STAT 216 Coursepack

Fall 2025
Montana State University

Melinda Yager
Jade Schmidt
Stacey Hancock

Contents

Preface

This coursepack accompanies the textbook for STAT 216: Montana State Introductory Statistics with R, which can be found at https://mtstateintrostats.github.io/IntroStatTextbook/. The syllabus for the course (including the course calendar), data sets, and links to D2L Brightspace, Gradescope, and the MSU RStudio server can be found on the course webpage: https://math.montana.edu/courses/s216/. Other notes and review materials are linked in D2L.

Each of the activities in this workbook is designed to target specific learning outcomes of the course, giving you practice with important statistical concepts in a group setting with instructor guidance. In addition to the in-class activities for the course, video notes are provided to aid in taking notes while you complete the required videos. Bring this workbook with you to class each class period, and take notes in the workbook as you would your own notes. A well-written completed workbook will provide an optimal study guide for exams!

All activities and labs in this coursepack will be completed during class time. Parts of each lab will be turned in on Gradescope. To aid in your understanding, read through the introduction for each activity before attending class each day.

STAT 216 is a 3-credit in-person course. In our experience, it takes six to nine hours per week outside of class to achieve a good grade in this class. By "good" we mean at least a C because a grade of D or below does not count toward fulfilling degree requirements. Many of you set your goals higher than just getting a C, and we fully support that. You need roughly nine hours per week to review past activities, read feedback on previous assignments, complete current assignments, and prepare for the next day's class. A typical week in the life of a STAT 216 student looks like:

- *Prior to class meeting*:
    - Read assigned sections of the textbook, using the provided reading guides to take notes on the material.
    - Watch the provided videos, taking notes in the coursepack.
    - Read through the introduction to the day's in-class activity.
    - Read through the week's homework assignment and note any questions you may have on the content.
- *During class meeting*:
    - Work through the guided activity, in-class activity or weekly lab with your classmates and instructor, taking detailed notes on your answers to each question in the activity.
- *After class meeting*:
    - Complete any parts of the activity you did not complete in class.
    - Review the activity solutions in the Math and Stat Center, and take notes on key points.
    - Complete any remaining assigned readings for the week.
    - Complete the week's homework assignment.

Basics of Data and Sampling Methods

1.1 Vocabulary Review and Key Topics

At the beginning of each module is a summary of key topics and new vocabulary terms for that module. As you read through the material in the textbook and watch the videos prior to class, look for these terms. Reference the definitions to guide your understanding.

1.1.1 Key topics

Module 1 introduces the foundations of data: observational units, types of variables, and how to collect sample data from a population of interest in a way that allows us to generalize our results back to the population.

1.1.2 Vocabulary

- **Data**: observations used to answer research questions.
- **Observational units (cases)**: the subjects or entities on which data are collected.
    - The rows in a data set represent the observational units.
- **Sample size**: the number of observational units in a data set, denoted by n.
- **Variable**: the characteristics collected on each observational unit.
- **Types of variables**:
    - **Categorical**: cases are grouped into categories.
    - **Quantitative**: numerical measurements, where performing arithmetic operations makes sense.
- **Target population**: group of observational units of interest.
- **Sample**: subset of the population.
- **Sampling methods**:
    - **Unbiased sampling method (e.g., a random sample)**: on average, the sample will be representative of the target population; all observational units in the target population have the same chance of being selected.
    - **Biased sampling method (e.g., convenience sample)**: on average, the sample will not be representative of the target population; some part of the target population will be over- or under-represented.
- **Type of sampling bias**:
    - **Selection bias**: method of sampling is biased; some part of the target population is over- or under-represented.
    - **Non-response bias**: part of a pre-selected sample does not respond or cannot be reached.
    - **Response bias**: responses are not truthful (poor/leading question phrasing, social desirability).

- **Generalization**: to what group of observational units can the results be applied to?
  - If an unbiased method of selection was used and there is no non-response or response bias, we can generalize the results to the target population.
  - If a biased method of selection was used or if non-response or response bias is present, we can only generalize the result to the sample or similar observational units.

1.2 Activity 1: Intro to Data Analysis and Sampling Bias

1.2.1 Learning outcomes

- Identify observational units, variables, and variable types in a statistical study.

- Creating a data set

- Identify biased sampling methods.

1.2.2 Terminology review

Statistics is the study of how best to collect, analyze, and draw conclusions from data. In the next few days of class, we will learn the building blocks of the semester. This week in class you will be introduced to the following terms:

- Observational units or cases

- Variables: categorical or quantitative

- Selection bias

- Response bias

- Non-response bias

For more on these concepts, read Chapter 1 and 2 in the textbook.

Notes on Observational Units and Variables

Further analysis of class data set

1. What are the observational units or cases for the data collected in class on day 1?

2. How many observations are reported in the data set? This is the **sample size**.

3. The header for each column in the data set describes each variable measured on the observational unit. For each column of data, fill in the following table identifying the type of each variable.

- If the variable is categorical, indicate in column 3 whether the variable is binary.

- If the variable is quantitative, indicate in column 4 the units of measure used.

Column	Type of Variable	Binary?	Units?
Major			
Residency			
Num Credits			
Dominant hand			
Hand Span			
Grip strength dominant hand			
Grip strength non-dominant hand			

4. Review the completed data set with your class. Remember that when creating a data set for use in R it is important to use single words or an underscore between words. Each outcome must be written the same way each time to have consistency between responses. Do not include units of measure in the data set when reporting numerical values. Write down some issues found with the created class data set.

Notes on Sampling Methods and Types of bias

Types of bias

Complete Q5 together as a class:

5. A television station is interested in predicting whether or not local voters will pass a referendum to legalize marijuana for adult. The TV station asks its viewers to phone in and indicate whether they are in favor or opposed to the referendum. Of the 2241 viewers who phoned in, forty-five percent were opposed to legalizing marijuana.

   Sample size:

   Observational units sampled:

   Target population:

   Justify why there is selection bias in this study.

6. To determine if the proportion of out-of-state undergraduate students at Montana State University has increased in the last 10 years, a statistics instructor sent an email survey to 500 randomly selected current undergraduate students. One of the questions on the survey asked whether they had in-state or out-of-state residency. She only received 378 responses.

   Sample size:

   Observational units sampled:

   Target population:

   Justify why there is non-response bias in this study.

7. To gauge the interest of Bozeman City Voters in a new swimming pool, a local organization stood outside of the Bogart Pool in Bozeman, MT, during open hours. One of the questions they asked was, "Since the Bogart Pool is in such bad repair, don't you agree that the city should fund a new pool?"

Sample size:

Observational units sampled:

Target population:

Justify why there is response bias in this study.

Justify why there is selection bias in this study.

1.2.3 Take-home messages

1. There are two types of variables: categorical (groups) and quantitative (numerical measures).

2. We will learn more about summarizing variable later in the semester. Categorical variables are summarized by calculating a proportion from the data and quantitative variables are summarized by finding the mean and the standard deviation.

3. There are three types of bias to be aware of when designing a sampling method: selection bias, non-response bias, and response bias.

1.2.4 Additional notes

Use this space to summarize your thoughts and take additional notes on today's activity and material covered, and to write down the names and contact information of your teammates.

1.3 Activity 2: American Indian Address

1.3.1 Learning outcomes

- Explain why a sampling method is unbiased or biased.

- Identify biased sampling methods.

- Explain the purpose of random selection and its effect on generalization.

1.3.2 Terminology review

In this activity, we will examine unbiased and biased methods of sampling. Some terms covered in this activity are:

- Random sample

- Unbiased vs biased methods of selection

- Generalization

To review these concepts, see Chapter 2 in the textbook.

1.3.3 Class Preparation

Prior to the next class, complete questions 1–3.

1.3.4 American Indian Address

For this activity, you will read a speech given by Jim Becenti, a member of the Navajo American Indian tribe, who spoke about the employment problems his people faced at an Office of Indian Affairs meeting in Phoenix, Arizona, on January 30, 1947 (Moquin and Van Doren 1973). His speech is below:

It is hard for us to go outside the reservation where we meet strangers. I have been off the reservation ever since I was sixteen. Today I am sorry I quit the Santa Fe [Railroad]. I worked for them in 1912–13. You are enjoying life, liberty, and happiness on the soil the American Indian had, so it is your responsibility to give us a hand, brother. Take us out of distress. I have never been to vocational school. I have very little education. I look at the white man who is a skilled laborer. When I was a young man I worked for a man in Gallup as a carpenter's helper. He treated me as his own brother. I used his tools. Then he took his tools and gave me a list of tools I should buy and I started carpentering just from what I had seen. We have no alphabetical language.

We see things with our eyes and can always remember it. I urge that we help my people to progress in skilled labor as well as common labor. The hope of my people is to change our ways and means in certain directions, so they can help you someday as taxpayers. If not, as you are going now, you will be burdened the rest of your life. The hope of my people is that you will continue to help so that we will be all over the United States and have a hand with you, and give us a brotherly hand so we will be happy as you are. Our reservation is awful small. We did not know the capacity of the range until the white man come and say "you raise too much sheep, got to go somewhere else," resulting in reduction to a skeleton where the Indians can't make a living on it. For eighty years we have been confused by the general public, and what is the condition of the Navajo today? Starvation! We are starving for education. Education is the main thing and the only thing that is going to make us able to compete with you great men here talking to us.

By eye selection

1. Circle ten words in Jim Becenti's speech which are a representative sample of the length of words in the entire text. Describe your method for selecting this sample.

2. Fill in the table below with your selected words from the previous question and the length of each word (number of letters/digits in the word):

Observation	Word	Length
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

3. Calculate the mean (average) word length in your selected sample. Is this value a parameter or a statistic?

Notes on sampling

1.3.5   Class Activity

1. Report your mean word length from question3 to your instructor. Your instructor will create a visualization of the distribution of results generated by your class. Draw a picture of the plot here. Include a descriptive x-axis label. Report the mean and standard deviation of the sample mean word lengths.

The plot created in question 1 is a sampling distribution of statistics. This sampling distribution plots the mean word length from many samples taken from the population of words.

2. The true mean word length of the population of all 359 words in the speech is 3.95 letters. Is this value a parameter or a statistic?

   Where does the value of 3.95 fall in the plot given? Near the center of the distribution? In the tails of the distribution?

3. Based on the class discussion, would you say the sampling method used ("by-eye" selection) by the class is biased or unbiased? Justify your answer.

4. If the sampling method is biased, what type of sampling bias (selection, response, non-response) is present? What is the direction of the bias, i.e., does the method tend to overestimate or underestimate the population mean word length?

Random selection

Suppose instead of attempting to select a representative sample by eye (which did not work), each student used a random number generator to select a simple random sample of 10 words. A **simple random sample** relies on a random mechanism to choose a sample, without replacement, from the population, such that every sample of size 10 is equally likely to be chosen.

To use a random number generator to select a simple random sample, you first need a numbered list of all the words in the population, called a **sampling frame**. You can then generate 10 random numbers from the numbers 1 to 359 (the number of words in the population), and the chosen random numbers correspond to the chosen words in your sample.

5. Use the random number generator at https://istats.shinyapps.io/RandomNumbers/ to select a simple random sample from the population of all 359 words in the speech.

- Set "Choose Minimum" to 1 and "Choose Maximum" to 359 to represent the 359 words in the population (the sampling frame).

- Set "How many numbers do you want to generate?" to 10 and ensure the "No" option is selected under "Sample with Replacement?"

- Click "Generate".

Fill in the table on the next page with the random numbers selected and use the Becenti.csv data file found on D2L to determine each number's corresponding word and word length (number of letters/digits in the word):

Observation	Number	Length
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

6. Calculate the mean word length in your selected sample in question 5. Is this value a parameter or a statistic?

7. Report your mean word length in the Google sheet. Your instructor will create a visualization of the distribution of results generated by your class. Draw a picture of the plot here. Include a descriptive x-axis label. Report the mean and standard deviation of the samples.

