P Chapter 1 Projects

Project A: Analyzing a Given Study

Space Tourism?

According to a survey conducted by the Pew Research Center, more US adults say they would not want to orbit the Earth than say they would. Among the 2,541 respondents from March 27-April 9, 2018, 42% say they would definitely or probably be interested in orbiting the Earth in a spacecraft in the future, while roughly 58% say they would not be interested. Among the many different viewpoints explored by the researchers, about one-half of men (51%) say they would be interested in orbiting the Earth in a spacecraft, compared with one-third of women (33%). Interest in being a space tourist is higher among younger generations and men overall. Also, 63% of Millennials (born 1981 to 1996) say they would definitely or probably be interested in space tourism. Only 39% of Gen Xers (born 1965 to 1980) and 27% of those in the Baby Boomer or older generations would be interested. The 58% of US adults who say they wouldn't want to orbit the Earth aboard a spacecraft believe that such a trip would be either "too expensive" (28% of those asked) or "too scary" (28%), or that their age or health wouldn't allow it (28%). Some 16% of those not interested in space travel offered reasons other than the three options in the survey. The survey was part of the American Trends Panel (ATP), a nationally representative panel of randomly selected US adults living in households recruited from landline and cell phone random-digit-dial (RDD) surveys. Panelists participate monthly via self-administered web surveys. Panelists who do not have internet access were provided with a tablet and wireless internet connection.

Source: Strauss, Mark and Kennedy, Brian. "Space tourism? Majority of Americans say they wouldn't be interested." Pew Research Center. 7 June 2018. http://www.pewresearch.org/fact-tank/2018/06/07/space-tourism-majority-of-americans-say-they-wouldnt-be-interested/ (9 Dec. 2018).

Analyze the surveys described above using the following exercises as a guide.

1. Identify the population being studied.
2. Identify the sample and the sample size. Is the sample representative of the population? Explain your answer.
3. Describe how the sample was chosen. Is there any potential bias in the sampling method? Explain your answer.
4. List the descriptive statistics given in the article.
5. What inferences does the article make from the descriptive statistics?
6. Who conducted the studies? Is there any potential researcher bias? Explain your answer.

1

Project B: Analyzing a Study You Find

Find a study of interest on the internet, in a newspaper or magazine, or in an educational journal. If you are reading studies in another course, an excellent choice is to use an assignment from that class. This would also be a good opportunity to begin reading journal articles from your major field of study. The periodical section of the library or library internet resources are wonderful starting places. Once you have found an article, make a copy to include with your analysis.

Note: Though it may be tempting to choose a very short article, if the article is too short, it might not contain all of the necessary elements.

Once you have found a study of interest, analyze it using the following exercises as a guide. Write a formal summary of your analysis.

1. Who conducted the study, when and where?
2. What question(s) does the study seek to answer?
3. Identify the population and variables being studied.
4. Identify the sample and the approximate sample size. Is the sample representative of the population? Explain your answer.
5. How do the researchers deal with the issues of confidentiality and informed consent?
6. Describe how the sample was chosen. Is there any potential bias in the sampling method? Explain your answer.
7. List the descriptive statistics given in the article.
8. Are there any confounding variables?
9. What inferences do the researchers draw from the descriptive statistics?
10. Is there any potential researcher bias?
11. Do you feel comfortable believing the results of the study based on your analysis? Explain your answer.

P Chapter 2 Project

Fast Food Comparison

For three fast-food restaurants, a random sample of service times (in seconds) was taken. The results are to the right.

1. For each restaurant's service times, do the following.

   a. Create a frequency distribution with the given classes. Include the frequency, class boundaries, relative frequency, and cumulative frequency of each class.

Class
30–49
50–69
70–89
90–109
110–129

   b. Use the frequency distribution to construct a histogram of the data.

2. Based on the graphical descriptions of the data you created in part 1, which fast-food restaurant do you believe consistently has the shortest service times? Explain why.

Service Times (in Seconds)

Kim's Kajun Kitchen	Chez Carolyn	Emily's Eatery
111	109	99
94	84	63
57	93	53
80	123	82
78	97	75
109	56	112
92	79	65
34	57	55
67	32	65
122	68	86
95	45	94
46	99	87
103	82	62
110	76	68
93	54	61
86	84	49
75	86	37
49	73	46
61	79	65
82	67	86
76	96	48
92	112	95
65	94	33
92	82	49
94	96	76
106	125	35
112	98	68
83	120	57
72	78	41
119	63	49

P Chapter 3 Projects

Project A: Olympic Gold

The following table gives the years of the Winter Olympics and the number of gold medals won in each year by the United States through 2014.

