SIT Session Lesson Plan

Week/Chapter: Week 1 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on this week? New Vocabulary, Frequency vs. Relative Frequency, and Charts/Graphs

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time

Opening/Warm-Up/Ice-breaker	**M&M game:**
	• 1. Get a bag of M&Ms or Skittles (anything with color) and have the students grab a handful.
	• 2. Assign categories or questions with each color.
	  o Ex) red- what is your favorite color?
	  o Ex) green: any topic that deals with food.
	  o Ex) blue: say your favorite _____ (any favorite)
	• 3. Go around the table and have the students pick one color randomly (eyes closed) and with whatever color they get answer the category or question.
	• I am going to give each student 1 of each color so this does not take too long.
	  o **Red:** What is your name, Age, major?
	  o **Blue**: Something about Family ☺ (It could be about pets too!)
	  o **Green**: Random!!! ANYTHING! Keep it PG 13 please >///<
	  o **Orange**: Something we cannot tell about you by looking at your appearance.
	  o **Brown**: What are your Hobbies?
	  o **Yellow**: Favorite memory!
		5-10 minutes

Challenging Content 1	**Vocabulary**!!! There seems to be all new vocabulary that we must know! I will have them write this list down on a piece of paper and we will do a scavenger hunt through the book and our minds and our notes to remember what the definitions were! Also we will write examples of the vocabulary word if it is applicable! • Also ask if anyone can come up with a fun idea to remember the definitions! • Vocab words: o **Statistics**- a collection of procedures/ principles to gather data and analyze information to help make uncertain decisions. ▪ HINT* IMPORTANT FOR TEST!! o **Raw data**- a collection of numbers or categories that have not been processed ▪ Ex: what you did during class on the first day in sample groups was forming RAW DATA! • So raw data in that survey was : o Male or female? : ex) raw data- male o **Variable**- a characteristic that is different among individuals o **Sample data**- collection of data from a population o **Population data**- collection of all measured individuals o **Parameter vs**. **Statistic**- ▪ **Parameter**: summary of population data ▪ **Statistic**: summary of sample data o **Categorical variables**- also called qualitative o **Qualitative vs**. **Quantitative**- ▪ **Qualitative**: group or categorical names that do not measure or have logical ordering • Ex: eye color, place of birth ▪ **Quantitative**: numerical values taken from individuals • Ex: weight, height	15-20 minutes
Challenging Content 2	**Summarizing One or Two Categorical Variables (2.3)** • **Frequency vs. Relative Frequency**: o **Frequency**: categorical variable listing with each frequencies. (like counting) ▪ Ex: learned to ride bike at age… • 7yrs: 6 people • 8yrs: 10 people • 9yrs: 4 people	15-20 minutes

	○ **Relative frequency**: like frequency but instead with relative frequencies next to the category (usually in percentages or proportions) ▪ Ex: back to the bike, let's say the total number was 20 students • 7 yrs: 6 ppl (3/10 or 30 %) • 8 yrs: 10 ppl (½ or 50%) • 9 yrs: 4 ppl (1/5 or 20 %) • **Pie Charts & Bar graphs**! ○ Look and find examples and make a simple bar graph & pie chart together! ▪ Ask students to write down their favorite color out of these 5 colors: • White, Blue, Pink, Green, Black • (reduce amount of color if not enough people) ▪ Then gather frequencies and make a chart with relative frequencies! ▪ Then make a pie chart and a Bar graph together! ▪ If the students can do this exercise, then they have shown full understanding of difficult content 2!	
Closing/Check for Understanding	• What does it mean to be a college student? ○ How is this different than high school student? • Good study tool ideas: ○ Evernote Application (Great for taking notes!) ○ Cornell notes! ○ Time management! (Learn to manage time and make to-do lists!) ○ Keep up to date with homework and tests and projects and papers!	5-10 minutes
Tutoring = Q & A	Question Time!	5-10 minutes

After session comments/thoughts:

References

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

SIT Session Lesson Plan

Week/Chapter: Week 2 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on today? My two difficult concepts for this week are standard deviation and z score.

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
Warm up to see if the students remembered what they learned last week	Grab bag of examples and give definition	10 mins
Concept 1: Standard Deviation	First Line Only: I will give the group a problem and each student will do one step until the solve it together	15 mins
Concept 2: Z Score	Rally Race: I will split them up into 2 groups and then present them with questions for them to answer, where they are going to race up to the board to complete it	15 mins
Closing	KWL	5 mins
Tutoring = Q & A		5-10 minutes

After session comments/thoughts:

Nationality (l)	**Softness of a cat**
Level of education	**Number of boys and girls in your classroom**
Number of children in a household	**Majors of students**
Household income	**Number of a jersey**
Intake of grains	**Phone numbers**

Graphing data (n)	Nation of origin

Grab bag (Utts & Heckard, 2012)

Raw data
Variable
Observational unit
Sample size
Dataset
Sample data
Population data
Statistic
Parameter
Categorical variable
Ordinal variable
Quantitative variable

WK2: Standard Deviation: show if the data point is close or far from the mean or if it is dispersed (measure of spread)

Find the population standard deviation. (Utts & Heckard, 2012)

82,77,90,71

X_i	X_i^2
82	82^2 = 6724
77	77^2 = 5929
90	8100
71	5041
EX_i = 320	EX_i^2 = 25794

ANSWER: 6.964

- Find the sample population standard deviation (Utts & Heckard, 2012)

82,77,90,71

X_i	X_i^2
82	82^2 = 6724
77	77^2 = 5929
90	8100

71	5041
$EX_i = 320$	$EX_i^2 = 25794$

ANSWER: 8.042

- Find the sample standard deviation. Round to one decimal place (Utts & Heckard, 2012)
  44,66,31,43,86,32,76,53,60
  $EX_i = 491$ $EX_i^2 = 29707$ n-1 → 9-1 = 8
  ANSWER: 19.1