8. Where does the value 3.95, the true mean word length, fall in the distribution given? Near the center of the distribution? In the tails of the distribution? Circle this value on the provided distribution.

One set of randomly generated sample mean word lengths from a single class may not be large enough to visualize the distribution results. Let's have a computer generate 1,000 sample mean word lengths for us.

The following plot illustrates a sampling distribution of 1000 samples of size 10 selected at random from the sample.

Statistic: ⦿ Mean ◯ Median ◯ *t*-statistic

Mean=3.913
SD =0.716
Num samples=1000
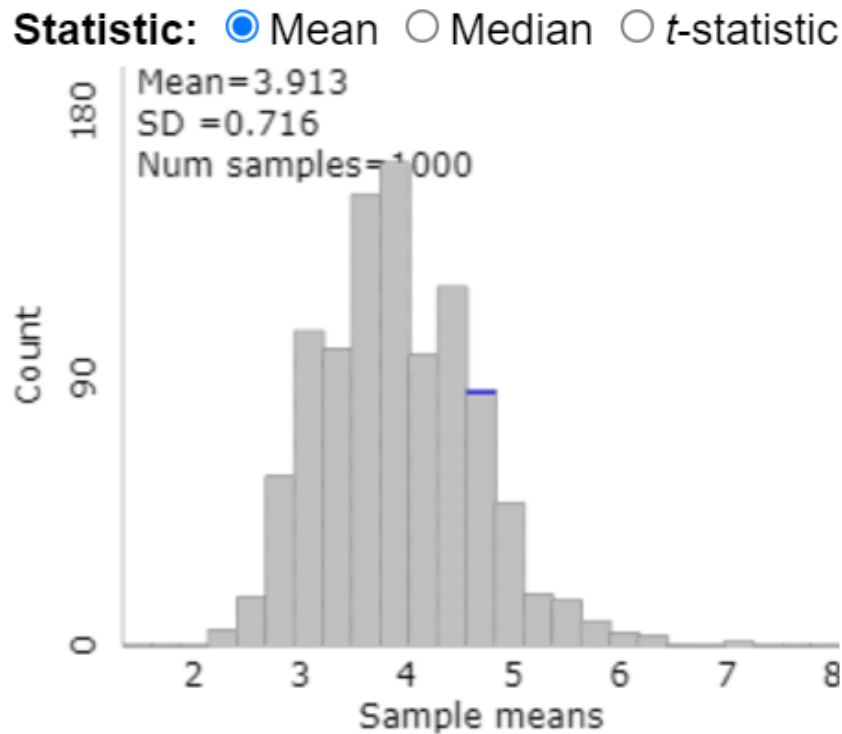
9. What is the center value (mean) of the distribution displayed above?

10. Explain why the sampling method of using a random number generator to generate a sample is a "better" method than choosing 10 words "by eye".

11. Is random selection an unbiased method of selection? Explain your answer. Be sure to reference the plot from before Q15.

Effect of sample size

We will now consider the impact of sample size.

12. First, consider if each student had selected 30 words, instead of 10, by eye. Do you think this would make the plot from the previous activity centered on 3.95 (the true mean word length)? Explain your answer.

Now we will select 30 words instead of 10 words at random. The following plot illustrates a sampling distribution of 1000 samples of size 30 selected at random from the sample.

Statistic: ● Mean   ○ Median   ○ *t*-statistic

Mean=3.954
SD =0.411
Num samples=1000
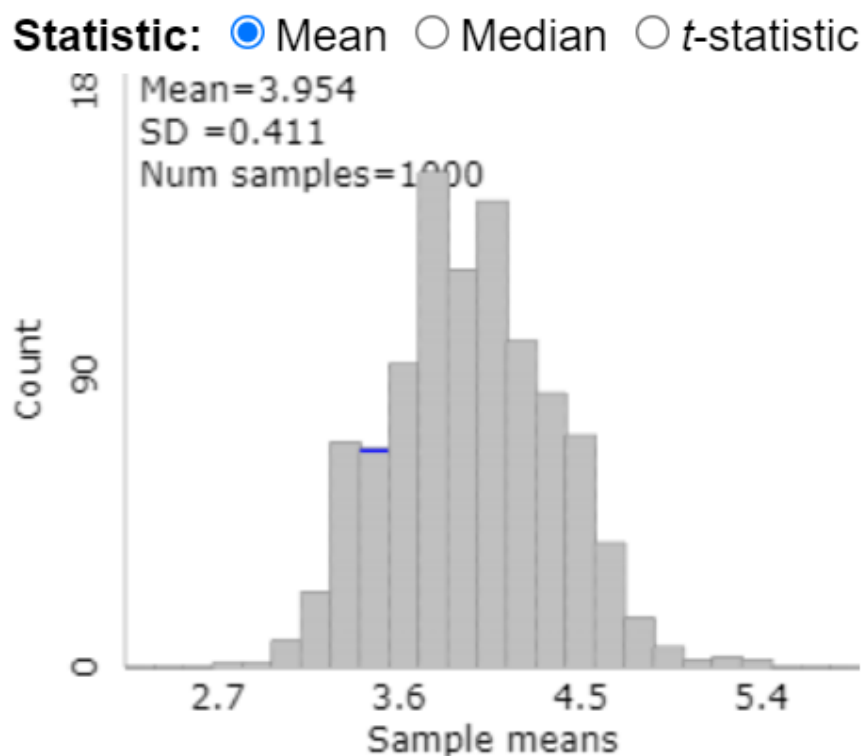
13. Compare the values of the standard deviation of the plots before question 9 and before question 13. Which plot shows the smallest standard deviation?

14. Using the evidence from your simulations, answer the following research questions:

    Does changing the sample size impact whether the sample estimates are unbiased? Explain your answer.

    Does changing the sample size impact the variability (spread) of sample estimates? Explain your answer

15. What is the purpose of random selection of a sample from the population?

1.3.6   Take-home messages

1. When we use a biased method of selection, we will over or underestimate the parameter.

2. If the sampling method is biased, inferences made about the population based on a sample estimate will not be valid.

3. Random selection is an unbiased method of selection.

4. To determine if a sampling method is biased or unbiased, we compare the distribution of the estimates to the true value. We want our estimate to be on target or unbiased. When using unbiased methods of selection, the mean of the distribution matches or is very similar to our true parameter.

5. Random selection eliminates selection bias. However, random selection will not eliminate response or non-response bias.

6. The larger the sample size, the more similar (less variable) the statistics will be from different samples.

7. Sample size has no impact on whether a *sampling method* is biased or not. Taking a larger sample using a biased method will still result in a sample that is not representative of the population.

1.3.7   Additional notes

Use this space to summarize your thoughts and take additional notes on today's activity and material covered.

Probability

2.1 Vocabulary Review and Key Topics

2.1.1 Key topics

Module 2 introduces the concept of probability as a long-run relative frequency and demonstrates how to use hypothetical two-way tables to set up a probability problem and solve for unconditional and conditional probabilities.

2.1.2 Vocabulary

- **Probability** (of an event): the long-run proportion of times the event would occur if the random process were repeated indefinitely (under identical conditions).

- **Conditional probability** (of an event *given* another event): probability of an event calculated dependent on another event having occurred.

- **Probability notation**:

  - $P(A)$: the probability of event A.

    * This is the probability of a single event, *unconditional* probability calculated out of the overall population.

  - $P(A^C)$: the probability of the **complement** of event A, or "A complement".

    * This is the probability of the opposite of event A, or "not A".

    * $P(A^C) = 1 - P(A)$

  - $P(A \text{ and } B)$: the probability of event A and B.

    * The is the probability of an "and" event, *unconditional* probability calculated out of the overall population.

  - $P(A|B)$: the probability of event A given (conditional on) event B.

    * This is a *conditional* probability calculated out of the total population for which event B occurred.

2.2 Activity 3: Probability Studies

2.2.1 Learning outcomes

- Recognize and simulate probabilities as long-run frequencies.
- Construct two-way tables to evaluate conditional probabilities.

2.2.2 Terminology review

In today's activity, we will cover two-way tables and probability. Some terms covered in this activity are:

- Proportions
- Probability
- Conditional probability
- Two-way tables

To review these concepts, see Chapter 23 in the textbook.

Notes on probability

The probability of an event is the long-run proportion of times the event would occur if the random process were repeated indefinitely (under identical conditions).

To calculate the probability of an event happening:

$$\text{probability} = \frac{\text{number of ways an event can happen}}{\text{total number of possible outcomes}}$$

For example, to calculate the probability of a coin flip landing on heads; there are only two outcomes (heads or tails) and only one possibility way to land on heads.

$$P(heads) = \frac{1}{2} = 0.5$$

The figure below shows the long-run proportion of times a simulated coin flip lands on heads on the y-axis, and the number of tosses on the x-axis. Notice how the long-run proportion starts converging to 0.5 as the number of tosses increases.

19
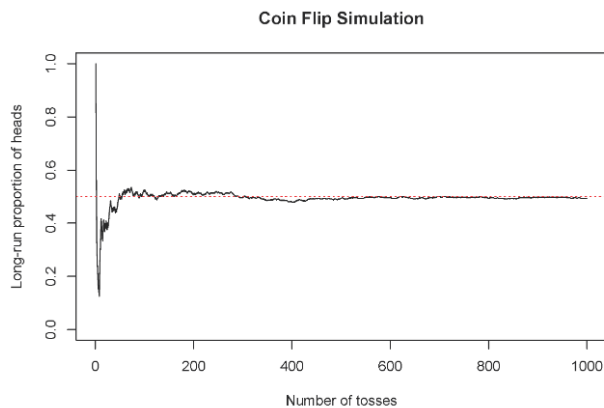
In today's activity we will discuss the probability of a single event, the probability of an "and" event, and the probability of a conditional event.

Probability notation

We will use the notation P(event) to represent the probability of an event and use letters to represent events. The following are notations for different probabilities where we are discussing event A and event B:

- $P(A)$ represents the probability of event A

- $P(A^C)$ represents the probability of the complement of event A

    - $P(A^C) = 1 - P(A)$

- $P(A and B)$ represents the probability of events A and B

- $P(A|B)$ represents the probability of event A given event B

- $P(B|A)$ represents the probability of event B given event A

2.2.2.1   Probability questions

For the beginning of this activity we will start with discussing the probabilities associated with drawing a card from a standard card deck. In a card deck there are:

- 52 cards

- Half are red, half are black

- Four suits: spades, hearts, diamonds, and clubs

- Each suit has 13 cards: cards 2–10, ace, jack, queen, and king

- Let A represent the event that a card is an ace

- Let B represent the event that a card is red

To find the probability of selecting an ace, first start with determining how many aces are possible (four) and how many cards will we select from (total of 52).

Find the probability of selecting a card that is not an ace. This is the complement of event A.

Find the probability of selecting a red ace.

Find the probability of selecting an ace given that the card is red.

If a card drawn is an ace, what is the probability the card drawn is red.

2.2.3 Calculating probabilities from a two-way table

1. In 2014, the website FiveThirtyEight examined the works of Bob Ross to see what trends could be found. They determined that of all the paintings he created, 95% of them contained at least one "happy tree." Of those works with a happy tree, 43% contained at least one "almighty mountain." Of the paintings that did not have at least one happy tree, only 10% contained at least one almighty mountain.

Let A = Bob Ross painting contains a happy tree, and B = Bob Ross painting contains an almighty mountain

	A	A^C	Total
B	40850	500	41350
B^C	54150	4500	58650
Total	95000	5000	100000

a. What is the probability that a randomly selected Bob Ross painting contains both a "happy tree" and an "almighty mountain"? Use appropriate probability notation.

b. What is the probability that a selected Bob Ross painting without an "almighty mountain" contains a "happy tree." Use appropriate probability notation.

c. What is the probability that a selected Bob Ross painting does not contain a "happy tree" given it does not contain an "almighty mountain". Use appropriate probability notation.