USA Gold Medal Count

Year	1924	1928	1932	1936	1948	1952	1956	1960	1964	1968	1972	1976
Number of Gold Medals	1	2	6	1	3	4	2	3	1	1	3	3
Number of Events	16	14	14	17	22	22	24	27	34	35	35	37

Year	1980	1984	1988	1992	1994	1998	2002	2006	2010	2014	2018
Number of Gold Medals	6	4	2	5	6	6	10	9	9	9	?
Number of Events	38	39	46	57	61	68	78	84	86	98	102

Source: Olympic.org. "Olympic Games." http://www.olympic.org/olympic-games (4 April 2019).

Analyze the data in the table by calculating descriptive statistics for the USA gold medal count. We will use these statistics to estimate the expected number of gold medals for the United States in the 2018 Winter Games. Note that the number of events is not the same for every year. Therefore, it will be necessary to take into account the number of gold medals as a percentage of the number of events.

1. Begin by calculating the number of gold medals won each year as a percentage of the number of events. To do this, divide the number of gold medals by the number of events, and then multiply by 100. Round your answers to the nearest whole percentage. Create a table similar to the one above with the percentage of gold medals won in each year.
2. Calculate the range of the number of gold medals and the range of the percentage of gold medals.
3. Calculate the median number of gold medals and the median percentage of gold medals.
4. Calculate the mode of the number of gold medals and the mode of the percentage of gold medals.
5. Calculate the mean number of gold medals won.
6. Calculate the mean percentage of gold medals won by adding to get the total number of gold medals won in all years, dividing by the total number of events in all years and then multiplying by 100. Round your answer to the nearest whole percentage. (Note that you cannot average percentages. This is why we have to go back to the numbers of gold medals and the numbers of events.)
7. Calculate the five-number summary for the number of gold medals won.
8. Draw a box plot for the number of gold medals won.
9. Using a value of 102 events in the 2018 Olympics, estimate the expected number of gold medals by multiplying 102 by the median percentage of gold medals won by the United States.
10. The United States actually won 9 gold medals in the 2018 Winter Games in Pyeongchang, South Korea. How does your calculation for the expected number of gold medals compare to the actual number?

Project B: Where Would You Invest Your Money?

Let's look at several sets of stock prices. The following prices were obtained from historical records from 2019, and are listed in dollars per share.

Stock Prices

Coca-Cola (KO)	Facebook (FB)	General Electric (GE)
46.38	166.69	10.01
46.18	165.55	10.10
46.57	165.87	10.24
46.72	167.68	10.10
46.86	166.29	9.99
46.58	164.34	9.89
46.61	166.08	9.96
46.64	165.44	10.10
46.03	161.57	9.88
45.93	160.47	9.98
45.51	165.98	10.27
45.53	170.17	10.22

Source: Yahoo! Finance. http://finance.yahoo.com (4 April. 2019).

1. Find the mean of each set of stock prices.
2. Find the median of each set of stock prices.
3. Calculate the variance and standard deviation of each set of stock prices.
4. If you have $10,000 to invest, what stock would you buy under the following circumstances? Justify your reasoning.
   a. You are nearing retirement and need a stable investment for the future.
   b. You are a wealthy entrepreneur hoping to make a large profit in a short amount of time.

P Chapter 4 Project

Law of Large Numbers

This project is designed to be completed in groups, but it may be completed by an individual. Each group is to follow the steps below using a standard six-sided die.

Step 1: Calculate the probability of rolling a single die and getting a four. Round the probability to four decimal places.

$$\text{Probability} = \frac{\#\ \text{of ways of getting a four}}{\#\ \text{of ways to roll a die}} = \boxed{}$$

Step 2: Each group member is to roll a die ten times and record the outcomes. Compile the outcomes for all group members, and compute the proportion of the group's rolls that were fours. Round the proportion to four decimal places.

$$\text{Proportion of group rolls that were fours} = \frac{\text{Total \# of group rolls that were fours}}{\text{Total \# of group rolls}}$$

$$= \boxed{}$$

Step 3: This time, each group member is to roll a die an additional 40 times, for a total of 50 rolls per person. Again, combine the outcomes for all group members, and compute the proportion of the group's rolls that were fours. Round the proportion to four decimal places.

$$\text{Proportion of group rolls that were fours} = \frac{\text{Total \# of group rolls that were fours}}{\text{Total \# of group rolls}}$$

$$= \boxed{}$$

Step 4: Let's combine the information from all groups in order to look at the results for the class as a whole. On the board, make a chart with one column for "number of fours" and one for "number of rolls." Fill in the chart with the information from each group.