CALC CHEAT:
Stat→ ENTER→ enter data into L1
Stat→ Calc → 1:1-Var Stat → ENTER
Look at σx or sx (depending on what you are looking for)

- DO IT ANOTHER WAY Population deviation (Utts & Heckard, 2012)
  4, 9, 11, 12, 17, 5, 8, 12, 14
  1) Find Mean = 10.222
  2) Subtract the mean individually from each of the numbers given
     6.22, 1.22, 0.77, 1.778, 6.774, 5.225, 2.22, 1.77, 3.78
  3) Square each of the result
     38.7, 1.49, 0.60, 3.16, 45.9, 27.3, 4.94, 3.16, 14.3
  4) Add up all of the numbers that you squared
     139.55
  5) Divide by N
     N= 9
     15.51
  6) Square root it
     ANSWER: 3.94
- Sample deviation: 9, 2, 5, 4, 12, 7, 8, 11, 9, 3, 7, 4, 12, 5, 4, 10, 9, 6, 9, 4 (Utts & Heckard, 2012)
  ANSWER: 3.06

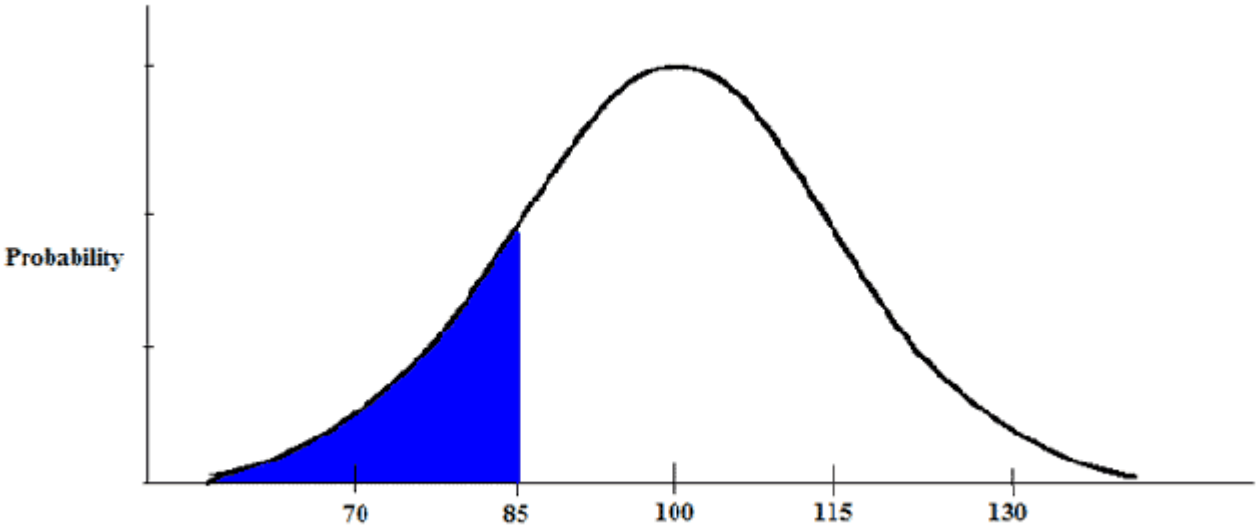
Z score: the distance that a data value is from the mean ((in terms of standard deviation)

$$z = \frac{\text{Observed value} - \text{Mean}}{\text{Standard deviation}}$$

Test	Mean	Standard Deviation	Score	Z-Score
Clerical Ability	50	15	41	-0.6
Logical Reasoning	40	4	47	1.75
Mechanical Ability	120	25	100	-0.8
Numerical Reasoning	100	10	105	0.5
Spatial Relations	70	20	90	1
Verbal Fluency	60	6	70	1.67

- The grades on a language midterm at Gardner Bullis are normally distributed with the population mean of the class being 83, sample mean 70, and standard deviation 3.5. Gabriela scored 82 on the exam. What is her z-score? (Utts & Heckard, 2012)
  ANSWER: -.28

- In the United States, the average IQ is 100, with a standard deviation of 15. What percentage of the population would you expect to have an IQ lower than 85? (Utts & Heckard, 2012)



ANSWER: .1587

1) First find z score: (85-100)/15 = -1.0
2) Look up z score on z table, and find out that the area is .8413
3) 1-0.8413 = 0.1587

References

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

SIT Session Lesson Plan

Week/Chapter: Week 3 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on today? My two challenging concepts this week are survey bias and sampling methods.

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
Warmup and little review over the last 2 weeks	Mini quiz with 5 questions covering the last 2 weeks	10 mins
Concept 1: Survey Bias	Opposite charade: Each student is going to be describing a survey scenario and the other students are going to guess which type of survey bias is being committed. They are then going to explain why it is that type of survey bias.	15 mins
Concept 2: Sampling Methods	Would give a visual example using a deck of cards and ask them to identify what method it was. They would later give me an example of when it would be used	15 mins
Conclusion	One Minute Paper	5 mins
Tutoring = Q & A		5-10 minutes

After session comments/thoughts:

Week 3: Mini Quiz Warm Up

1) Zip code. Qualitative OR Quantitative? **QUALITATIVE**

2) Collected from a subset of a larger population. What is the term? **SAMPLE** (Utts & Heckard, 2012)

3) Find the population standard deviation. (Utts & Heckard, 2012)
82,77,90,71
ANSWER: 6.964

4) Find the sample standard deviation. Round to one decimal place. (Utts & Heckard, 2012)
44,66,31,43,86,32,76,53,60
ANSWER: 19.1

5) The temperature is recorded at 60 airports in a region. The average temperature is 67 degrees Fahrenheit with standard deviation of 5 degrees. What is the *z*-score for a temperature of 68 degrees? (Utts & Heckard, 2012)
ANSWER: (68-67)/5 = 0.2

Correct data collecting method: Opposite charade method question of interest (Utts & Heckard, 2012)

A teaching software company wants to see if professors would use their software, but only emailed those college professors that registered in their website.	A call in radio show that ask for audience participation in a survey over gun control
Survey of high school students to measure teenage use of illegal drugs	Police officers are going door to door asking how well the citizens think they are being protected.
Calling different phone numbers in the phone book to see what interest service they liked better	A mail in survey that asks what is the best cereal brand
Asking people at a prolife convention if they support or disapprove of abortion	A law official conducts a survey by asking civilians if they are drug dealers.