2. A recent study of population decline of white-tailed deer in Wyoming due to chronic wasting disease (Edmunds 2016) (CWD) reported that 35.4% of white-tailed deer have CWD. The survival rate of deer with CWD is 39.6% and the survival rate of deer without CWD is 80.1%.

Let A = the event a deer has CWD, and B = the event the deer survives.

a. Identify what each numerical value given in the problem represents in probability notation.

0.354 =

0.396 =

0.801 =

b. Create a hypothetical two-way table to represent the situation.

	A	A^C	Total
B			
B^C			
Total			100,000

c. Find $P(A \text{ and } B)$. What does this probability represent in the context of the problem?

d. Find the probability that a deer that has CWD does not survive. What is the notation used for this probability?

e. What is the probability that a deer does not survive given they do not have CWD? What is the notation used for this probability?

3. Since the early 1980s, the rapid antigen detection test (RADT) of group A *streptococci* has been used to detect strep throat. A recent study of the accuracy of this test shows that the **sensitivity**, the probability of a positive RADT given the person has strep throat, is 86% in children, while the **specificity**, the probability of a negative RADT given the person does not have strep throat, is 92% in children. The **prevalence**, the probability of having group A strep, is 37% in children. (Stewart et al. 2014)

Let A = the event the child has strep throat, and B = the event the child has a positive RADT.

a. Identify what each numerical value given in the problem represents in probability notation.

$0.86 =$

$0.92 =$

$0.37 =$

b. Create a hypothetical two-way table to represent the situation.

	A	A^C	Total
B			
B^C			
Total			100,000

c. Find $P(B)$. What does this probability represent in the context of the problem?

d. Find the probability that a child with a positive RADT actually has strep throat. What is the notation used for this probability?

e. What is the probability that a child does not have strep given that they have a positive RADT? What is the notation used for this probability?

23

2.2.4 Take home messages

1. Conditional probabilities are calculated dependent on a second variable. In probability notation, the variable following | is the variable on which we are conditioning. The denominator used to calculate the probability will be the total for the variable on which we are conditioning.

2. When creating a two-way table we typically want to put the explanatory variable on the columns of the table and the response variable on the rows.

3. To fill in the two-way table, always start with the unconditional variable in the total row or column and then use the conditional probabilities to fill in the interior cells.

2.2.5 Additional notes

Use this space to summarize your thoughts and take additional notes on today's activity and material covered.

Exploring Categorical Data: Exploratory Data Analysis and Inference using Simulation-based Methods

---

3.1 Vocabulary Review and Key Topics

Review the Golden Ticket posted in the resources at the end of the coursepack for a summary of a single categorical variable.

3.1.1 Key topics

Module 3 introduces the steps of the statistical investigation process. We conduct **exploratory data analysis** (summary statistics and plots) and simulation-based **inference** (hypothesis testing and confidence intervals) in the single categorical variable (one proportion) scenario.

- Notation for a sample proportion: $\hat{p}$

- Notation for a population proportion: π

- Types of plots for a single categorical variable:

    - Frequency bar plot

    - Relative frequency bar plot

Exploratory data analysis is step 3 of the statistical investigation process. We will then use simulation-based methods **to find evidence of an effect by finding a p-value** and **estimating how large the effect is by creating a confidence interval** in the one proportion (one categorical variable) scenario. These are steps 4 and 5 from the steps of the statistical investigation process.

Steps of the statistical investigation process

As we move through the semester we will work through the six steps of the statistical investigation process.

1. Ask a research question.

2. Design a study and collect data.

3. Summarize and visualize the data.

4. Use statistical analysis methods to draw inferences from the data.

5. Communicate the results and answer the research question.

6. Revisit and look forward.

3.1.2 Vocabulary

- **Summary measure**: a numerical quantity that summarizes data. Summary measures covered in STAT 216 include: single proportion, difference in proportions, single mean, paired mean difference, difference in means, correlation, and slope of a regression line.

    - For a single categorical variable, a proportion is calculated.

- **Summary statistic (point estimate)**: the value of a numerical summary measure computed from *sample* data.
    - To interpret in context include:
        * Summary measure (in context)
        * Value of the statistic
- **Parameter of interest**: a numerical summary measure of the entire *population* in which we are interested.
    - The value of the parameter of interest is unknown (unless we have access to the entire population).
    - To write in context:
        * Population word (true, long-run, population)
        * Summary measure (depends on the type of data)
        * Context
            · Observational units
            · Variable(s)
- For a single categorical variable, the category that we are counting the proportion of is generically called a "**success**", with categories not a success labeled "**failure**". Thus, a sample proportion is the "proportion of successes" in the sample: the total number of successes divided by the sample size (n).

Plotting one categorical variable

- **Frequency bar plot**: plots the count (frequency) of observational units in each level of a categorical variable. R code to create a frequency bar plot:

```
object %>% # Data set piped into...
ggplot(aes(x = variable)) +   # This specifies the variable
geom_bar(stat = "count") +  # Tell it to make a bar plot
labs(title = "Don't forget to title your plot!",
   # Give your plot a title
   x = "x-axis label",   # Label the x axis
   y = "Frequency")  # Label the y axis
```

- **Relative frequency bar plot**: plots the proportion (relative frequency) of observational units in each level of a categorical variable. R code to create a relative frequency bar plot:

```
object %>% # Data set piped into...
ggplot(aes(x = variable)) +   # This specifies the variable
geom_bar(aes(y = after_stat(prop), group = 1)) +  # Tell it to make a bar plot with proportions
labs(title = "Don't forget to title your plot!",
   # Give your plot a title
   x = "x-axis label",   # Label the x axis
   y = "Relative Frequency")  # Label the y axis
```

Inference

- **Sampling distribution** (of a statistic): the distribution of possible values of a statistic across repeated samples of the same size and under the same conditions.
    - We can create a *simulated* sampling distribution using simulation-based methods to simulate many samples, or we can mathematically model the sampling distribution (theory-based methods).

- **Hypothesis testing**: a formal statistical technique for evaluating two competing possibilities about a population: the null hypothesis and alternative hypothesis.

  - When we observe an effect in a sample, we would like to determine if this observed effect represents an actual effect in the population, or whether it was simply due to random chance.

  - A hypothesis test helps us answer the following question about the population: How strong is the *evidence* of an effect?

- **Null hypothesis**: typically represents a statement of "no difference", "no effect", or the status quo.

  - The null hypothesis is what we assume is true when calculating the p-value. Thus, we can never have evidence *for* the null hypothesis—we cannot "accept" a null hypothesis—we can only find evidence *against* the null hypothesis if the observed data is very unlikely to have occurred under the assumption that the null hypothesis is true.

- **Alternative hypothesis**: represents an alternative claim under consideration and is often represented by a range of possible values for the parameter of interest.

  - The alternative hypothesis is determined by the research question.

- **Hypotheses in notation for a single proportion**: In the hypotheses below, π_0 is the **null value**.

$$H_0 : \pi = \pi_0$$

$$H_A : \pi \left\{ \begin{array}{c} < \\ \neq \\ < \end{array} \right\} \pi_0$$

- **P-value**: the probability of the value of the observed sample statistic or a value more extreme, if the null hypothesis were true.

  - To write in context include:

    * Statement about probability or proportion of samples

    * Statistic (summary measure and value)

    * Direction of the alternative

    * Null hypothesis (in context)

- **Strength of evidence**: the p-value indicates the amount of evidence there is against the null hypothesis. The smaller the p-value the more evidence there is against the null hypothesis.

27
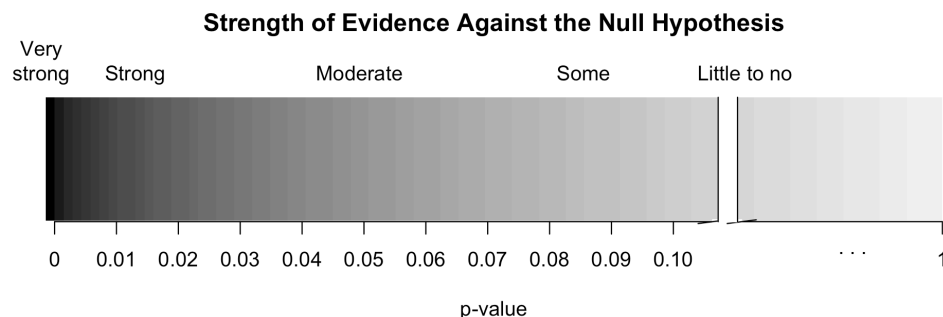
- **Conclusion** (to a hypothesis test): answers the research question. How much evidence is there in support of the alternative hypothesis?

    – To write in context include:

      * Amount of evidence

      * Parameter of interest

      * Direction of the alternative hypothesis

- **Confidence interval**: an interval estimate for the parameter of interest; an interval of *plausible values* for the parameter.

    – A confidence interval helps us answer the following question about the population: How *large* is the effect?

    – To write in context include:

      * How confident you are (e.g., 90%, 95%, 98%, 99%)

      * Parameter of interest

      * Calculated interval

3.1.2.1 Simulation-based inference for a single proportion

- **Conditions necessary to use simulation-based methods for inference for a single categorical variable**:

    – **Independence**: observational units must be independent of one another; the outcome of one observational unit should have no influence on the outcome of another.

- **Null distribution**: a sampling distribution of simulated sample statistics created under the assumption that the null hypothesis is true

- **Simulation-based methods to create the null distribution**: a process of using a computer program (e.g., R) to simulate many samples that we would expect based on the null hypothesis.

R code to use simulation methods for one categorical variable to find the p-value, `one_proportion_test` (from the `catstats` package), is shown below.

```
one_proportion_test(probability_success = xx, # Null hypothesis value
      sample_size = xx, # Enter sample size
      number_repetitions = 10000, # Enter number of simulations
      as_extreme_as = xx, # Observed statistic
      direction = "xx", # Specify direction of alternative hypothesis
      summary_measure = "proportion") # Reporting proportion or number of successes?
```

- **Bootstrapping**: creating a simulated sample of the same size as the original sample by sampling with replacement from the original sample.

- **Simulation-based methods to create the bootstrap distribution**: a process of using a computer program to simulate many bootstrapped samples.

R code to use simulation methods for one categorical variable to find a confidence interval, `one_proportion_bootstrap_CI` (from the `catstats` package), is shown below.

```
one_proportion_bootstrap_CI(sample_size = xx, # Sample size
            number_successes = xx, # Observed number of successes
            number_repetitions = 10000, # Number of bootstrap samples to use
            confidence_level = 0.95) # Confidence level as a decimal
```

- **Percentile method**: process to find the confidence interval from the bootstrap distribution.
    - A 90% confidence interval will be found between the 5th and 95th percentiles of the bootstrap distribution.
    - A 95% confidence interval will be found between the 2.5th and 97.5th percentiles of the bootstrap distribution.
    - A 99% confidence interval will be found between the 0.5th and 99.5th percentiles of the bootstrap distribution.

3.2 Activity 4: Helper-Hinderer Part 1 — Simulation-based Hypothesis Test

3.2.1 Learning outcomes

- Identify the two possible explanations (one assuming the null hypothesis and one assuming the alternative hypothesis) for a relationship seen in sample data.

- Given a research question involving a single categorical variable, construct the null and alternative hypotheses in words and using appropriate statistical symbols.

- Describe and perform a simulation-based hypothesis test for a single proportion.

3.2.2 Terminology review

In today's activity, we will work through a simulation-based hypothesis testing for a single categorical variable. Some terms covered in this activity are:

- Parameter of interest

- Null hypothesis

- Alternative hypothesis

- Simulation

To review these concepts, see Chapters 9 & 14 in your textbook.

3.2.3 Steps of the statistical investigation process

We will work through a five-step process to complete a hypothesis test for a single proportion, first introduced in the activity in week 1.

- **Ask a research question** that can be addressed by collecting data. What are the researchers trying to show?

- **Design a study and collect data**. This step involves selecting the people or objects to be studied and how to gather relevant data on them.