Number of Fours	Number of Rolls

Using this information, calculate the proportion of all rolls for the class that were fours. Round the proportion to four decimal places.

$$\text{Proportion of class rolls that were fours} = \frac{\text{Total \# of class rolls that were fours}}{\text{Total \# of class rolls}}$$

$$= \boxed{}$$

Step 5: Let's evaluate the results of this experiment. Compare the proportions calculated in Steps 2, 3, and 4. You should see that as the number of times the dice were tossed increases, the proportion of fours rolled becomes closer to the probability calculated in Step 1. This is precisely what the Law of Large Numbers says: the greater the number of trials, the closer the experimental probability comes to the classical probability. In fact, if it were possible to complete an infinite number of die tosses, the proportion of all rolls that were fours would indeed be equal to the classical probability.

P Chapter 5 Project

Playing Roulette

We all dream of winning big, becoming an instant millionaire; but how likely is that? Let's say we decide to pursue our goal of winning big money by going to a casino and continually playing what we think will be an easy game: roulette. Can we expect to win big in the long run? Is one bet better than another? How do the casinos make so much money anyway? If we are betting against the casino, how do they make sure that they always win? This project will help you answer these questions.

Let's begin with a lesson in roulette. Roulette is a casino game that involves spinning a ball on a wheel that is marked with numbered squares that are red, black, or green. Half of the numbers 1–36 are colored red and half are black and the numbers 0 and 00 are green. Each number occurs only once on the wheel.

We can make many different types of bets, but two of the most common are to bet on a single number (1–36) or to bet on a color (either red or black). These will be the two bets we will consider in this project. After all players place their bets on the table, the wheel is spun and the ball tossed onto the wheel. The pocket in which the ball lands on the wheel determines the winning number and color. The ball can land on only one color and number at a time.

We begin by placing a bet on a number between 1 and 36. This bet pays 36 to 1 in most casinos, which means we will be paid \$36 for each \$1 we bet on the winning number. If we lose, we simply lose whatever amount of money we bet.

1. Calculate the probability that we will win on a single spin of the wheel.
2. Calculate the probability that we will lose.
3. If we bet \$8 on the winning number, how much money will we win?
4. What is the expected value of a bet on a single number if we bet \$1?
5. For a \$5 bet, what is the expected value of a bet on a single number?
6. What is the expected value of a bet on a single number if we bet \$10?
7. Do you see a pattern in the answers to the last three questions?

We decide that we can certainly increase our chances of winning if we bet on a color instead of a number. Roulette allows us to bet on either red or black and if the number is that color, we win. This bet pays even money in most casinos. This means that for each dollar we bet, we will win \$1 for choosing the winning color. So, if we bet \$5 and win, we would keep our \$5 and win \$5 more. If we lose, we lose whatever amount of money we bet, just as before.

8. What is the probability that we will win on a single spin if we bet on red?

9. What is the probability that we will lose on a single spin if we bet on red?

10. If we bet \$60 on the winning color, how much money will we win? Is this more or less than we will win by betting \$8 on our favorite number? Explain why.

11. What is the expected value of a bet on red if we bet \$1?

12. For a \$5 bet, what is the expected value of a bet on red?

13. What is the expected value of a bet on red if we bet \$10?

14. Do you see a pattern in the answers to the last three questions?

15. How does the expected value of betting on a number compare to the expected value of betting on a color? Is one bet more profitable than another?

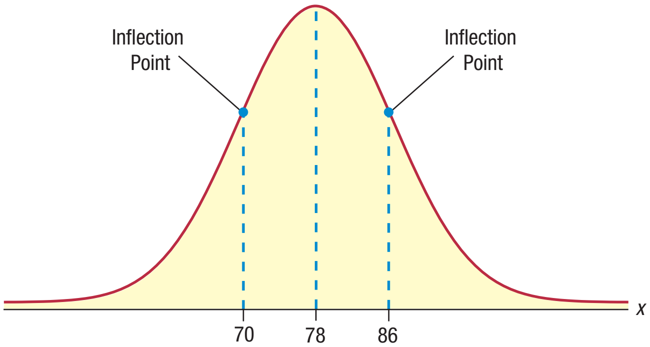
16. If our goal was to play roulette so that we can "win it big," what does the expected value of a bet tell us about our chances of winning a large amount of money?

17. Are the casinos really gambling when we place a bet against them? Explain.

P Chapter 6 Project

Curving Grades Using a Normal Distribution

Dr. Smith, a biology professor at Bradford University, has decided to give his classes a standardized biology exam that is nationally normed. This indicates that the normal distribution is an appropriate approximation for the probability distribution of students' scores on this exam. The probability distribution of students' scores on this standardized exam can be estimated using the normal distribution shown below.



