We randomly sample 400 boxes (with replacement) in a very large, national shipping warehouse and record their weight in kilograms. Have we met the conditions for inference?
- 3. A cookie company makes packages of cookies, where the weight of the packages is normally distributed with $\mu = 500$ grams and $\sigma = 4$ grams. If the cookie company's production manager randomly selects 100 packages of cookies, what is the probability that the sample mean is within 7.5 grams of the population mean?
- 4. A sushi chef builds a sushi roll approximately every 3 minutes, with a standard deviation of 15 seconds, every night while his restaurant is open between 5 : 00 p.m. and 10 : 00 p.m., Tuesday through Sunday. The time spent to build sushi rolls is normally distributed. If the chef takes a random sample of 20 sushi rolls over the course of a week, what is the probability that the sample mean is within 5 seconds of the population mean?



- 5. The time spent playing video games by competitive gamers is normally distributed with $\mu = 40$ hours per week, and $\sigma = 2.5$ hours. If we take a random sample with replacement of 75 players and record the number of hours they spend playing this week, what's the probability that our sample mean is within 30 minutes of the population mean?
- 6. The time it takes for a roofing company to install a new roof on a single-story house normally distributed with $\mu = 6$ hours and $\sigma = 1$ hour. If the company's owner takes a random sample with replacement of 10 roofing jobs, what's the probability that his sample mean is within 45 minutes of the population mean?



SAMPLING DISTRIBUTION OF THE SAMPLE PROPORTION

- 1. The state representatives want to know how their constituents feel about the new tax to fund road improvements, so they send out a survey. Of the 5 million who reside in the state, 150,000 people respond. 40 % disapprove of the new tax and 60 % are in favor of the new tax because of the improvements they've seen to the roads. Does this sample meet the conditions for inference?

- 2. An ice cream shop states that only 5 % of their 1,200 customers order a sugar cone. We want to verify this claim, so we randomly select 120 customers to see if they order a sugar cone. Does this sample meet the conditions for inference?

- 3. The zoo conducts a study about the demographics of its patrons, and wants to learn about how many groups that visit the zoo bring children under age 12. Every 10th customer or group is recorded as a “family,” and classified as either “including children under 12” or “not including children under 12.” The zoo collected data on 65 families, and 45 of them are classified as “not including children under 12.” That day, 650 families came to the zoo. What is the standard error of the sampling distribution of the sample proportion?



- 4. A pizza shop finds that 80 % of the 75 randomly selected pizzas ordered during the week have pepperoni. What is the standard error of the proportion if the pizza shop has a total of 1,000 pizzas ordered during the week?
- 5. A hospital conducts a survey and finds that 10 patients of 30 who are randomly selected on a given day have high blood pressure. There were 325 patients in the hospital that day. What is the standard error of the proportion?
- 6. A study claims that first-born children are more likely to become leaders. The study finds that 72 % of 2,000 first-born children are currently in or have held leadership roles in their careers. Another group of scientists wants to verify the claim, but can't survey all 2,000 people, so they randomly sample 175 of the participants. What is the probability that their results are within 2 % of the first study's claim?



CONDITIONS FOR INFERENCE WITH THE SDSP

- 1. A gym owner takes a random sample of 10 local fitness instructors and asks them whether or not they train clients at multiple gyms. He finds that $\hat{p} = 70\%$ of them report training clients at multiple gyms. Can he proceed with a hypothesis test?
- 2. A professional basketball player makes 87.5% of his free throws. If he takes a random sample (without replacement) of 100 of his own free throws, can he move forward with a hypothesis test?
- 3. If the basketball player from the previous question finds a sample proportion $\hat{p} = 0.85$ in his sample of 100 free throws, calculate his test statistic. Remember that $p = 0.875$.
- 4. A grocery chain claims that 75% of their customers say that they are “satisfied” with their local store. We want to verify this claim, so we take a random sample of 45 of their customers and ask them whether or not they are “satisfied.” How likely is it that our results are within 2% of the chain’s claim?