- **Summarize and visualize the data**. Calculate summary statistics and create graphical plots that best represent the research question.

- **Use statistical analysis methods to draw inferences from the data**. Choose a statistical inference method appropriate for the data and identify the p-value and/or confidence interval after checking assumptions. In this study, we will focus on using randomization to generate a simulated p-value.

- **Communicate the results and answer the research question**. Using the p-value and confidence interval from the analysis, determine whether the data provide statistical evidence against the null hypothesis. Write a conclusion that addresses the research question.

Notes on Hypothesis Testing

3.2.4 Helper-Hinderer

A study by Hamblin, Wynn, and Bloom reported in Nature (Hamblin, Wynn, and Bloom 2007) was intended to check young kids' feelings about helpful and non-helpful behavior. Non-verbal infants ages 6 to 10 months were shown short videos with different shapes either helping or hindering the climber. As a class we will watch this short video to see how the experiment was run: https://youtu.be/anCaGBsBOxM. Researchers were hoping to assess: Are infants more likely to choose the helper toy? In the study, of the 16 infants age 6 to 10 months, 14 chose the *helper* toy and 2 chose the *hinderer* toy.

- Observational units:

- Variable:

- Type of variable:

- Success:

Ask a research question

- Research question:

Design a study and collect data

Before using statistical inference methods, we must check that the cases are independent. The sample observations are independent if the outcome of one observation does not influence the outcome of another. One way this condition is met is if data come from a simple random sample of the target population.

1. Are the cases independent? Justify your answer.

R code

For almost all activities and labs it will be necessary to upload the provided R script file from D2L for that day. Your instructor will highlight a few steps in uploading files to and using RStudio.

The following are the steps to upload the necessary R script file for this activity:

- Download the Activity R script file from D2L.

- Click "Upload" in the "Files" tab in the bottom right window of RStudio. In the pop-up window, click "Choose File", and navigate to the folder where the Activity R script file is saved (most likely in your downloads folder). Click "Open"; then click "Ok".

- You should see the uploaded file appear in the list of files in the bottom right window. Click on the file name to open the file in the Editor window (upper left window).

Notice that the first threelines of code contain a prompt called `library`. Packages needed to run functions in R are stored in directories called libraries. When using the MSU RStudio server, all the packages needed for the class are already installed. We simply must tell R which packages we need for each R script file. We use the prompt `library` to load each **package** (or library) needed for each activity. Note, these `library` lines MUST be run each time you open a R script file in order for the functions in R to work.

- Highlight and run lines 1–3 to load the packages needed for this activity. Notice the use of the # symbol in the R script file. This symbol is not part of the R code. It is used by these authors to add comments to the R code and explain what each call is telling the program to do.

R will ignore everything after a # symbol when executing the code. Refer to the instructions following the # symbol to understand what you need to enter in the code.

```
library(tidyverse)
library(ggplot2)
library(catstats)
```

Throughout activities, we will often include the R code you would use in order to produce output or plots. These "code chunks" appear in gray. In the code chunk below, we demonstrate how to read the data set into R using the `read.csv()` function. The line of code shown below (line 7 in the R script file) reads in the data set and names the data set `infants`.

Summarize and visualize the data

The following code reads in the data set and gives the number of infants in each level of the variable, whether the infant chose the helper or the hinderer.

- Highlight and run lines 7 and 8 to check that you get the same counts as shown below

```
# Read in data set
infants <- read.csv("https://math.montana.edu/courses/s216/data/infantchoice.csv")
infants %>% count(choice)  # Count number in each choice category
```

```
#>     choice  n
#> 1   helper 14
#> 2 hinderer  2
```

The following formula is used to calculate the proportion of successes in the sample.

$$\hat{p} = \frac{\text{number of successes}}{\text{total number of observational units}}$$

2. Using the R output and the formula given, calculate the summary statistic (sample proportion) to represent the research question. Recall that `choosing the helper toy` is a considered a success. Use appropriate notation.

To visually display this data we can use either a frequency bar plot or a relative frequency bar plot.

- Enter the variable name `choice` for `variable` in the R code to create the frequency bar plot.

- Note the name of the title is given in line 16 and includes the **type of plot**, **observational units**, and **variable name**.

- Highlight and run lines 13–19 to create the plot

```
infants %>% # Data set piped into...
    ggplot(aes(x = variable)) +    # This specifies the variable
    geom_bar(stat = "count") +   # Tell it to make a bar plot
    labs(title = "Frequency Bar Plot of Toy Choice for Pre-verbal Infants",
       # Give your plot a title
       x = "Toy Choice",    # Label the x axis
       y = "Frequency")   # Label the y axis
```

3. Sketch the frequency bar plot created below.

We could also choose to display the data as a proportion in a **relative frequency** bar plot. To find the relative frequency, the count in each level of `choice` is divided by the sample size. This calculation is the sample proportion for each level of `choice`. Notice that in the following code we told R to create a bar plot with proportions.

- In the R script file, highlight and run lines 23–29 to create the relative frequency bar plot.

```
infants %>% # Data set piped into...
    ggplot(aes(x = choice)) +    # This specifies the variable
    geom_bar(aes(y = after_stat(prop), group = 1)) +   # Tell it to make a bar plot with proportions
    labs(title = "Relative Frequency Bar Plot of Toy Choice for Pre-verbal Infants",
       # Give your plot a title
       x = "Toy Choice",    # Label the x axis
       y = "Relative Frequency")   # Label the y axis
```

Relative Frequency Bar Plot of Toy Choice for Pre-verb...
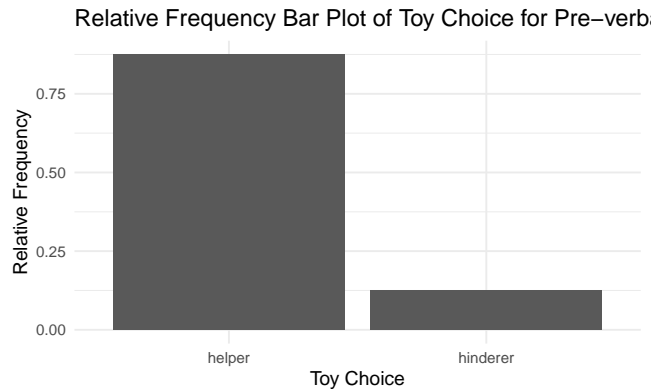
4. Which features in the relative frequency bar plot are the same as the frequency bar plot? Which are different?

We cannot assess whether infants are more likely to choose the helper toy based on the statistic and plot alone. The next step is to analyze the data by using a hypothesis test to discover if there is evidence against the null hypothesis.

Use statistical analysis methods to draw inferences from the data

When performing a hypothesis test, we must first identify the null hypothesis. The null hypothesis is written about the parameter of interest, or the value that summarizes the variable in the population.

The parameter of interest is a statement about what we want to find about the population. The following must be included when writing the parameter of interest.

- Population word (true, long-run, population)

- Summary measure (depends on the type of data)

- Context

  - Observational units

  - Variable(s)

Parameter of interest:

If the children are just randomly choosing the toy, we would expect half (0.5) of the infants to choose the helper toy. This is the null value for our study.

Null Hypothesis (in words):

The notation used for a population proportion (or probability, or true proportion) is π. Since this summarizes a population, it is a parameter. When writing the **null hypothesis** in notation, we set the parameter equal to the null value, $H_0 : \pi = \pi_0$.

Null Hypothesis (in notation):

The **alternative hypothesis** is the claim to be tested and the direction of the claim (less than, greater than, or not equal to) is based on the research question.

- Based on the research question from question 1, are we testing that the parameter is greater than 0.5, less than 0.5 or different than 0.5?

Alternative hypothesis (in words):

Alternative hypothesis (in notation):

Remember that when utilizing a hypothesis test, we are evaluating two competing possibilities. For this study the **two possibilities** are either...

- The true proportion of infants who choose the helper is 0.5 and our results just occurred by random chance; or,
- The true proportion of infants who choose the helper is greater than 0.5 and our results reflect this.

Notice that these two competing possibilities represent the null and alternative hypotheses.

We will now simulate one sample of a **null distribution** of sample proportions. The null distribution is created under the assumption the null hypothesis is true. In this case, we assume the true proportion of infants who choose the helper is 0.5, so we will create 10000 (or more) different simulations of 16 infants under this assumption.

Let's think about how to use a coin to create one simulation of 16 infants under the assumption the null hypothesis is true. Let heads equal infant chose the helper toy and tails equal infant chose the hinderer toy.

5. How many times would you flip a coin to simulate the sample of infants?

6. Flip a coin 16 times recording the number of times the coin lands on heads. This represents one simulated sample of 16 infants randomly choosing the toy. Calculate the proportion of coin flips that resulted in heads.

7. Is the value from question 6 closer to 0.5, the null value, or closer to the sample proportion, 0.875?

Report the number of coin flips you got as indicated by your instructor.

8. Sketch the graph created by your instructor of each student's proportion of heads out of 16 coin flips.

9. Circle the observed statistic (value from question 2) on the distribution shown above. Where does this statistic fall in this distribution: Is it near the center of the distribution (near 0.5) or in one of the tails of the distribution?

10. Is the observed statistic likely to happen or unlikely to happen if the true proportion of infants who choose the helper is 0.5? Explain your answer using the plot.

In the next class, we will continue to assess the strength of evidence against the null hypothesis by using a computer to simulate 10000 samples when we assume the null hypothesis is true.

3.2.5 Take-home messages

1. Two types of plots are used for plotting categorical variables: frequency bar plots, relative frequency bar plots.

2. In a hypothesis test we have two competing hypotheses, the null hypothesis and the alternative hypothesis. The null hypothesis represents either a skeptical perspective or a perspective of no difference or no effect. The alternative hypothesis represents a new perspective such as the possibility that there has been a change or that there is a treatment effect in an experiment.

3. In a simulation-based test, we create a distribution of possible simulated statistics for our sample if the null hypothesis is true. Then we see if the calculated observed statistic from the data is likely or unlikely to occur when compared to the null distribution.

4. To create one simulated sample on the null distribution for a sample proportion, spin a spinner with probability equal to π_0 (the null value), n times or draw with replacement n times from a deck of cards created to reflect π_0 as the probability of success. Calculate and plot the proportion of successes from the simulated sample.

3.2.6 Additional notes

Use this space to summarize your thoughts and take additional notes on today's activity and material covered.

3.3 Activity 5: Helper-Hinderer (continued)

3.3.1 Learning outcomes

- Describe and perform a simulation-based hypothesis test for a single proportion.

- Interpret and evaluate a p-value for a simulation-based hypothesis test for a single proportion.

- Explore what a p-value represents

3.3.2 Steps of the statistical investigation process

In today's activity we will continue with steps 4 and 5 in the statistical investigation process. We will continue to assess the Helper-Hinderer study from last class.

- **Ask a research question** that can be addressed by collecting data. What are the researchers trying to show?

- **Design a study and collect data**. This step involves selecting the people or objects to be studied and how to gather relevant data on them.

- **Summarize and visualize the data**. Calculate summary statistics and create graphical plots that best represent the research question.

- **Use statistical analysis methods to draw inferences from the data**. Choose a statistical inference method appropriate for the data and identify the p-value and/or confidence interval after checking assumptions. In this study, we will focus on using randomization to generate a simulated p-value.

- **Communicate the results and answer the research question**. Using the p-value and confidence interval from the analysis, determine whether the data provide statistical evidence against the null hypothesis. Write a conclusion that addresses the research question.

3.3.3 Helper-Hinderer

In class today, we will revisit the study on infants as described below.

A study by Hamblin, Wynn, and Bloom reported in Nature (Hamblin, Wynn, and Bloom 2007) was intended to check young kids' feelings about helpful and non-helpful behavior. Non-verbal infants ages 6 to 10 months were shown short videos with different shapes either helping or hindering the climber. As a class we will watch this short video to see how the experiment was run: https://youtu.be/anCaGBsBOxM. Researchers were hoping to assess: Are infants more likely to choose the helper toy? In the study, of the 16 infants age 6 to 10 months, 14 chose the *helper* toy and 2 chose the *hinderer* toy.