Answers to Concept 1: (Utts & Heckard, 2012)

Selection bias	Selection bias
Selection bias bc does not include home schooled or dropouts	Biased response
Nonparticipation/nonresponse bias	Nonresponse bias
Selection bias	Biased response

References

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

SIT Session Lesson Plan

Week/Chapter: Week 4- Exam Review Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on this week? Focus on the rest of Chapter 8 and exam review for Exam

Beginning reminders:

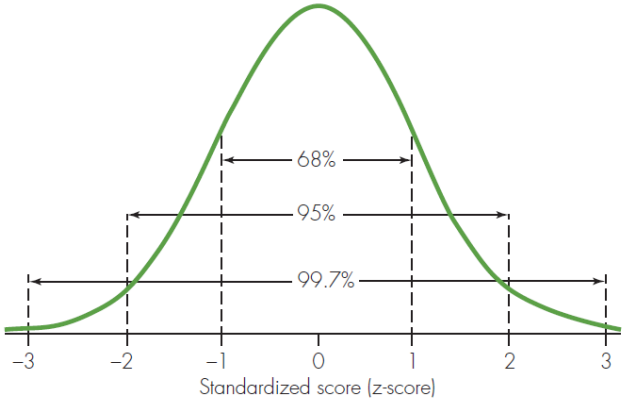
1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
Opening/Warm-Up/Ice-breaker	Get seats arranged, ask each other names and try to memorize! Also get signed in and seated	5-10 minutes
Challenging Content 1	-Sample vs. Population& statistics vs. parameter -Knowing Symbols (Summation, population, sample, means, etc.) -Frequencies, Relative frequency -Qualitative vs. Quantitative -Graphs: Histogram, Box plot, stem and leaf plot.	15-20 minutes
Challenging Content 2	-Bell shaped curves: Symmetric, Left and Right skewed. -5 number summary: Q1, Q2, Q3, Min, Max. Plus Range, percentiles, resistance! -Sampling methods, Bias, and variables (discrete and continuous)	15-20 minutes
	We will focus on these topics and go around making sure everyone is familiar with each topic and understands the concept.	
Closing/Check for Understanding	In order to check for understanding, we will have the students draw out a box plot, histogram, the empirical rule, equations and symbols needed for this test and by doing the review together.	5-10 minutes

Tutoring = Q & A	If students have questions!	5-10 minutes

After session comments/thoughts:

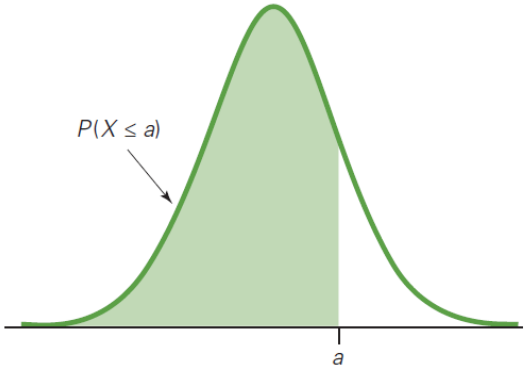
- **Empirical Rule & Standard deviation**:
  - Remember: **Empirical Rule** is exact characteristics of a normal curve model.

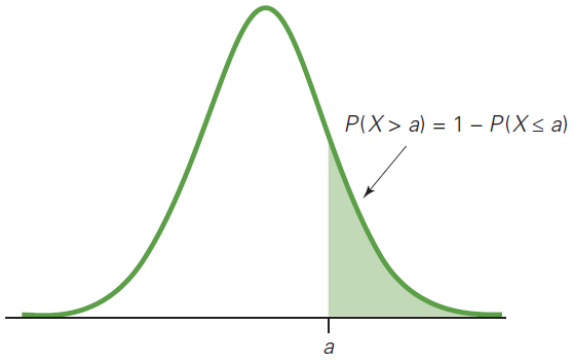


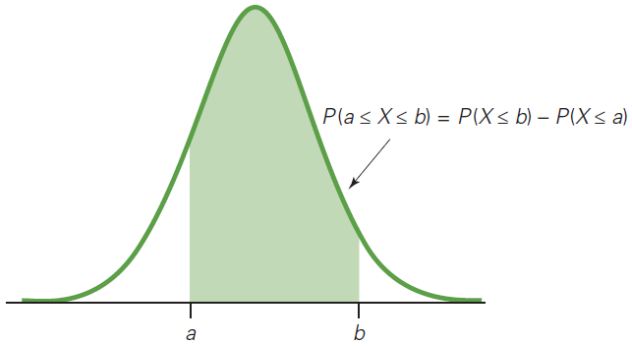
(Utts and Heckard, 2012, pg. 739)