- 5. A professional pickleball player claims that he wins 60 % of the points he plays in championship matches. We want to verify this claim, so we take a random sample of 25 of his points in championship matches and record whether or not he wins each point. How likely is it that our results are within 5 % of the player's claim?
- 6. A company reports that the proportion of its invoices that get paid on time is $p = 35 \%$. A clerk on the Accounts Receivable team wants to verify this claim, so she takes a random sample of 80 invoices and records whether or not they were paid on time. How likely is it that her sample proportion will fall within 10 % of the company's claim?



THE STUDENT'S T-DISTRIBUTION

- 1. We take a random sample of size $n = 25$, and we want to be 99 % confident about our results. What t -score will we find?

- 2. We take a random sample of size $n = 18$, and we want to be 90 % confident about our results. What t -score will we find?

- 3. We take a random sample of size $n = 8$, and we want to be 95 % confident about our results. What t -score will we find?

- 4. We take a random sample of size $n = 14$, and our upper-tail probability will be 0.05. What t -score will we find?

- 5. We take a random sample of size $n = 21$, and our upper-tail probability will be 0.001. What t -score will we find?

- 6. We take a random sample of size $n = 3$, and our upper-tail probability will be 0.025. What t -score will we find?



CONFIDENCE INTERVAL FOR THE MEAN

- 1. We want to determine the mean of calories served in a restaurant meal in America. The government has already done a study to find this mean, and they found $\sigma = 350.2$. We randomly sample 31 meals and find $\bar{x} = 1,500$. Construct and interpret a 95 % confidence interval for the mean number of calories in a restaurant meal.

- 2. A bus travels between Kansas City and Denver. We take a sample of 30 trips and find a mean travel time of $\bar{x} = 12$ hours with standard deviation $s = 0.25$ hours. Construct and interpret a 95 % confidence interval for the mean bus trip time in hours from Kansas City to Denver.

- 3. A student wanted to know how many chocolates were in the small bags of chocolate candies her school was selling for a fundraiser. She took a simple random sample of 20 small bags of chocolate candy. From the sample, she found an average of 17 pieces of candy per bag with a standard deviation of 2.03.

A box-plot of the data from the sample showed the distribution to be approximately normal. Compute and interpret a 95 % confidence interval for the mean number of chocolate candies per bag.



- 4. Consider the formula for a confidence interval for a population mean with an unknown sample standard deviation. How does doubling the sample size affect the confidence interval?

$$(a, b) = \bar{x} \pm t^* \cdot \frac{s}{\sqrt{n}}$$

- 5. A magazine took a random sample of 30 people and reported the average spending on an Easter basket this year to be \$44.78 per basket with a sample standard deviation of \$18.10. Construct and interpret a 98 % confidence interval for the data.

- 6. A confidence interval for a study is (11.5,18.5). What was the value of the sample mean?



CONFIDENCE INTERVAL FOR THE PROPORTION

- 1. According to a recent poll, 47 % of the 648 Americans surveyed make weekend plans based on the weather. Construct and interpret a 99 % confidence interval for the percentage of Americans who make weekend plans based on the weather.
- 2. We want to determine the proportion of teenagers who own their own cell phone. We take a random sample of 100 teenagers and find that 86 of them own a cell phone. At 90 % confidence, build a confidence interval for the population proportion.
- 3. A biologist is trying to determine the proportion of plants in a jungle that are ferns. She takes a random sample of 82 plants and finds that 31 of them can be classified as ferns. At 95 % confidence, what is the confidence interval for the population proportion?
- 4. A statistics teacher at a university conducted a study and found that 80 % of university students are interested in taking a statistics class. We want to see if this proportion holds at your own university. Find the minimum sample size we can use to keep a margin of error of 0.02 at a 99 % confidence level.



■ 5. Sarah is conducting a class survey to determine if the percentage of juniors in favor of having the next dance at a local bowling alley is 65 %. How many juniors should she survey in order to be 90 % confident with a margin of error of 0.08?

■ 6. A study suggests that 10 % of practicing physicians are cognitively impaired. What random sample of practicing physicians is needed to confirm this finding at a confidence level of 95 % with a margin of error of 0.05?