1. Report the sample proportion (summary statistic) calculated in the previous activity.

2. Write the alternative hypothesis in words in context of the problem. Remember the direction we are testing is dependent on the research question.

Today, we will use the computer to simulate a null distribution of 10000 different samples of 16 infants, plotting the proportion who chose the helper in each sample, based on the assumption that the true proportion of infants who choose the helper is 0.5 (or that the null hypothesis is true).

To use the computer simulation, we will need to enter the

- assumed "probability of success" (π_0),
- "sample size" (the number of observational units or cases in the sample),
- "number of repetitions" (the number of samples to be generated - typically we use 10000),
- "as extreme as" (the observed statistic), and
- the "direction" (matches the direction of the alternative hypothesis).

3. What values should be entered for each of the following into the one proportion test to create 10000 simulations?

- Probability of success (null value):

- Sample size (n):

- Number of repetitions (always use 10000 simulations):

- As extreme as (value of statistic):

- Direction (`"greater"`, `"less"`, or `"two-sided"`):

We will use the `one_proportion_test()` function in R (in the `catstats` package) to simulate the null distribution of sample proportions and compute a p-value. Using the provided R script file, fill in the values/words for each `xx` with your answers from question 3 in the one proportion test to create a null distribution with 10000 simulations. Then highlight and run lines 1–16.

```r
one_proportion_test(probability_success = xx, # Null hypothesis value
          sample_size = xx, # Enter sample size
          number_repetitions = 10000, # Enter number of simulations
          as_extreme_as = xx, # Observed statistic
          direction = "xx", # Specify direction of alternative hypothesis
          summary_measure = "proportion") # Reporting proportion or number of successes?
```

4. Sketch the null distribution created from the R code here.

Notes on the null distribution

5. Circle the observed statistic (value from question 1) on the distribution you drew in question 4. Where does this statistic fall in the null distribution: Is it near the center of the distribution (near 0.5) or in one of the tails of the distribution?

6. Is the observed statistic likely to happen or unlikely to happen if the true proportion of infants who choose the helper is 0.5? Explain your answer using the plot.

7. Using the simulation, what is the proportion of simulated samples that generated a sample proportion at the observed statistic or greater, if the true proportion of infants who choose the helper is 0.5? *Hint*: Look under the simulation.

Notes on the p-value

The value in question 7 is the **p-value**. The smaller the p-value, the more evidence we have against the null hypothesis.

Strength of Evidence Against the Null Hypothesis
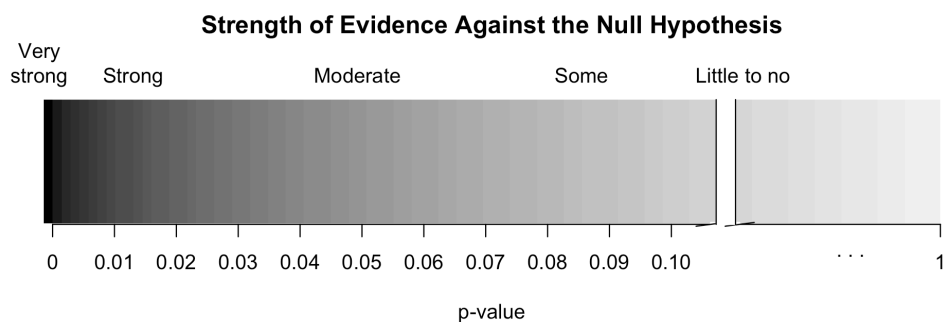
Interpret the p-value

The p-value measures the probability that we observe a sample proportion as extreme as what was seen in the data or more extreme (matching the direction of the Ha) IF the null hypothesis is true. This is a conditional probability, calculated dependent on the null hypothesis being true. Represented in probability notation:

$$P(\text{statistic or more extreme}|\text{the null hypothesis is true})$$

p-value interpretation:

Communicate the results and answer the research question

When we write a conclusion we answer the research question by stating how much evidence there is for the alternative hypothesis.

Conclusion:

Generalization

8. To what group of observational units can the results be generalized to?

3.3.4 Take-home messages

1. The null distribution is created based on the assumption the null hypothesis is true. We compare the sample statistic to the distribution to find the likelihood of observing this statistic.

2. The p-value measures the probability of observing the sample statistic or more extreme (in direction of the alternative hypothesis) is the null hypothesis is true.

3. The smaller the p-value of the test, the more evidence there is **against** the null hypothesis.

3.3.5   Additional notes

Use this space to summarize your thoughts and take additional notes on today's activity and material covered.

3.4 Activity 6: Helper-Hinderer — Simulation-based Confidence Interval

3.4.1 Learning outcomes

- Use bootstrapping to find a confidence interval for a single proportion.

- Interpret a confidence interval for a single proportion.

3.4.2 Terminology review

In today's activity, we will introduce simulation-based confidence intervals for a single proportion. Some terms covered in this activity are:

- Parameter of interest

- Bootstrapping

- Confidence interval

To review these concepts, see Chapters 10 & 14 in your textbook.

3.4.3 Helper-Hinderer

In the last class, we found very strong evidence that the true proportion of infants who will choose the helper character is greater than 0.5. But what *is* the true proportion of infants who will choose the helper character? We will use this same study to estimate this parameter of interest by creating a confidence interval.

As a reminder: A study by Hamblin, Wynn, and Bloom reported in Nature (Hamblin, Wynn, and Bloom 2007) was intended to check young kids' feelings about helpful and non-helpful behavior. Non-verbal infants ages 6 to 10 months were shown short videos with different shapes either helping or hindering the climber. Researchers were hoping to assess: Are infants more likely to preferentially choose the helper toy? In the study, of the 16 infants age 6 to 10 months, 14 chose the *helper* toy and 2 chose the *hinderer* toy.

A **point estimate** (our observed statistic) provides a single plausible value for a parameter. However, a point estimate is rarely perfect; usually there is some error in the estimate. In addition to supplying a point estimate of a parameter, a next logical step would be to provide a plausible *range* of values for the parameter. This plausible range of values for the population parameter is called an **interval estimate** or **confidence interval**.

Activity intro

1. What is the value of the point estimate (sample statistic)?

2. If we took another random sample of 16 infants, would we get the exact same point estimate? Explain why or why not.

In today's activity, we will use bootstrapping to find a 95% confidence interval for π, the parameter of interest.

Notes on Confidence Intervals

Use statistical analysis methods to draw inferences from the data

3. Write out the parameter of interest in words, in context of the study. What does π represent?

To create the null distribution we flipped a coin 16 times to simulate infants randomly choosing the helper toy with a probability of 50%.

4. Why can't we use a coin to simulate the bootstrap distribution.

To create the bootstrap distribution.

- First we would label the cards to represent the sample statistic: 14 helper and 2 hinderer.

- Sample with replacement 16 times

5. Using the cards provided by your instructor, create one bootstrap sample. Report your simulated sample proportion on the whiteboard.

To use the computer simulation to create a bootstrap distribution, we will need to enter the

- "sample size" (the number of observational units or cases in the sample),
- "number of successes" (the number of cases that choose the helper character),
- "number of repetitions" (the number of samples to be generated), and
- the "confidence level" (which level of confidence are we using to create the confidence interval).

6. What values should be entered for each of the following into the simulation to create the bootstrap distribution of sample proportions to find a 95% confidence interval?

- Sample size (n):

- Number of successes:

- Number of repetitions (at least 10000):

- Confidence level (as a decimal):

We will use the `one_proportion_bootstrap_CI()` function in R (in the `catstats` package) to simulate the bootstrap distribution of sample proportions and calculate a confidence interval. Using the provided R script file, fill in the values/words for each `xx` with your answers from question 6 in the one proportion bootstrap confidence interval (CI) code to create a bootstrap distribution with 10000 simulations. Then highlight and run lines 1–9.

```
one_proportion_bootstrap_CI(sample_size = xx, # Sample size
                    number_successes = xx, # Observed number of successes
                    number_repetitions = 10000, # Number of bootstrap samples to use
                    confidence_level = xx) # Confidence level as a decimal
```

7. Sketch the bootstrap distribution created below.

Notes on the bootstrap distribution

95% Confidence Interval:

Interpretation of the 95% confidence interval in context.

**Communicate the results and answer the research question**

8. Is the value 0.5 (the null value) in the 95% confidence interval?

Explain how this indicates that the p-value provides strong evidence against the null.

Effect of confidence level

9. Suppose instead of finding a 95% confidence interval, we found a 90% confidence interval. Would you expect the 90% confidence interval to be narrower or wider? Explain your answer.

10. The following R code produced the bootstrap distribution with 10000 simulations that follows. Circle the value that changed in the code.

```
one_proportion_bootstrap_CI(sample_size = 16, # Sample size
                    number_successes = 14, # Observed number of successes
                    number_repetitions = 10000, # Number of bootstrap samples to use
                    confidence_level = 0.90) # Confidence level as a decimal
```


Bootstrapped Proportions

Mean: 0.876, SD: 0.082, 90% CI: (0.75, 1)
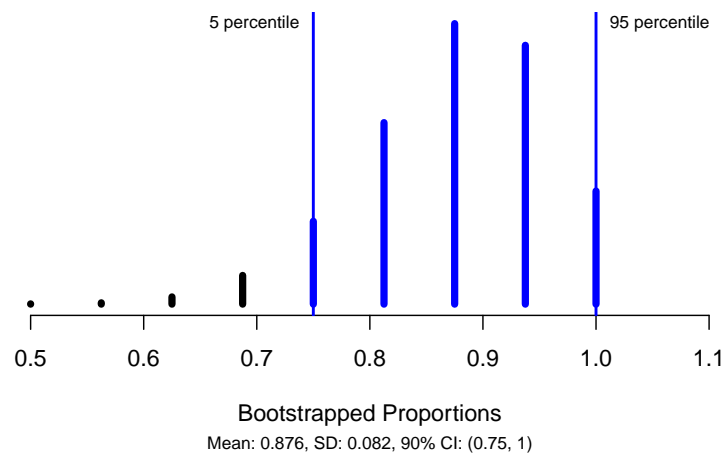
11. Report both the 95% confidence interval and the 90% confidence interval (question 10). Is the 90% confidence interval narrower or wider than the 95% confidence interval?

Concluding paragraph

In many of our studies we will write a paragraph summarizing the results of the study.

Researchers were interested if infants observe social cues and would be more likely to choose the helper toy. In a sample of _____ infants, _____ chose the helper toy. A simulation null distribution with 10000 simulations was created in RStudio. The p-value was found by calculating the proportion of simulations in the null distribution at the sample statistic of 0.875 and greater. This resulted in a p-value of _____. We would observe a sample proportion of _____ or _____ with a probability of _____

IF we assume _____. Based on this

p-value, there is _____ evidence that the _____ proportion of infants age 6 to 10 months who will choose the helper toy is _____ 0.5. In addition, a 95% confidence interval was found for the parameter of interest. We are 95% confident that the _____ proportion of infants age 6 to 10 months who will choose the helper toy is between _____ and _____. The results of this study can be generalized to _____ as the researchers _____ select a random sample.

3.4.4 Take-home messages

1. The goal in a hypothesis test is to assess the strength of evidence for an effect, while the goal in creating a confidence interval is to determine how large the effect is. A **confidence interval** is a range of *plausible* values for the parameter of interest.

2. A confidence interval is built around the point estimate or observed calculated statistic from the sample. This means that the sample statistic is always the center of the confidence interval. A confidence interval includes a measure of sample to sample variability represented by the **margin of error**.

3. In simulation-based methods (bootstrapping), a simulated distribution of possible sample statistics is created showing the possible sample-to-sample variability. Then we find the middle X percent of the distribution around the sample statistic using the percentile method to give the range of values for the confidence interval. This shows us that we are $X\%$ confident that the parameter is within this range, where X represents the level of confidence.