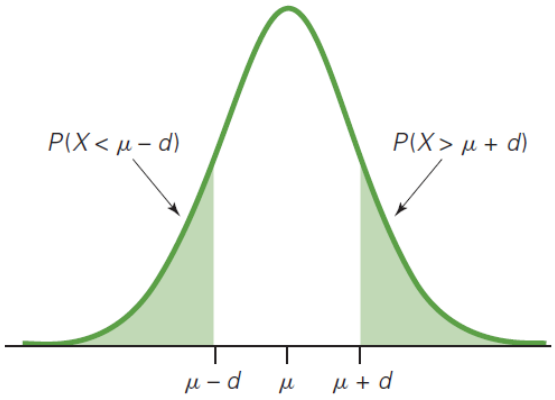
  - For bell-shaped data,
    - About **68%** of values have *z*-scores between –1 and +1.
    - About **95%** of values have *z*-scores between –2 and +2.
    - About **99.7%** of values have *z*-scores between –3 and +3.
  - **Standard deviation formula** and how to solve it from last week's notes.
    - Population & Sample:
      - .
      - . $\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{n}}$ $s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$

- **Z-scores and Probabilities and Z-score chart/table!**
  - .
  - $$z = \frac{\text{Value} - \text{Mean}}{\text{Standard deviation}} = \frac{x - \mu}{\sigma}$$

- 

- 

- 

- 

(Utts and Heckard, 2012, pg. 739)

References

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

SIT Session Lesson Plan

Week/Chapter: Week5 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on today? Parameter vs. Statistics and Review

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
WarmUp: Ice Breaker	2 lies and 1 truth game so that they would feel relaxed after taking the test.	5 mins
Concept1: Parameter Vs. Statistics	Around the World: This is where everyone would give an example of parameter and statistics. I would then give them some questions for them to work over, where one would do even and one would do odd questions.	15 mins
Concept2: Review	Use the study review, for them to have a refresher of the materials that were being taught 4 weeks ago.	20 mins
Tutoring = Q & A		5-10 minutes

Week5: Questions (Parameter vs. Statistics)

In each situation, explain whether the value given is a statistic or a parameter:
(Utts & Heckard, 2012)

1. A polling organization samples 1000 adults nationwide and finds that 72% of those sampled factor tougher penalties for persons convicted of drunk driving
2. N the year 2000 census, the US Census Bureau found that the median age of all American citizens was about 35 years old
3. For a sample of 20 men and 25 women, there is a 14 cm difference in the mean heights of the men and women.
4. A writer wants to know how many typing mistakes there are in his manuscript, she hires a proofreader, who reads the entire manuscript and finds 34 mistakes.
5. A polling organization plans to sample 1000 adult Americans to estimate the proportion of Americans who think crime is a serious problem in this country. What is the parameter of interest to the polling organization? What will be the statistic
6. Explain what symbol ($\hat{p}$, μ, etc.) would be used to represent each of the following:
   a. The mean of the sampling distribution of the sample mean.
   b. One value from the sampling distribution of the sample mean.
7. Explain what symbol ($\hat{p}$, μ, etc.) would be used to represent each of the following:
   a. The mean of the sampling distribution of $\hat{p}$
   b. One value from the sampling distribution of $\hat{p}$.

Review

1) Discrete or Continuous? The numbers of students in a class.
2) Discrete or Continuous? A dog's weight.
3) Discrete or Continuous? Time you wake up
4) Identify the type of sampling method. 100 students are randomly selected, and a survey is conducted to see if they have or is enrolled in a statistic class.
5) Find the population standard deviation. Round to one decimal place.
   44 66 31 43 (Utts & Heckard, 2012)
6) Parameter or Statistics? All of the students in TWU surveyed, said that they like statistics.
7) Parameter or statistics? The students surveyed in in Ms. Owen's class, majority said that they love statistics.
8) Quantitative or Qualitative? Jersey number
9) Quantitative or quantitative? House number
10) Quantitative or qualitative? Body temperature
11) The TEAS national average was 65, and standard deviation was 7. What is the probability Anabeth will get lower than a 75 on her TEAS test?

Week5: Answers

1. Statistic
2. Parameter
3. Statistic
4. Parameter
5. STAT: proportion of the 1000 adults in the sample who think crime is a serious problem
   POP: adult Americans
   (Utts & Heckard, 2012)
6. A. the mean of the sampling distribution of the sample means is the population mean μ. (Utts & Heckard, 2012)
   c. One value from the sampling distribution of the sample mean is one sample mean, denoted by $\bar{x}$. (Utts & Heckard, 2012)
7. A. the mean of the sample distribution of the sample proportion $\hat{p}$ is the population proportion *p*. (Utts & Heckard, 2012)
   c. One value from the sampling distribution of $\hat{p}$ is one sample proportion, denoted by $\hat{p}$. (Utts & Heckard, 2012)

Review

1) Discrete
2) Continuous
3) Continuous
4) Cluster sampling
5) 12.6

44	46	-2	4
66	46	20	400
31	46	15	225
43	46	-3	9
			638/4 = (159.5)^1/2 = 12.6

6) Parameter
7) Statistics
8) Qualitative
9) Qualitative
10) Quantitative
11) Z score = 75-65/7= 1.4285 ≈1.43
    Then look at z table at 1.43 to find the area/probability = **0.9236**

References

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning, pg 739.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

SIT Session Lesson Plan

Week/Chapter: Week 7 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on today? 1. Constructing Confidence Intervals for Proportions; and 2. Interpreting a Confidence Interval

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
Warm Up: If Game	Ask students questions about "if" scenarios (questions at the end of doc)	5 min.
Challenging Content: Constructing Confidence Intervals	Peer Lessons: Assign each person a part of the process. Give them time to figure their part out and as each person gets their part, they will go up to the board and show how they got their answer and explain their thought process.	40 min
Challenging Content: Interpreting Confidence Intervals		
Closer: Assess the Session	Have students write own one thing they liked about SIT and one thing they learned or understood better from going. Then have them write down one thing they would like to see done differently and any questions they feel didn't get answered or something they wanted to learn. Have them hand in their responses to me.	5 min
Tutoring = Q & A		5-10 minutes