1. State the mean of the distribution of the biology exam scores.

2. State the standard deviation of the distribution of the biology exam scores.

Grading Curve Option I

Originally, Dr. Smith decides to curve his students' exam grades as follows.

- Students whose scores are at or above the 90th percentile will receive an A.
- Students whose scores are in the 80th–89th percentiles will receive a B.
- Students whose scores are in the 70th–79th percentiles will receive a C.
- Students whose scores are in the 60th–69th percentiles will receive a D.
- Students whose scores are below the 60th percentile will receive an F.

3. Find the z-scores that correspond to the following percentiles.

- 90th percentile
- 80th percentile
- 70th percentile
- 60th percentile

4. Using that information, find the exam scores that correspond to the curved grading scale. Assume that the exam scores range from 0 to 100. (Round to the nearest whole number.)

A: __________–100

B: __________–__________

C: __________–__________

D: __________–__________

F: 0–__________

5. The following is a partial list of grades for students in Dr. Smith's class. Using the grading scale you just created, find the new curved *letter grades* that the students will receive on their tests given their raw scores.

Biology Exam Grades

Name	Raw Score / Grade
Adam	82 / B
Bill	77 / C
Susie	91 / A
Troy	86 / B
Sharon	75 / C
Laura	66 / D
Eric	88 / B
Marcus	69 / D
Stephanie	79 / C

Grading Curve Option II

After reviewing the results, Dr. Smith decides to consider an alternate curving method. He decides to assign exam grades as follows.

- A: Students whose scores are at least two standard deviations above the mean of the standardized test.
- B: Students whose scores are from one up to two standard deviations above the mean of the standardized test.
- C: Students whose scores are from one standard deviation below the mean up to one standard deviation above the mean of the standardized test.
- D: Students whose scores are from two standard deviations below the mean up to one standard deviation below the mean of the standardized test.
- F: Students whose scores are more than two standard deviations below the mean of the standardized test.

6. Using the previous information, create Dr. Smith's new grading scale. (Round to the nearest whole number.)

A: __________ – 100

B: __________ – __________

C: __________ – __________

D: __________ – __________

F: 0 – __________

7. Using the grading scale you just created, return to the partial list of grades and find the new curved *letter grades* that the students will receive on their tests given their raw scores.

8. Review the grades each student received using the two grading scales. Which grading scale do you feel is fairer? Explain why.

P Chapter 7 Projects

Project A: Central Limit Theorem Experiment

You will need a standard six-sided die and at least six sets of data to complete this project.

Consider the distribution of the possible outcomes from rolling a single die; that is, 1, 2, 3, 4, 5, and 6. Let's use this distribution as our theoretical population distribution. We want to use this population distribution to explore the properties of the Central Limit Theorem. Let's begin by determining the shape, center, and dispersion of the population distribution.

1. What would you expect the distribution of the outcomes from repeated rolls of a single die to look like; in other words, what is its shape? (**Hint:** What is the probability of getting each value?)

   Shape: ____________

2. Calculate the mean of the population. (**Hint:** What is the mean outcome for rolling a single die?)

   $\mu =$ ________________

3. Calculate the standard deviation of the population. (**Hint:** What is the standard deviation of all possible outcomes from rolling a single die?)

   $\sigma =$ ________________

Let's continue by exploring the distribution of the original population empirically. To do so, follow these steps.

Step 1: Roll your die 60 times and record each outcome.

Step 2: Combine your results with at least two other students and tally the frequency of each roll of the die from the combined results. Record your results in a table similar to the following.

Outcome	Frequency
1	
2	
3	
4	
5	
6	

Step 3: Draw a bar graph of these frequencies.

Step 4: Does the distribution appear to be a normal distribution? Is this what you expected from question 1?

The Central Limit Theorem is not about individual rolls like we just looked at, but is about the averages of sample rolls. Thus we need to create samples in order to explore the properties of the Central Limit Theorem.

Step 5: Return to your original data from Step 1. To create samples from your data you can group the rolls into sets of 10. For each sequence of 10 rolls, calculate the mean of that sample. Round your answers to one decimal place. (You should have six sample means.)

Step 6: Combine your sample means with those of as many of your classmates as you can. Record the sample means of each of your classmates' six samples.

Step 7: Tally the frequencies of the sample means from your combined results in a table like the one that follows.