4. When the null value is within the confidence interval, it is a plausible value for the parameter of interest; thus, we would find a larger p-value for a hypothesis test of that null value. Conversely, if the null value is NOT within the confidence interval, we would find a small p-value for the hypothesis test and strong evidence against this null hypothesis.

5. To create one simulated sample on the bootstrap distribution for a sample proportion, label n cards with the original responses. Draw with replacement n times. Calculate and plot the resampled proportion of successes.

3.4.5 Additional notes

Use this space to summarize your thoughts and take additional notes on today's activity and material covered.

Inference for a Single Categorical Variable: Theory-based Methods

4.1 Vocabulary Review and Key Topics

Review the Golden Ticket posted in the resources at the end of the coursepack for a summary of a single categorical variable.

4.1.1 Key topics

Module 4 introduces theory-based inference methods (hypothesis testing and confidence intervals) for a single categorical variable. We also explore what "confidence level" means and which parts of a study impact the width of a confidence interval and the p-value.

- Theory-based methods should give the same results as simulation-based methods if the sample size is large enough. For a single categorical variable, the sample size is large enough if the success-failure condition is met.

- If repeated samples of the same size are taken from the population, 95% of samples will create a 95% confidence interval that contains the value of the parameter of interest.

4.1.2 Vocabulary

- **Theory-based methods**: when specific conditions are met, the distribution of sample statistics if we were to repeatedly sample from the population can be fit with a theoretical distribution.

- **Conditions for the sampling distribution of $\hat{p}$ to follow an approximate normal distribution**:
  - **Independence**: the sample's observations are independent, e.g., are from a simple random sample. (*Remember*: This also must be true to use simulation-based methods!)
  - **Success-failure condition**: we *expect* to see at least 10 successes and 10 failures in the sample, $n\pi \geq 10$ and $n(1-\pi) \geq 10$. Since π is typically unknown, we consider this condition to be met if we observe at least 10 successes and 10 failures in our data set: $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$.

- **Standard normal distribution**: a theoretical distribution that is bell-shaped, centered on the mean of zero, and has a standard deviation of one, denoted in notation by $N(0,1)$.

- **Standard error of a statistic**: an estimated standard deviation of the statistic as it would vary across repeated samples of the same size under the same conditions.
  - The standard error tells us about how far we would expect an observed sample statistic to fall from the true parameter value for which it is estimating, on average.

- **Standardized statistic**: calculation to standardize the sample statistic in order to compare the standardized value to the theoretical distribution.
  - Calculated by subtracting the null value from the sample statistic, then dividing by the standard error:
  $$\frac{\text{statistic} - \text{null value}}{\text{standard error}}$$

- Measures the number of standard errors the sample statistic is above (if positive) or below (if negative) the null value.

- **Standard error of the sample proportion assuming the null is true**: measures the how far each possible sample proportion is from the true proportion, on average, and is calculated using the null value:

$$SE_0(\hat{p}) = \sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}$$

.

- **Standardized sample proportion**: standardized statistic for a single categorical variable calculated using:

$$Z = \frac{\hat{p} - \pi_0}{SE_0(\hat{p})},$$

  If the conditions for the sampling distribution of $\hat{p}$ to follow an approximate normal distribution are met, and if the true value of π is equal to the null value of π_0, the standardized sample proportion, Z, will have an approximate *standard* normal distribution.

- The theory-based **p-value** for hypothesis testing involving proportions can be found in R by using the `pnorm` function to find the probability of the observed standardized statistic or one more extreme (in the direction of H_A). This probability is the area under a *standard normal distribution* at or more extreme than the observed standardized statistic.

  - Enter the value of the standardized statistic for `xx`.

  - If a "greater than" alternative, change `lower.tail = TRUE` to `FALSE`.

  - If a two-sided test, multiply by 2.

  ```
  pnorm(xx, lower.tail=TRUE)
  ```

- **Margin of error**: half the width of the confidence interval. For a single proportion, the margin of error is:

$$ME = z^* \times SE(\hat{p})$$

  where z^* is the **multiplier**, corresponding to the desired confidence level found from the standard normal distribution. For example, for a 95% confidence level, the middle 95% of the standard normal distribution falls between $-z^* = -1.96$ and $z^* = 1.96$.

- **Standard error of the sample proportion for a confidence interval** (not assuming the null is true):

$$SE(\hat{p}) = \sqrt{\frac{\hat{p} \times (1 - \hat{p})}{n}}$$

- To find the endpoints of a confidence interval, add and subtract the margin of error to the sample statistic. The confidence interval for a population proportion is:

$$\hat{p} \pm ME$$

- R code to find the **multiplier** for the confidence interval using theory-based methods involving proportions.

  - `qnorm` will give you the multiplier using the standard normal distribution.

  - Enter the percentile for the given level of confidence (e.g., 0.975 for a 95% confidence level).

  ```
  qnorm(percentile, lower.tail=TRUE)
  ```

4.2 Activity 7: Handedness of Male Boxers

4.2.1 Learning outcomes

- Describe and perform a theory-based hypothesis test for a single proportion.
- Check the appropriate conditions to use a theory-based hypothesis test.
- Calculate and interpret the standardized sample proportion.
- Interpret and evaluate a p-value for a theory-based hypothesis test for a single proportion.
- Use the normal distribution to find the p-value.

4.2.2 Terminology review

In this activity, we will introduce theory-based hypothesis tests for a single categorical variable. Some terms covered in this activity are:

- Parameter of interest
- Standardized statistic
- Normal distribution
- p-value

To review these concepts, see Chapter 11 & 14 in your textbook.

Activities from module 3 covered simulation-based methods for hypothesis tests involving a single categorical variable. This activity covers theory-based methods for testing a single categorical variable.

4.2.3 Handedness of male boxers

Left-handedness is a trait that is found in about 10% of the general population. Past studies have shown that left-handed men are over-represented among professional boxers (Richardson and Gilman 2019). Is there evidence that there is an over-prevalence of left-handed fighters? In this random sample of 500 professional male boxers, 81 were left-handed.

- Observational units:
- Variable:
- Type of variable:
- Success:

4.2.4 Summary statistics review

- Download the R file for today's activity from D2L
- Upload the file to the R server
- Run lines 1–15 to load the needed packages and the data set and create a plot of the data

```
 # Read in data set
boxers <- read.csv("https://math.montana.edu/courses/s216/data/Male_boxers_sample.csv")
boxers %>% count(Stance)  # Count number in each Stance category

#>        Stance   n
#> 1  left-handed  81
#> 2 right-handed 419
```

```
boxers %>% # Data set piped into...
    ggplot(aes(x = Stance)) +    # This specifies the variable
    geom_bar(aes(y = after_stat(prop), group = 1)) +  # Tell it to make a bar plot with proportions
    labs(title = "_____ of Handedness of Male Professional Boxers",
        # Give your plot a title
        x = "Handedness",   # Label the x axis
        y = "Relative Frequency")  # Label the y axis
```

_____ of Handedness of Male Profes
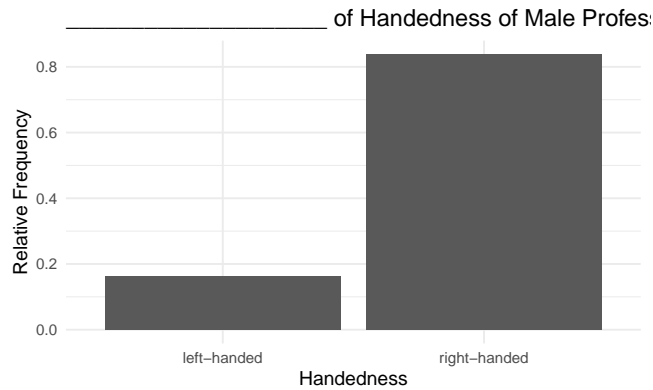
1. What type of plot was created of these data?

Hypotheses and summary statistics

2. Write out the parameter of interest in words, in context of the study.

3. Write out the null hypothesis **in words**.

4. Write out the alternative hypothesis **in notation**.

5. Calculate the value of the summary statistic (sample proportion) for this study. Use proper notation.

Theory-based methods

The sampling distribution of a single proportion — how that proportion varies from sample to sample — can be mathematically modeled using the normal distribution if certain conditions are met.

Conditions for the sampling distribution of $\hat{p}$ to follow an approximate normal distribution:

- **Independence**: The sample's observations are independent, e.g., are from a simple random sample. (*Remember*: This also must be true to use simulation methods!)

- **Large enough sample size**: Success-failure condition: We *expect* to see at least 10 successes and 10 failures in the sample, $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$.

Additional notes on Theory-based methods

- Verify that the independence condition is satisfied.

- Is the success-failure condition met to model the data with the normal distribution? Explain your answer in context of the problem.

To calculate the standardized statistic we use the general formula

$$Z = \frac{\text{point estimate} - \text{null value}}{SE_0(\text{point estimate})}.$$

For a single categorical variable the standardized sample proportion is calculated using

$$Z = \frac{\hat{p} - \pi_0}{SE_0(\hat{p})},$$

where the standard error is calculated using the null value:

$$SE_0(\hat{p}) = \sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}$$

.

For this study, the null standard error of the sample proportion is calculated using the null value, 0.1.

$$SE_0(\widehat{p}) = \sqrt{\frac{0.1 \times (1 - 0.1)}{500}} = 0.013$$

.

Each sample proportion of male boxers that are left-handed is 0.013 from the true proportion of male boxers that are left-handed, on average.

Label the standard normal distribution shown below with the null value as the center value (below the value of zero). Label the tick marks to the right of the null value by adding 1 standard error to the null value to represent 1 standard error, 2 standard errors, and 3 standard errors from the null. Repeat this process to the left of the null value by subtracting 1 standard error for each tick mark.
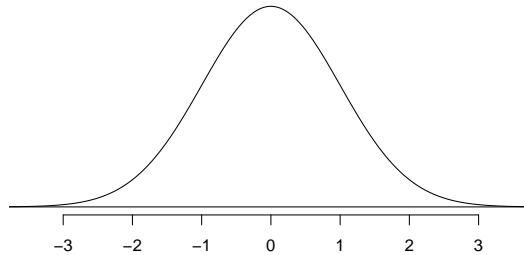

Figure 4.1: Standard Normal Curve

6. Using the null standard error of the sample proportion, calculate the standardized sample proportion (Z). Mark this value on the standard normal distribution above.

The standardized statistic is used as a ruler to measure how far the sample statistic is from the null value. Essentially, we are converting the sample proportion into a measure of standard errors to compare to the standard normal distribution.

The standardized statistic measures the *number of standard errors the sample statistic is from the null value.*

** Interpretation of the standardized sample proportion:**

We will use the `pnorm()` function in R to find the p-value. In the code below, notice that we used `lower.tail = FALSE` to find the p-value. R will calculate the p-value *greater* than the value of the standardized statistic.

Notes:

- Use `lower.tail = TRUE` when doing a left-sided test.
- Use `lower.tail = FALSE` when doing a right-sided test.
- To find a two-sided p-value, use a left-sided test for negative Z or a right-sided test for positive Z, then multiply the value found by 2 to get the p-value.

```
pnorm(4.769, # Enter value of standardized statistic
      m=0, s=1, # Using the standard normal mean = 0, sd = 1
      lower.tail=FALSE) # Gives a p-value greater than the standardized statistic
```

7. Report the p-value obtained from the R output.

8. Write a conclusion based on the p-value.

9. To what group of observational units can the results be generalized to.

Impacts on the P-value

Suppose that we want to show that the true proportion of male boxers **differs** from that in the general population.

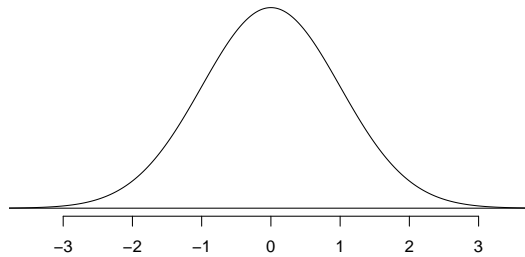10. Write out the alternative hypothesis in notation for this new research question.