After session comments/thoughts:

Practice Problem:

1. A researcher took a simple random sample of 675 elementary school kids, ages 7 – 12. He asked who liked broccoli. Of the 675 kids surveyed, 215 said they liked broccoli. Obtain an 87% confidence interval for the proportion of kids who said they like broccoli.
    a. Step 1: Compute $\hat{p}$
        i. $\hat{p} = \frac{x}{n} = \frac{215}{675} = .3185$
    b. Step 2: Compute confidence interval if $n\hat{p}\ and\ n(1-\hat{p}) \geq 10$
        i. $n\hat{p} = 675(.3185) = 214.988$
        ii. $n(1-\hat{p}) = 675(1-.3185) = 460.013$
    c. Step 3: Determine the critical value
        i. 87% = .87
        ii. 1- .87 = .13
        iii. .13/2 = .065
        iv. Find in chart – Z = -1.51 , 1.51
    d. Step 4: Determine lower and upper bounds of the confidence interval
        i. Find standard error of p-hat
            1. $se(\hat{p}) = \sqrt{\frac{.3185\ (1-.3185)}{675}} = .0179$
        ii. Confidence interval
            1. $\hat{p} \pm CV(se) = .3185 \pm 1.51(.0179) = (.2914, .3456)$
    e. Step 5: Interpret results
        i. We can be 87% confident that the proportion of kids between 7 and 12 that like broccoli is between .2914 and .3456

If Game Questions:

1. If you could talk to anyone in the world, who would it be?

2. If you could wish one thing to come true this year, what would it be?

3. If you could be someone else, who would you be?

4. If you could have any question answered, what would it be?

5. If you had one day to live over again, what day would you pick?

References

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

SIT Session Lesson Plan

Week/Chapter: Week 8 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on today? 1. Constructing Confidence Intervals for Means, using z and t; and 2. Interpreting a Confidence Interval

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
Warm Up: Get to Know your Neighbor	Have each person ask the person to their right one question. Be sure to give your name and the class section you are in	5 minutes
Challenging Concept: Constructing Confidence Intervals for Means, using Z and T	Structured Problem Solving: List out the steps for finding and interpreting means, using z and t. Have each student go to the board to work out problem. The person at the board is only the scribe and the people sitting down will tell them what to write. After everyone agrees on each part we will move through the other parts. We will work one of each kind together in the group.	45 minutes
Challenging Concept: Interpreting Confidence Intervals		
Tutoring = Q & A		5-10 minutes

After session comments/thoughts:

Session Problems:

1.) A researcher was interested in the average amount of time college students spent studying each week. She randomly polled 475 students and asked them to report the amount of time they spent studying. The results indicate that $\bar{x} = 6.5$ hours and s = 1.45 hours. Construct and interpret an 80% confidence level for the mean amount of time college students spend studying each week.

Step 1: Identify information given

- **n = 475**
- $\bar{x} = 6.5$
- **s = 1.45**
- **t* = 1.29**
- **df = 474**
- **CL = .80**

Step 2: Write out equation

- $\bar{x} \pm t^*(s/\sqrt{n})$

Step 3: Find t* and Find the standard error

- **t* = 1.29**
- $(s/\sqrt{n})$ **= .0665**

Step 4: Fill in equation and solve

- **$6.5 \pm 1.29(.0665)$ = (6.41, 6.58)**

Step 5: Interpret results

- **We can be 80% confident the amount of time college students spend studying each week is between 6.41 hours and 6.58 hours**

2.) A study, which surveyed 50 members of the U.S. Congress found that average amount of time spent watching television each week. $\bar{x}$ was found to be 3.5. The population is known to be normally distributed with a standard deviation of .45. Compute and interpret a 98% confidence interval for the average amount of time spent by U.S. Congress members watching television each week.

Step 1: Identify Information given in problem

- $\sigma = .45$
- $\bar{x} = 3.5$
- **n = 50**
- $Z^* = 2.33$

Step 2: Write down equation and Find Z*

- $\bar{x} \pm Z^*(\sigma)$
- Z^* **= 2.33**

Step 3: Fill in equation and solve

- $3.5 \pm 2.33(.45)$ **= 2.45**
- **3.5** $\pm$ **2.33(.45) = 4.55**

Step 4: interpret Results

- **We can be 98% confident that the average time U.S. Congress members spend watching television each week is between 2.45 and 4.55**

References

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

Planning the SIT Session

Week/Chapter: Week 9 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on this week? Test 2 Review

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
Opening/Warm-Up/Ice-breaker	Introduction: Names and favorite food	5 min
Content	Test Review: Jumbo Tic-Tac-Toe: Split into two teams, Xs and Os. Each tic tac toe space has a notecard with one of the questions on it. One scribe goes up to board to read question and team has to tell the scribe how to answer the question.	35-40 minutes
Closing/Check for Understanding	Q&A	5 min
Tutoring = Q & A	Question Time!	5-10 minutes

After session comments/thoughts:

9.1, 9.3, 9.4

1. In each situation, explain whether the value given in bold print is a statistic or a parameter: (2012, Utts & Heckard)
   a. A polling organization samples 1000 adults nationwide and finds that 72% of those sampled favor tougher penalties for persons convicted of drunk driving.
   b. In the year 2000 census, the U.S. Census Bureau found that the median age of all American citizens was about 35 years.
   c. For a sample of 20 men and 25 women, there is a 14-cm difference in the mean heights of the men and women.
   d. A writer wants to know how many typing mistakes there are in his manuscript, so he hires a proofreader, who reads the entire manuscript and finds 15 errors.
2. In a random sample of adults aged 18 to 25, individuals said that they drink alcohol at least once a month. (2012, Utts & Heckard)
   a. Calculate the value of $\hat{p}$ = sample proportion that drinks alcohol at least once a month.
   b. Calculate the standard error of $\hat{p}$.
3. In a random sample of n = 500 adults, 300 individuals say that they believe in love at first sight. (2012, Utts & Heckard)
   a. Calculate the value of $\hat{p}$ = sample proportion that believes in love at first sight.
   b. Calculate the standard error of $\hat{p}$.
4. Suppose that of all voters in the US, 40% are in favor of Candidate C for president. Pollsters take a sample of 2400 voters. What proportion of the sample would be expected to favor Candidate C? (2012, Utts & Heckard)
5. Past records show the statistics class average of a test taken by all 50 students has a mean of 82 and a standard deviation of 3.6.
   a. Describe the sampling distribution of the mean test score for a random selected sample of 50 students.
   b. Assume that the students in any stats class are similar to a random sample. P($\bar{x}$ < 83.5)

9.6, 9.9, 9.10

1. What is the mean of the sampling distribution for a sample mean?
2. What is the standard deviation of $\bar{x}$?
3. Given the following information, give the standard error of the mean. $\bar{x}$ = 325, s = 25, σ = 30, n = 1,532
4. Out of 3,000 participants in a study, 570 chose cup A in a taste test. Assume the percentage that chose cup A for the population is 20%. What is the probability that the proportion of those choosing cup A is greater than the 510?
5. What states that if n is sufficiently large, the sample means of random samples from a population with mean μ and finite standard deviation σ are approximately normally distributed with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$? (Utts & Heckard, 2012)
6. What states that sample mean $\bar{x}$ will eventually get "close" to the population mean μ, no matter how small a difference you use to define close. Larger the sample size, the more you can count on $\bar{x}$ to be an accurate representation of μ? (Utts & Heckard, 2012)

10.1,10.2, Pre 11

1) What is the confidence interval formula for proportion?
2) What is the formula for Margin of Error?
3) A poll conducted a survey, asking, "Do you think the use of marijuana should be legal or not?" In the nationwide poll of n= 2102 adults, 35% said that they favored legalization. The margin of error was given was 2%. Construct a 95% confidence interval for the proportion of those who favor legalization of marijuana. Interpret it.
4) In a survey of 190 college students, 134 students said that they believe there is extraterrestrial life.
   a. Find sample proportion who believe that there is extraterrestrial life (Utts & Heckard, 2012)
   B. Find an 86% confidence interval estimate of the proportion of all college students who believe that extraterrestrial life exists.
5) What is the confidence interval formula for population mean?
6) A store manager wanted to see how many TVs he sold for that month, so he gathered some data. He wants to find out the 93% confidence interval for the mean TVs sold for that month. His data were: 23, 13, 6, 3, 7, 3, 9, 10, 11, 29
7) Suppose that we conduct a survey of all the teachers at TWU (1000 teachers) to find out what percent of their income the average teacher donates to charity. We find out that the mean is 20%, where the standard deviation is 8%. Find the 97% confidence interval for the mean.

11.1,11.2, Sample Size Determination

1) A teacher is interested in the population mean score for a test that she gives every year. In the past the standard deviation was found to be 6.78 points. How large must her sample of test scores be to estimate the average score on the test within 2 points with a 90% confidence level?
2) The same teacher is interested in how long, on average, it takes students to finish the test. The known standard deviation is 8.9 minutes. How many students should be sample to estimate the average time it takes a student to finish the test within 2 minutes at a 95% confidence level.
3) A researcher is interested in the average number of televisions in American households. From previous studies, the standard deviation is known to be 2. How many households should he survey to find out the average number of television in an American household, within 1 TV with a confidence level of 95%?
4) A researcher is interested in the proportion of U.S. children who don't like watching television. How many people should be surveyed to estimate within 3% points of the true proportion with a confidence level of 90%?
5) What if we have a previously estimated proportion of 12.3%
6) A researcher was interested in the average distance a college student drives in amonth. She randomly surveyed 524 students and asked them to report the distance theydrive every month. The results indicate that the average distance driven by the surveyedstudents is 65 miles with a standard deviation of 5.3 miles. Construct and interpret an 80%confidence level for average distance college students drive in a month.

9.1, 9.3, 9.4 ANSWER KEY

1) A. Statistic because it is a sample value.
   b. Parameter because it is a population value.
   c. Statistic because it is a sample value
   d. Parameter because it is a population value (errors in the entire manuscript).
   (2012, Utts & Heckard)
2) A. . $\hat{p}$ = 590/1000 = .59
   B. s.e.($\hat{p}$) = sqrt/($\hat{p}$(1- $\hat{p}$)/n = .0156
   (2012, Utts & Heckard)
3) A. $\hat{p}$ = 300/500 = .60
   B. s.e.($\hat{p}$) = sqrt/($\hat{p}$(1- $\hat{p}$)/n = .022
   (2012, Utts & Heckard)
4) Mean = p = .4 (40%)
   Standard Deviation = s.d.($\hat{p}$) = sqrt/(p(1-p)/n = .01
   (2012, Utts & Heckard)
5) a. s.d.($\bar{x}$) = $\frac{\sigma}{\sqrt{n}}$ = 3.6 / sqrt 50 = .509117
   b. 83.5 – 88 /.509117 = 2.95
   z= 2.95
   = .9984
   =9.94

9.6, 9.9, 9.10 ANSWER KEY

1) μ
2) $sd(\bar{x}) = \frac{\sigma}{\sqrt{n}}$
3) $\frac{25}{\sqrt{1,532}}$ = 0.6387
4) $\hat{p}$ = x/n = 570/3,000 = .19 = 19%

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.19 - 0.20}{\sqrt{\frac{0.20(1-0.20)}{3{,}000}}} = -1.37$$

P($\hat{p} \geq .19$) = P($z \geq -1.37$) = 1- 0.0853 = 0.9147

5) Central Limit Theorem (Utts & Heckard, 2012)
6) Law of Large Numbers (Utts & Heckard, 2012)

10.1,10.2, Pre 11 ANSWER KEY

1) $\hat{p} \pm Z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ (Utts & Heckard, 2012)

2) $Z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ (Utts & Heckard, 2012)
Or anything after the +/-

3) $\hat{p} = .35$, margin of error = .02
.35 +/- .02 = **0.33** AND **0.37**
We are 95% confidence that between 33% to 37% of all American adults favored the legalization of marijuana.