Sample Mean	Frequency
1.0–1.2	
1.3–1.5	
1.6–1.8	
1.9–2.1	
2.2–2.4	
2.5–2.7	
2.8–3.0	
3.1–3.3	
3.4–3.6	
3.7–3.9	
4.0–4.2	
4.3–4.5	
4.6–4.8	
4.9–5.1	
5.2–5.4	
5.5–5.7	
5.8–6.0	

Step 8: Draw a histogram of the sample means.

Step 9: What is the shape of this distribution?

Step 10: What is the mean of your sample means? (**Hint:** Use the sample means you collected in Step 6.)

$\mu_{\bar{x}}$ = __________

How does $\mu_{\bar{x}}$ compare to μ from question 2?

Step 11: What is the standard deviation of the sample means? (Again, go back to the sample means you collected in Step 6 and use a calculator or statistical software.)

$\sigma_{\bar{x}}$ = __________

How does $\sigma_{\bar{x}}$ compare to σ from question 3?

Since our samples were groups of 10 rolls, $n = 10$. Using σ from question 3, calculate $\frac{\sigma}{\sqrt{n}}$.

$\frac{\sigma}{\sqrt{n}}$ = __________

How does $\sigma_{\bar{x}}$ compare to $\frac{\sigma}{\sqrt{n}}$?

The Central Limit Theorem says that the distribution of the sample means should be closer to a normal distribution when the sample size becomes larger. To see this effect, group your original data from Step 1 into two samples of 30 rolls instead of six sets of 10.

Repeat Steps 5–11 using the new sample size of $n = 30$.

Step 12: Do your results seem to verify the Central Limit Theorem?

Project B: Sampling Distribution Simulation

In the Hawkes Learning courseware, *Beginning Statistics*, open Lesson 7.1, Introduction to the Central Limit Theorem. This lesson is a simulation designed to help you better understand sampling distributions as well as the Central Limit Theorem. Begin the simulation by choosing a parent distribution. Select the Settings icon at the top right of the screen, choose Distribution, then pick a distribution from the menu and press OK. If you do not have a preference, the computer will automatically begin with a uniform parent distribution. For each iteration, the computer randomly chooses 30 numbers from the parent distribution and displays them in the parent histogram. At the same time, the computer chooses samples of size 5, 15, and 30 from those same 30 numbers. The mean of each sample is calculated and then displayed in its respective histogram. Answer the following questions.

1. How many numbers are displayed in the parent histogram?
2. How many numbers are displayed in each of the sampling distribution histograms?

Click **Next** to obtain a second iteration.

3. After the second iteration, how many numbers are displayed in the parent histogram?
4. After the second iteration, how many numbers are displayed in each of the sampling distribution histograms?

Now let's see what happens after many iterations. Click **Auto** and let the simulation run until about 100 iterations have passed. Click the same button, which now says **Stop**, to stop the process.

5. How many numbers are displayed in the parent histogram? (This number will vary, depending on how many iterations have passed.)
6. How many numbers are displayed in each of the sampling distributions?
7. Which, if any, of the sampling distributions appear to have a normal shape?

Finally, click **Auto** again and allow the program to process at least 1500 iterations before clicking **Stop**.

8. Which of the sampling distributions appear to have a normal shape?
9. Compare the means of sample means listed in the table. Do these numbers behave as you would expect?
10. Compare the standard deviations of sample means listed in the table. Do these numbers behave as you would expect?

Take some time to explore the other parent distributions available. Although the shape of the parent histogram will differ with each parent used, the Central Limit Theorem will always remain constant.

P Chapter 8 Projects

Project A: Constructing a Confidence Interval for a Population Mean

Choose a study question and follow the given steps to construct a confidence interval for a population mean.

Choose one of the study questions below, or write your own study question for which you could collect data from students on your campus to estimate a population mean.

- How many hours per week do college students on your campus study?
- What is the mean number of parking tickets college students on your campus receive per semester?
- What is the mean price of rent paid per month by college students on your campus?

For the study question you chose, construct a confidence interval for that parameter by doing each of the following.

1. Collect data on your study question from at least 30 students on your campus.
2. Calculate the sample mean and sample standard deviation of your sample data.
3. Calculate the margin of error for a 95% confidence interval using your sample statistics.
4. Construct a 95% confidence interval using your sample statistics.
5. Write a summary of your results.

Project B: Constructing a Confidence Interval for a Population Proportion

Choose a study question and follow the given steps to construct a confidence interval for a population proportion.

Choose one of the study questions below, or write your own study question for which you could collect data from students on your campus to estimate a population proportion.

- What percentage of college students on your campus own a smartphone?
- What percentage of college students on your campus are seeking a degree in your field of study?
- What percentage of college students on your campus live in on-campus housing?

For the study question you chose, construct a confidence interval for that parameter by doing each of the following.