Figure 4.2: Standard Normal Curve

11. How would this impact the p-value? Would the p-value be larger or smaller?

Suppose instead of 500 male boxers the researchers only took a sample of 300 male boxers and found the same proportion ($\hat{p} = 0.162$) of male boxers that are left-handed. Since we are still assuming the same null value, 0.1, the standard error would be calculated as below:

$$SE_0(\hat{p}) = \sqrt{\frac{0.1(1 - 0.1)}{300}} = 0.017$$

.

The standardized statistic for this new sample is calculated below:

$$Z = \frac{0.162 - 0.1}{0.017} = 3.64$$

12. How does the decrease in sample size affect the p-value?

Suppose another sample of 500 male boxers was taken and 68 were found to be left-handed. Since we are still assuming the same null value, 0.1, the standard error would be calculated as before:
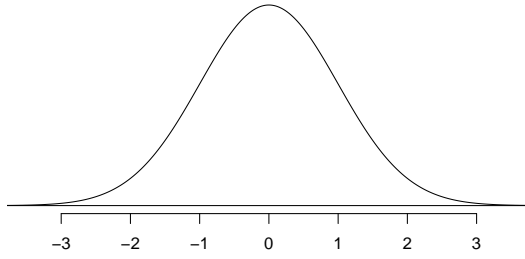
Figure 4.3: Standard Normal Curve

$$SE_0(\widehat{p}) = \sqrt{\frac{0.1(1 - 0.1)}{500}} = 0.013$$

.

The standardized statistic for this new sample is calculated below:
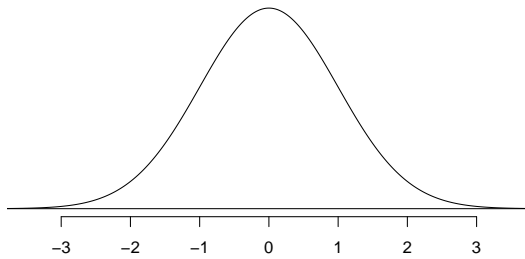
$$Z = \frac{0.136 - 0.1}{0.013} = 2.769$$

Figure 4.4: Standard Normal Curve

13. How does a statistic closer to the null value affect the p-value?

14. Summarize how each of the following affected the p-value:

a) Switching to a two-sided test.

b) Using a smaller sample size.

c) Using a sample statistic closer to the null value.

4.2.5   Take-home messages

1. Both simulation and theory-based methods can be used to find a p-value for a hypothesis test. In order to use theory-based methods we need to check that both the independence and the success-failure conditions are met.

2. The standardized statistic measures how many standard errors the statistic is from the null value. The larger the standardized statistic the more evidence there is against the null hypothesis.

3. The p-value for a two-sided test is approximately two times the value for a one-sided test. A two-sided test provides less evidence against the null hypothesis.

4. The larger the sample size, the smaller the sample to sample variability. This will result in a larger standardized statistic and more evidence against the null hypothesis.

5. The farther the statistic is from the null value, the larger the standardized statistic. This will result in a smaller p-value and more evidence against the null hypothesis.

4.2.6   Additional notes

Use this space to summarize your thoughts and take additional notes on today's activity and material covered.

4.3 Activity 8: Confidence interval and what confidence means

4.3.1 Learning outcomes

- Explore what confidence means
- Interpret the confidence level
- Explore impact of sample size, direction of the alternative hypothesis, and value of the sample statistic on the p-value.

4.3.2 Terminology review

In this activity, we will explore what being 95% confidence means. Some terms covered in this activity are:

- Parameter of interest
- Two-sided vs. one-sided tests
- Confidence level

4.3.3 Handedness of male boxers continued

In today's activity, we will use the male boxer study to look at what confidence means.

Left-handedness is a trait that is found in about 10% of the general population. Past studies have shown that left-handed men are over-represented among professional boxers (Richardson and Gilman 2019). Is there evidence that there is an over-prevalence of left-handed fighters? In this random sample of 500 professional male boxers, 81 were left-handed.

```
# Read in data set
boxers <- read.csv("https://math.montana.edu/courses/s216/data/Male_boxers_sample.csv")
boxers %>% count(Stance)  # Count number in each Stance category
```

```
#>        Stance    n
#> 1  left-handed   81
#> 2 right-handed  419
```

What does *confidence* mean?

In the interpretation of a 95% confidence interval, we say that we are 95% confident that the parameter is within the confidence interval. Why are we able to make that claim? What does it mean to say "we are 95% confident"?

1. In the last activity we found very strong evidence that the true proportion of male professional boxers that are left-handed is greater than 0.1. As a class, determine a plausible value for the true proportion of male boxers that are left-handed. *Note: we are making assumptions about the population here. This is not based on our calculated data, but we will use this applet to better understand what happens when we take many, many samples from this believed population.*

2. Go to this website, http://www.rossmanchance.com/ISIapplets.html and choose 'Simulating Confidence Intervals'. In the input on the left-hand side of the screen enter the value from question 1 for π (the true value), 500 for n, and 100 for 'Number of intervals'. Click 'sample'.

- In the graph on the bottom right, click on a green dot. Write down the confidence interval for this sample given on the graph on the left. Does this confidence interval contain the true value chosen in question 1?

- Now click on a red dot. Write down the confidence interval for this sample. Does this confidence interval contain the true value chosen in question 1?

- How many intervals out of 100 contain π, the true value chosen in question 1? *Hint*: This is given to the left of the graph of green and red intervals.

3. Click on 'sample' nine more times. Write down the 'Running Total' for the proportion of intervals that contain π.

4. Change the confidence level to 90%. What happened to the width of the intervals?

5. Write down the `Running Total` for the proportion of intervals that contain π using a 90% confidence level.

Interpretation of the level of confidence:

Notes on theory-based confidence intervals

To calculate a theory-based 95% confidence interval for π, we will first find the **standard error** of $\widehat{p}$ by plugging in the value of $\widehat{p}$ for π in $SD(\widehat{p})$:

$$SE(\widehat{p}) = \sqrt{\frac{\widehat{p} \times (1 - \widehat{p})}{n}}$$

Note that we do not include a "0" subscript, since we are not assuming a null hypothesis.

Calculate the standard error of the sample proportion to find a 95% confidence interval.

We will calculate the margin of error and confidence interval later in this activity. **The margin of error (ME) is the value of the z^* multiplier times the standard error of the statistic.**

$$ME = z^* \times SE(\widehat{p})$$

The z^* multiplier is the percentile of a standard normal distribution that corresponds to our confidence level. If our confidence level is 95%, we find the Z values that encompass the middle 95% of the standard normal distribution. If 95% of the standard normal distribution should be in the middle, that leaves 5% in the tails, or 2.5% in each tail.

The `qnorm()` function in R will tell us the z^* value for the desired percentile (in this case, $95\% + 2.5\% = 97.5\%$ percentile).
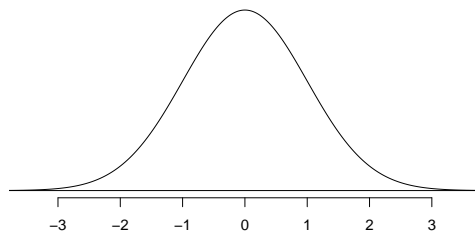

Figure 4.5: Standard Normal Curve

The following code will find the z^* value for a 95% confidence interval.

```r
qnorm(c(0.025, 0.975), lower.tail = TRUE) # Multiplier for 95% confidence interval
```

Calculate the margin of error for the 95% confidence interval.

To find the confidence interval, we will add and subtract the **margin of error** to the point estimate:

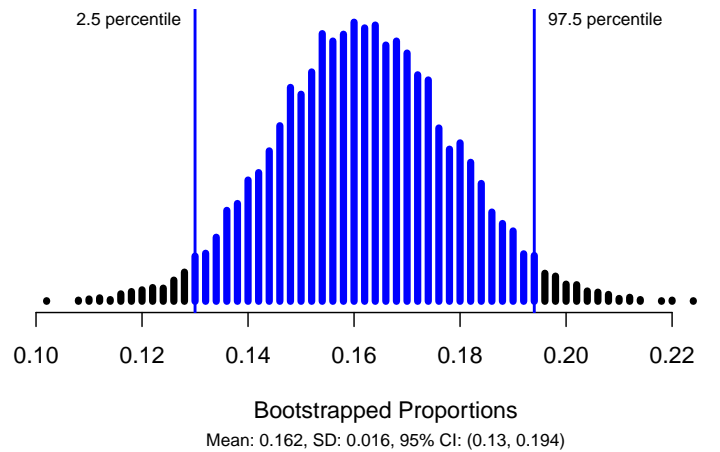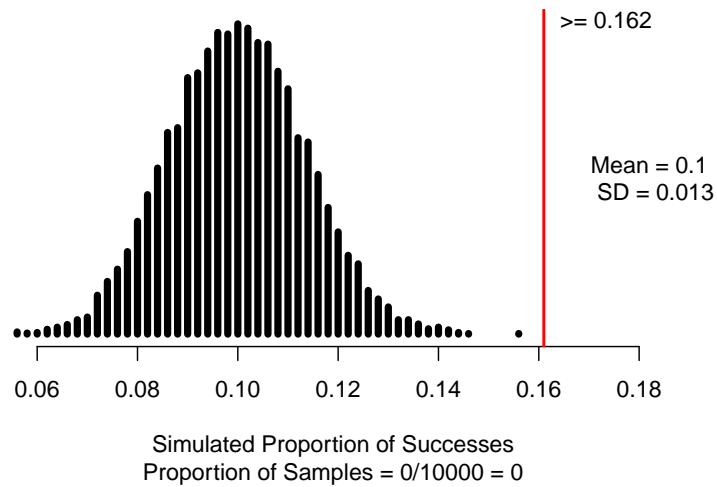$$\text{point estimate} \pm \text{margin of error}$$

$$\hat{p} \pm z^* \times SE(\hat{p})$$

Calculate the 95% confidence interval for the parameter of interest.

6. Interpret the 95% confidence **interval** in the context of the problem.

Simulation methods

We could also use simulation-based methods to analyze these data.

Simulated Proportion of Successes
Proportion of Samples = 0/10000 = 0

>= 0.162

Mean = 0.1
SD = 0.013

2.5 percentile

97.5 percentile

Bootstrapped Proportions
Mean: 0.162, SD: 0.016, 95% CI: (0.13, 0.194)

7. Explain why the results for simulation methods and theory-based methods are similar.

4.3.4 Take-home messages

1. If repeat samples of the same size are selected from the population, approximately 95% of samples will create a 95% confidence interval that contains the parameter of interest.

2. The calculation of the confidence interval uses the standard error calculated using the sample proportion rather than the null value.

4.3.5 Additional notes

Use this space to summarize your thoughts and take additional notes on today's activity and material covered.

4.4 Module 3 and 4 Lab: Mixed Breed Dogs in the U.S.

4.4.1 Learning outcomes

- Determine whether simulation or theory-based methods of inference can be used.

- Analyze and interpret a study involving a single categorical variable.

4.4.2 Mixed Breed Dogs in the U.S.

The American Veterinary Medical Association estimated in 2010 that approximately 49% of dog owners in the U.S. own dogs that are classified as "mixed breed." As part of a larger 2022 international study (Banton 2022) about overall dog health, survey participants were asked, among other things, to report whether their dog was purebred or a mixed breed. Seven hundred and fifty (750) dog owners from the U.S. were recruited to complete an online survey via an email indicating they had been randomly selected by Qualtrics (an "experience management" company that specializes in surveys). Three hundred sixty-four (364) out of 675 respondents from the U.S. reported they owned a mixed breed dog. Is there evidence that, in the last decade, the proportion of dog owners in the U.S. that own a mixed breed dog has changed from the value reported in 2010?