4) **a.** $\hat{p} = 134/190 = .705.$ (Utts & Heckard, 2012)
B.
z multiplier: (1-0.86)/ 2 = 0.07 → -1.48→ 1.48
0.705 +/- 1.48 (.0330848)
0.705 +/- 0.48965
= **0.656** AND **0.754** → **65.6%** AND **75.4%**

5) $\bar{x} \pm Z * (\sigma/\sqrt{n})$

6) $\bar{x} = 6.2$
σ = 3.4
$Z^* = \frac{1-0.93}{2}$ = 0.035 → -1.81→ 1.81
N = 10

$$\bar{x} \pm Z * (\sigma/\sqrt{n})$$
$$6.2 \pm 1.81\ (3.4/\sqrt{10})$$
6 .2 ± 1.946065 =**4.254 & 8.146 → 425.4% & 814.6%**

7) $\bar{x} = .20$
σ = .08
$Z^* = \frac{1-0.97}{2}$ = 0.015 → -2.17
N = 1000

$$\bar{x} \pm Z * (\sigma/\sqrt{n})$$
$$.20 \pm -2.17\ (0.08/\sqrt{1000})$$
.20 ± 0.005489714 =**0.1945 & 0.2055 → 19.45% & 20.55%**

11.1,11.2,Sample Size Determination ANSWER KEY

1) $n = \left(\frac{z^*\sigma}{E}\right)^2 = \left(\frac{1.645 \times 6.78}{2}\right)^2 = 31$
2) $n = \left(\frac{z^*\sigma}{E}\right)^2 = \left(\frac{1.96 \times 8.9}{2}\right)^2 = 76$
3) $n = \left(\frac{z^*\sigma}{E}\right)^2 = \left(\frac{1.96 \times 2}{1}\right)^2 = 16$
4) $n = .25\left(\frac{z^*}{E}\right)^2 = .25 \times \left(\frac{1.96}{.03}\right)^2 = 1068$
5) $n = \hat{p}(1-\hat{p})\left(\frac{z^*}{E}\right)^2 = (.123 \times (1-.123)) \times \left(\frac{1.96}{.03}\right)^2 = 461$
6) Step 1: Identify information given
n = 524
$\bar{x}$ = 65

s = 5.3
t* = 1.29
df = 523
CL = .80

Step 2: Write out equation
$\overline{x} \pm t*(s \sqrt{n})$

Step 3: Find t* and Find the standard error
t* = 1.29
$(s \sqrt{n})$ = (5.3 √524) = .2315

Step 4: Fill in equation and solve
65 ± 1.29(.2315) = (64.7, 65.3)

Step 5: Interpret results
We can be 80% confident the average distance driven by college students in a month is between 64.7 miles and 65.3miles.

References

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

SIT Session Lesson Plan

Week/Chapter: Week 10 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on today? 1. Type I and Type II Errors; and 2. Conducting Hypothesis Test for Proportions – when to reject the null using critical values and alpha

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
Opener: Conversation Wheel	Give students a list of questions for topics of conversation. Each student gets a chance to ask another one question	5 Minutes
Challenging Concept: Type I and Type II Errors	Find the mistake. I will scramble the setup for the hypothesis statements. Each student will be given a print out of the problem used to make the statements. Everyone will work together to unscramble the hypotheses and make them all correct. This will help them teach each other how to set them up and understand why they are set up a certain way for certain problems	15 Minutes
Challenging Concepts: Conducting Hypothesis Test for Proportions	Summarizing the Steps: Have students work as a group through each step of the problem. Have each person act as a scribe to show work for each step on the board. Next person will write down steps it took to get to that answer. Rotate until problem is finished	35 Minutes
Tutoring = Q & A		5 Minutes

After session comments/thoughts:

Type I and Type II Errors:

- H_o: μ = 24 H_1: μ ≠ 24
  - Previous research shows that in 2003 the average number of television watched by American in a month was 24 hours. A researcher wants to know if the mean has changed since then.
- H_o: p = .25 H_1: p < .25
  - Previous research shows that in 2004 the proportions of Americans who liked reality television was .25. A researcher wants to know if that proportion has decreased since then.
- H_o: σ = 60 H_1: σ > 60
  - The standard deviation for the amount of candy bars sold from the Willy Wonka Factory in 2007 was 60. A researcher wants to know if the standard deviation has increased since then

Conducting Hypothesis Tests for Proportions: (rejecting the null)

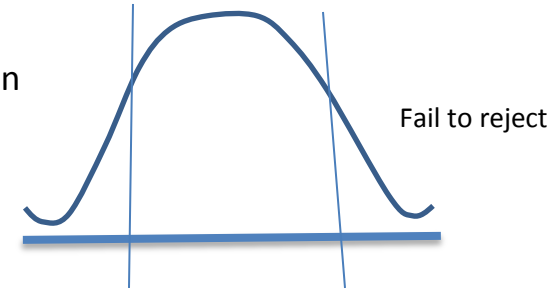
x = 40 n = 175 p_o = .1453 two-tailed $\alpha = .05$ $\hat{p}$ = .2286