1. Collect data on your study question from at least 30 students on your campus.
2. Calculate the sample proportion from your sample data.
3. Calculate the margin of error for a 90% confidence interval using your sample statistics.
4. Construct a 90% confidence interval using your sample statistics.
5. Write a summary of your results.

P Chapter 9 Projects

Project A: Screentime Challenge

Students sometimes get a bad reputation for the amount of time they spend on their smartphones. But what about the screen times of faculty and staff? Do you think students really have higher screen time averages than the faculty and staff at their schools? This project will allow you to investigate this claim for yourself.

Step 1: For this project you will need 30 volunteers, with approximately half being students and half being faculty or staff. (It doesn't have to be exactly 15 each.) Ask each volunteer to look up *yesterday's* total screen time as recorded on their smartphone. It's important that the screen times recorded are all the same day of the week for consistency. It is also necessary to record a full day of screen time, which is why you cannot use the screen time for the current day, which would be incomplete. Keep your data organized in a chart.

Step 2: Divide the results into two groups, students and faculty/staff. Compute the mean and sample standard deviation of each group. Record your statistics below.

$n\_\text{students} =$ ________________ $n\_\text{faculty / staff} =$ ________________

$\bar{x}\_\text{students} =$ ________________ $\bar{x}\_\text{faculty / staff} =$ ________________

$s\_\text{students} =$ ________________ $s\_\text{faculty / staff} =$ ________________

Step 3: Construct a 95% confidence interval for the true difference between the mean screen time of students and faculty/staff. Assume that the population variances are not the same and that the population distributions of screen times are approximately normal for both students and faculty/staff.

Step 4: Consider your results and write up your conclusion. Does your study show that students average more screen time than faculty and staff? Are your results convincing? Why or why not? Consider other factors such as how the sample was chosen, possible sources of bias, and the limits of interpreting the confidence interval when you make your conclusion.

Project B: Knowledge of Historical Dates

Many adults assume that college students don't know their history. Let's do an experiment to see if a simple crash course can improve college students' knowledge of important historical dates. This experiment will involve some research and preparation on your part.

Step 1: Begin by preparing a set of flashcards with the following historical events and their corresponding dates. Each flashcard should have one event on the front of the card and its date on the back. If you wish, you can even include pictures or colors to help distinguish the different historical events.

Historical Event	Event Date
Fall of Rome	476 AD
Dark Ages	500 - 1000 AD
US Civil War	1861 - 1865
Moon Landing	1969
Boston Tea Party	1773
September 11	2001
Invention of the Telegraph	1844
WWI	1914 - 1918
Pearl Harbor	1941
Great Depression	1929 - 1939
Columbus Discovers the New World	1492
Wright Brothers First Flight	1903

Step 2: Next, you will need to find 10 willing college student volunteers. You will give each volunteer a pretest and a posttest using the flashcards. For each volunteer, show them the historical events on the fronts of the flashcards and ask them to name the date(s) on the back. Record how many answers he gets correct out of 12. After the pretest, give them the answers. Coach them for five minutes on the historical events and dates. Next, visit with them for a few minutes about the class, what you are doing in this project, the weather, your favorite sports team, and so forth. Then, shuffle the flashcards and quiz your participant a second time. Record how many answers the participant gets correct out of 12. Thank them for their help with the project.

Step 3: Once data have been collected from 10 college students, calculate the paired difference for each student by subtracting the pretest score from the posttest score.

Pretest Score, x										
Posttest Score, y										
$d_i = y_i - x_i$										

Step 4: Calculate each of the following sample statistics.

$\bar{d} =$ ____________________ $s_d =$ ____________________

Step 5: Construct a 95% confidence interval for the true mean difference between the students' scores on the posttest and pretest. Assume that the population distribution of the paired differences is approximately normal.

Step 6: Did your "crash course" on historical dates improve the students' knowledge? Give a conclusion for your study and discuss what these results mean. Consider how you chose the participants in your sample and whether their initial knowledge of the events on your flashcards impacted your results.

P Chapter 10 Projects

Project A: Hypothesis Testing for Population Means

Choose one of the three claims, collect data from members of the appropriate population, and perform a hypothesis test to determine if the evidence supports the claim or does not support the claim. After you have written your conclusion, look at the "real" value of the population mean (given below) and determine if your hypothesis test produced a correct decision, a Type I error, or a Type II error.

Pick one of the following claims to test:

Claim 1: Parents of college freshmen believe that freshmen spend a mean of at most $50 per week on eating out.

Claim 2: It is believed that college sophomores see at least 4 movies per month at the theater.

Claim 3: The librarian claims that college juniors visit the library twice a week.