- Observational units:

- Variable:

- Type of variable:

- Success:

Activity intro

- Download the R script file and the data file (US_dogs.csv) from D2L

- Upload both files to D2L and open the R script file

- Enter the name of the dataset for datasetname.csv.

- Highlight and run lines 1 - 6

1. What is the value of the point estimate?

2. Create a plot of the data using the R code. Make sure to include an appropriate title with type of plot, observational units, and variable.

```
dogs %>% # Data set piped into...
    ggplot(aes(x = variable)) +    # This specifies the variable
    geom_bar(aes(y = after_stat(prop), group = 1)) +  # Tell it to make a bar plot with proportions
    labs(title = "Don't forget to title your plot",
        # Give your plot a title
        x = "Breed of Dog",    # Label the x axis
        y = "Relative Frequency")  # Label the y axis
```

Use statistical analysis methods to draw inferences from the data

3. Write out the parameter of interest in words, in context of the study.

4. Write out the null and alternative hypotheses in notation.

H_0 :

H_A :

5. Will theory-based methods give the same results as simulation based methods? Explain your answer.

Null Distribution

To use the computer simulation, we will need to enter the

- assumed "probability of success" (π_0),
- "sample size" (the number of observational units or cases in the sample),
- "number of repetitions" (the number of samples to be generated),
- "as extreme as" (the observed statistic), and
- the "direction" (matches the direction of the alternative hypothesis).

We will use the `one_proportion_test()` function in R (in the `catstats` package) to simulate the null distribution of sample proportions and compute a p-value.

- Using the provided R script file, fill in the values/words for each `xx` in the one proportion test to create a null distribution with 10000 simulations.
- Then highlight and run lines 21–26.

```
one_proportion_test(probability_success = xx, # Null hypothesis value
          sample_size = xx, # Enter sample size
          number_repetitions = 10000, # Enter number of simulations
          as_extreme_as = xx, # Observed statistic
          direction = "xx", # Specify direction of alternative hypothesis
          summary_measure = "proportion") # Reporting proportion or number of successes?
```

6. Report the p-value from the study.

Bootstrap distribution

We will use the `one_proportion_bootstrap_CI()` function in R (in the `catstats` package) to simulate the bootstrap distribution of sample proportions and calculate a confidence interval. Using the provided R script file, fill in the values/words for each `xx` in the one proportion bootstrap confidence interval (CI) code to create a bootstrap distribution with 10000 simulations. Then highlight and run lines ??? to create a 90% confidence interval.

```
one_proportion_bootstrap_CI(sample_size = xx, # Sample size
                    number_successes = xx, # Observed number of successes
                    number_repetitions = 10000, # Number of bootstrap samples to use
                    confidence_level = xx) # Confidence level as a decimal
```

7. Report the 90% confidence interval.

Summarize the results of the study

8. Write a paragraph summarizing the results of the study. Be sure to describe:

- Summary statistic and interpretation
    - Summary measure (in context)
    - Value of the statistic
    - Order of subtraction when comparing two groups
- P-value and interpretation
    - Statement about probability or proportion of samples
    - Statistic (summary measure and value)
    - Direction of the alternative
    - Null hypothesis (in context)
- Confidence interval and interpretation
    - How confident you are (e.g., 90%, 95%, 98%, 99%)
    - Parameter of interest
    - Calculated interval
    - Order of subtraction when comparing two groups
- Conclusion (written to answer the research question)
    - Amount of evidence
    - Parameter of interest
    - Direction of the alternative hypothesis
- Scope of inference
    - To what group of observational units do the results apply (target population or observational units similar to the sample)?
    - What type of inference is appropriate (causal or non-causal)?

Upload a copy of your group's paragraph to Gradescope.

Paragraph (continued):

"Average Driving Distance and Fairway Accuracy." 2008. https://www.pga.com/ and https://www.lpga.com/.

Banton, et al, S. 2022. "Jog with Your Dog: Dog Owner Exercise Routines Predict Dog Exercise Routines and Perception of Ideal Body Weight." *PLoS ONE* 17(8).

Bhavsar, et al, A. 2022. "Increased Risk of Herpes Zoster in Adults ≥50 Years Old Diagnosed with COVID-19 in the United States." *Open Forum Infectious Diseases* 9(5).

Bulmer, M. n.d. "Islands in Schools Project." https://sites.google.com/site/islandsinschoolsprojectwebsite/home.

"Bureau of Transportation Statistics." 2019. https://www.bts.gov/.

"Child Health and Development Studies." n.d. https://www.chdstudies.org/.

Darley, J. M., and C. D. Batson. 1973. ""From Jerusalem to Jericho": A Study of Situational and Dispositional Variables in Helping Behavior." *Journal of Personality and Social Psychology* 27: 100–108.

Davis, Smith, A. K. 2020. "A Poor Substitute for the Real Thing: Captive-Reared Monarch Butterflies Are Weaker, Paler and Have Less Elongated Wings Than Wild Migrants." *Biology Letters* 16.

Du Toit, et al, G. 2015. "Randomized Trial of Peanut Consumption in Infants at Risk for Peanut Allergy." *New England Journal of Medicine* 372.

Edmunds, et al, D. 2016. "Chronic Wasting Disease Drives Population Decline of White-Tailed Deer." *PLoS ONE* 11(8).

Education Statistics, National Center for. 2018. "IPEDS." https://nces.ed.gov/ipeds/.

"Great Britain Married Couples: Great Britain Office of Population Census and Surveys." n.d. https://discovery.nationalarchives.gov.uk/details/r/C13351.

Group, TODAY Study. 2012. "A Clinical Trial to Maintain Glycemic Control in Youth with Type 2 Diabetes." *New England Journal of Medicine* 366: 2247–56.

Hamblin, J. K., K. Wynn, and P. Bloom. 2007. "Social Evaluation by Preverbal Infants." *Nature* 450 (6288): 557–59.

Hirschfelder, A., and P. F. Molin. 2018. "I Is for Ignoble: Stereotyping Native Americans." Retrieved from https://www.ferris.edu/HTMLS/news/jimcrow/native/homepage.htm.

Hutchison, R. L., and M. A. Hirthler. 2013. "Upper Extremity Injuies in Homer's Iliad." *Journal of Hand Surgery (American Volume)* 38: 1790–93.

"IMDb Movies Extensive Dataset." 2016. https://kaggle.com/stefanoleone992/imdb-extensive-dataset.

Kalra, et al., Dl. 2022. "Trustworthiness of Indian Youtubers." Kaggle. https://doi.org/10.34740/KAGGLE/DSV/4426566.

Keating, D., N. Ahmed, F. Nirappil, Stanley-Becker I., and L. Bernstein. 2021. "Coronavirus Infections Dropping Where People Are Vaccinated, Rising Where They Are Not, Post Analysis Finds." *Washington Post.* https://www.washingtonpost.com/health/2021/06/14/covid-cases-vaccination-rates/.

Laeng, Mathisen, B. 2007. "Why Do Blue-Eyed Men Prefer Women with the Same Eye Color?" *Behavioral Ecology and Sociobiology* 61(3).

Levin, D. T. 2000. "Race as a Visual Feature: Using Visual Search and Perceptual Discrimination Tasks to Understand Face Categories and the Cross-Race Recognition Deficit." *Journal of Experimental Psychology* 129(4).

LUETKEMEIER, et al., M. 2017. "Skin Tattoos Alter Sweat Rate and Na+ Concentration." *Medicine and Science in Sports and Exercise* 49(7).

Madden, et al, J. 2020. "Ready Student One: Exploring the Predictors of Student Learning in Virtual Reality." *PLoS ONE* 15(3).

Miller, G. A. 1956. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63(2).

Moquin, W., and C. Van Doren. 1973. "Great Documents in American Indian History." Praeger.

"More Americans Are Joining the 'Cashless' Economy." 2022. https://www.pewresearch.org/short-reads/2022/10/05/more-americans-are-joining-the-cashless-economy/.

National Weather Service Corporate Image Web Team. n.d. "National Weather Service – NWS Billings." https://w2.weather.gov/climate/xmacis.php?wfo=byz.

O'Brien, Lynch, H. D. 2019. "Crocodylian Head Width Allometry and Phylogenetic Prediction of Body Size in Extinct Crocodyliforms." *Integrative Organismal Biology* 1.

"Ocean Temperature and Salinity Study." n.d. https://calcofi.org/.

"Older People Who Get Covid Are at Increased Risk of Getting Shingles." 2022. https://www.washingtonpost.com/health/2022/04/19/shingles-and-covid-over-50/.

"Physician's Health Study." n.d. https://phs.bwh.harvard.edu/.

Porath, Erez, C. 2017. "Does Rudeness Really Matter? The Effects of Rudeness on Task Performance and Helpfulness." *Academy of Management Journal* 50.

Quinn, G. E., C. H. Shin, M. G. Maguire, and R. A. Stone. 1999. "Myopia and Ambient Lighting at Night." *Nature* 399 (6732): 113–14. https://doi.org/10.1038/20094.

Ramachandran, V. 2007. "3 Clues to Understanding Your Brain." https://www.ted.com/talks/vs_ramachandran_3_clues_to_understanding_your_brain.

"Rates of Laboratory-Confimed COVID-19 Hospitalizations by Vaccination Status." 2021. CDC. https://covid.cdc.gov/covid-data-tracker/#covidnet-hospitalizations-vaccination.

Richardson, T., and R. T. Gilman. 2019. "Left-Handedness Is Associated with Greater Fighting Success in Humans." *Scientific Reports* 9 (1): 15402. https://doi.org/10.1038/s41598-019-51975-3.

Stephens, R., and O. Robertson. 2020. "Swearing as a Response to Pain: Assessing Hypoalgesic Effects of Novel "Swear" Words." *Frontiers in Psychology* 11: 643–62.

Stewart, E. H., B. Davis, B. L. Clemans-Taylor, B. Littenberg, C. A. Estrada, and R. M. Centor. 2014. "Rapid Antigen Group a Streptococcus Test to Diagnose Pharyngitis: A Systematic Review and Meta-Analysis" 9 (11). https://doi.org/10.1371/journal.pone.0111727.

Stroop, J. R. 1935. "Studies of Interference in Serial Verbal Reactions." *Journal of Experimental Psychology* 18: 643–62.

Subach, et al, A. 2022. "Foraging Behaviour, Habitat Use and Population Size of the Desert Horned Viper in the Negev Desert." *Soc.Open Sci* 9.

Sulheim, S., A. Ekeland, I. Holme, and R. Bahr. 2017. "Helmet Use and Risk of Head Injuries in Alpine Skiers and Snowboarders: Changes After an Interval of One Decade" 51 (1): 44–50. https://doi.org/10.1136/bjsports-2015-095798.

"Titanic." n.d. http://www.encyclopedia-titanica.org.

"US COVID-19 Vaccine Tracker: See Your State's Progress." 2021. Mayo Clinic. https://www.mayoclinic.org/coronavirus-covid-19/vaccine-tracker.

US Environmental Protection Agency. n.d. "Air Data – Daily Air Quality Tracker." https://www.epa.gov/outdoor-air-quality-data/air-data-daily-air-quality-tracker.

Wahlstrom, et al, K. 2014. "Examining the Impact of Later School Start Times on the Health and Academic Performance of High School Students: A Multi-Site Study." *Center for Applied Research and Educational Improvement.*

Watson, et al., N. 2015. "Recommended Amount of Sleep for a Heathy Adult: A Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society." *Sleep* 38(6).

Weiss, R. D. 1988. "Relapse to Cocaine Abuse After Initiating Desipramine Treatment." *JAMA* 260(17).

"Welcome to the Navajo Nation Government: Official Site of the Navajo Nation." 2011.Retrieved from https://www.navajo-nsn.gov/.

Wilson, Woodruff, J. P. 2016. "Vertebral Adaptations to Large Body Size in Theropod Dinosaurs." *PLoS ONE* 11(7).