Step 1: State the null and alternative hypotheses.

Based on the claim you chose, what are the null and alternative hypotheses?

Step 2: Determine which distribution to use for the test statistic, and state the level of significance.

In the next step, you will collect data from 10 students. Assuming that the population distribution is approximately normal, what formula should be used for the test statistic? Also, choose a level of significance of 0.10, 0.05, or 0.01.

Step 3: Gather data and calculate the necessary sample statistics.

Collect data from 10 students who are in the appropriate population. Discuss which method of data collection you used. List any potential for bias. Calculate the sample mean and sample standard deviation.

Calculate the test statistic using the values you just calculated from your sample.

Step 4: Draw a conclusion and interpret the decision.

Determine the type of your hypothesis test: left-tailed, right-tailed, or two-tailed.

Draw a picture of your rejection region.

What is your conclusion?

Interpret your decision.

Types of Errors

Let's assume we find out that the truths are as follows.

- Freshmen spend a mean of $40 per week eating out.
- The mean number of movies seen by sophomores at the theater each month is more than 4.
- The mean number of times per week that juniors visit the library is 4.

Based on your conclusion, did you make a Type I error, a Type II error, or a correct decision? Explain.

Project B: Chi-Square Test for Goodness of Fit

In this project, we will look at whether the makeup of your institution has changed significantly from Year 1 to Year 2, for two nonconsecutive years. In other words, are the percentages of students in every classification in Year 2 equal to the percentages of students in every classification in Year 1? Let's begin by collecting some data from the earlier academic year to determine the hypotheses for the more recent academic year.

1. Choose two years that are not consecutive from which to collect data. (For instance, you may choose the current academic year for Year 2 and the academic year four years earlier for Year 1.) Find out the number of students who were enrolled at your institution for each classification during Year 1 and enter them in a table like the one below.

	Year 1
Number of Freshmen Enrolled	
Number of Sophomores Enrolled	
Number of Juniors Enrolled	
Number of Seniors Enrolled	
Total Number of Students Enrolled	

2. Now calculate the percentage of students in each category during Year 1 and enter them in a new table.

	Year 1
Percentage of Freshmen Enrolled	
Percentage of Sophomores Enrolled	
Percentage of Juniors Enrolled	
Percentage of Seniors Enrolled	

3. State the null and alternative hypotheses in words.
4. Specify the null and alternative hypotheses with mathematical symbols.
5. Now let's gather data for Year 2. Find out the number of students enrolled at your institution for each classification during Year 2 and enter them in a table like the one below.

	Year 2
Number of Freshmen Enrolled	
Number of Sophomores Enrolled	
Number of Juniors Enrolled	
Number of Seniors Enrolled	
Total Number of Students Enrolled	

6. Now calculate the percentage of students in each category during Year 2 and enter them in a new table.

	Year 2
Percentage of Freshmen Enrolled	
Percentage of Sophomores Enrolled	
Percentage of Juniors Enrolled	
Percentage of Seniors Enrolled	

7. Calculate the expected value for each classification.

	Expected Values
Number of Freshmen Enrolled	
Number of Sophomores Enrolled	
Number of Juniors Enrolled	
Number of Seniors Enrolled	

8. Calculate the test statistic.
9. Determine your conclusion.
10. Is there sufficient evidence to conclude at $\alpha = 0.10$ that the makeup of your institution has significantly changed between Year 1 and Year 2? Explain.

P Chapter 11 Projects

Project A: Hypothesis Testing for Two Population Parameters

Choose one of the three questions to answer, collect data from members of the appropriate population, and perform a hypothesis test to answer the question. After you have written your conclusion, look at the "truth" and determine if your hypothesis test produced a correct decision, a Type I error, or a Type II error.

Pick one of the following questions to test:

- Is there a difference in study habits for college students who are 25 and older versus those students who are aged 18-24? Ask at least 30 students 25 and older and at least 30 students aged 18-24 to estimate the amount of time they spend studying each week. Assume that the population variances are equal. Use a 0.10 level of significance.
- Do college freshmen spend less money each week eating out than seniors? Ask between 10 and 20 freshmen and between 10 and 20 seniors to estimate the amount of money they spend each week eating out. Assume that the population variances are different and both population distributions are approximately normal. Use a 0.05 level of significance.
- Are the percentages of students and faculty who exercise regularly the same? Ask at least 30 students and at least 30 faculty if they exercise at least three times per week. Record the number of students and the number of faculty who say "yes." Use a 0.01 level of significance.

Step 1: State the null and alternative hypotheses.

What are the null and alternative hypotheses?

Step 2: Determine which distribution to use for the test statistic, and state the level of significance.

Based on the description of the test you chose, what formula should be used for the test statistic? Also, state the level of significance for your hypothesis test.

Step 3: Gather data and calculate the necessary sample statistics.

Collect data on the claim from the appropriate populations. Discuss which method of data collection you used. List any potential for bias. Calculate the sample statistics needed in order to compute the test statistic.

Calculate the test statistic using your sample statistics.

Step 4: Draw a conclusion and interpret the decision.

Determine the type of your hypothesis test: left-tailed, right-tailed, or two-tailed.

State the decision rule in terms of either the p-value or the rejection region for the test statistic.

What is your conclusion? Be sure to answer the original question.

Types of Errors

Suppose the “truths” are as follows.

- The mean amount of time spent studying each week is higher for students 25 and older in college than for college students aged 18-24.
- The mean amount of money spent eating out each week is higher for freshmen than for seniors.
- The percentage of students who exercise at least three times per week is the same as the percentage of faculty.

Based on your conclusion, did you make a Type I error, Type II error, or a correct decision? Explain.

Project B: ANOVA

A lack of adequate parking is one of the most common complaints of students on any college campus. Is there a difference in the mean number of parking tickets received in one semester by students who commute to campus, students who live in residential housing on campus, students who live in fraternity houses, and students who live in sorority houses? Let's perform a one-way ANOVA test to help us answer this question.

To begin, label the populations as follows.

Population 1: Students who commute to campus

Population 2: Students living in residential housing on campus

Population 3: Students living in fraternity houses

Population 4: Students living in sorority houses

Step 1: State the null and alternative hypotheses.

What are the null and alternative hypotheses?

Step 2: Determine which distribution to use for the test statistic, and state the level of significance.

Assuming that the population distributions are all approximately normal and the population variances are all equal, what formula should be used for the test statistic? Also, choose a level of significance of 0.10, 0.05, or 0.01.

Step 3: Gather data and calculate the necessary sample statistics.

Collect data from five students in each population. Record the number of parking tickets that each student received last semester in a table similar to the one below.

Parking Tickets

Commuters	Residence Hall	Fraternity House	Sorority House

Use the formulas given in Section 11.6, or available technology (as described in Example 11.6.2), to complete a one-way ANOVA table.

Step 4: Draw a conclusion and interpret the decision.

a. Draw a picture of your rejection region, labeling the critical value, or state the rejection rule for *p*-values depending on the method chosen to draw a conclusion.

b. Based on the calculated value of *F* from the ANOVA table, should you reject the null hypothesis?

c. Is there a difference in the mean number of parking tickets received by students who commute to campus, students who live in residential housing on campus, students who live in fraternity houses, and students who live in sorority houses?

d. Does your hypothesis test provide any indication that one of the populations receives more parking tickets than another? Explain.

P Chapter 12 Project

Arm Span vs. Height

Is it true that a person's arm span is equal to the person's height? If so, then there should be a perfect linear relationship between arm span and height. Let's find out if this is true by collecting data and creating a linear regression model.

Step 1: Collect data from 10 people. Measure each person's height (in inches) and then the person's arm span (in inches), which is the distance from fingertip to fingertip as the person's arms are outstretched. Record your data in a table similar to the one below. Your results will be more generalizable if you collect data from people of many different ages and heights.

Height										
Arm Span										

Step 2: Create a scatter plot of the data. Use height as the x-variable and arm span as the y-variable.

a. Does there appear to be a linear relationship between x and y?

b. Is the relationship positive or negative?

c. Is the relationship strong or weak?

d. Does the graph seem to support the claim that a person's height is equal to the person's arm span?

Step 3: Calculate the correlation coefficient, r.

a. Is the correlation coefficient statistically significant at the 0.05 level of significance?

b. What about the 0.01 level of significance?

c. Interpret what it means for the correlation coefficient to be statistically significant for this scenario.

Step 4: Determine the equation of the regression line, $\hat{y} = b_0 + b_1 x$.

a. Draw the regression line on your scatter plot.

b. If it is true that a person's arm span is the same as the person's height, what would you expect the slope of the regression line to be?

c. How close is the actual slope of the regression line to the expected value?

d. Calculate a 95% confidence interval for the slope of the regression line.

e. Based on your confidence interval, could you conclude that it is likely that a person's arm span is equal to the person's height?

Step 5: Calculate the coefficient of determination, r^2.

a. Interpret what r^2 means for this scenario.

b. Does the value of r^2 support the claim that arm span is equal to height?

Step 6: Based on your answers for Steps 1 through 5, formulate a conclusion as to whether a person's arm span is equal to the person's height.