1. Set up hypothesis statement
   a. H_o: p = .2453 H_1: p ≠ .2453
2. Find test statistic
   a. $z_0 = \frac{\hat{p} - p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}} = z_0 = \frac{.2286 - .2453}{\sqrt{\frac{.2453(1-.2453)}{175}}} = -.51$
3. Find p value
   a. .3050
4. Decide if p is less than alpha
   a. p < α .3050 < .05 (Not true)
5. Make decision about null
   a. Fail to reject the null

3. Find critical value and rejection region

   a. Critical value = ± 1.96

4. Conclusion



Test statistic not in rejection region

References

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

SIT Session Lesson Plan

Week/Chapter: Week 11 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on today? 1. Conducting hypothesis tests for t; 2. Power and when to use t and when to use z

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
Warm Up: The Sun Shines	Have one less chair than needed for the amount of students. The person without a chair must say the phrase "the sun shines on everyone who…" and then they have to finish the phrase with something that is true about them. If it applies to someone else that person must stand and the other person gets their seat and it continues until everyone has stood (Team Building Games).	10 Minutes
Challenging Concept: Conducting Hypothesis Test for t	Summarizing the Steps: Have students work through a problem together. As students go through each step they will discuss how they reached the answer for each step. We will attempt to do this without the notes. This will help student recall info similar to the way they would for a test.	40 Minutes
Closer: Power, When to use z and t	Note Review	5 Minutes
Tutoring = Q & A		5 Minutes

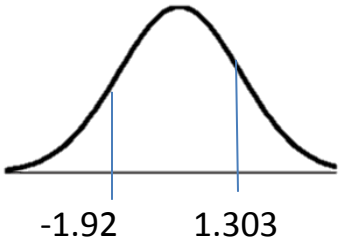
After session comments/thoughts:

Sample Question:

Info: μ = 17.5 n=45 $\bar{x}$ = 16.3 s=4.2 α= .1

- Step 1: H_o: μ = 17.5 H_1: μ < 17.5
- Step 2: $t = \frac{\bar{x} - \mu_o}{\frac{s}{\sqrt{n}}} = \frac{16.3 - 17.5}{\frac{4.2}{\sqrt{45}}} = -1.92$
- Step 3: find p-value -- .o274
- Step 4: p-value < α = .0274 < .1
  - Reject the null
- Step 5: There is statistically significant evidence

- Step 3: df = 44 – t = 1.303
- Step 4:



Reject Null

References

Team Building Games No.4 – The Sun Shines On...! (n.d.): n. pag. *Byc.org*. British Youth Council, 2010. Web. 31 Nov. 2014. <http://www.byc.org.uk/media/15231/Team%20Building%20Games%204_The%20Sun%20Shines%20On.pdf>.

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

SIT Session Lesson Plan

Week/Chapter: Week 12 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on today? 1. Conducting Hypothesis tests with z; and 2. Conducting Hypothesis tests with t

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
Warm Up: Meet Your Neighbor	Have each student introduce themselves and give one thing that they think no one would have in common with them that is not obvious to everyone else	5 minutes
Challenging Concept 1: Conducting Hypothesis Test for z	Going through the steps: Have students work as a team to list the steps for conducting hypothesis tests for z and then do the same for t. After steps are listed let student work in groups to solve sample problem. One group will solve using t and the other group will solve using t. Each group will report out what they find.	40 minutes
Challenging Concept 2: Conducting Hypothesis Test for t		
Tutoring = Q & A		10 minutes

After session comments/thoughts:

Sample Problem for z:

$\mu = 54.25 \quad \bar{x} = 60.43 \quad n = 32 \quad \sigma = 19.13 \quad \alpha = .05$

Step 1: $H_o: \mu = 54.25 \quad H_1: \mu \neq 54.25$

Step 2: $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{60.43 - 54.25}{\frac{19.3}{\sqrt{32}}}$ **= 1.81**

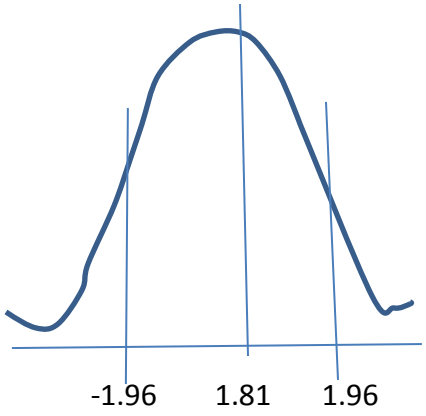
Step 3: .0351(2) = .0702 1-.9649 = .0351 (2) = .0702

Step 4: .0702 < .05, Fail to reject the null

Step 5: There is not sufficient evidence at alpha equals .05 to support the alternative

Step 3: $\alpha/2$ **= .025 C.V. = ± 1.96**

Step 4: There is not statistically significant evidence to support the alternative



Sample Problem for t:

$\mu = 54.25 \quad \bar{x} = 60.43 \quad n = 32 \quad s = 19.13 \quad \alpha = .1$ **df = 31**

Step 1: $H_o: \mu = 54.25 \quad H_1: \mu > 54.25$

Step 2: $t = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{60.43 - 54.25}{\frac{19.3}{\sqrt{32}}}$ **= 1.81**

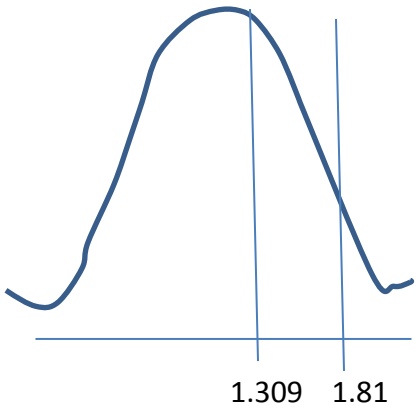
Step 3: p- value = .025 - .05

Step 4: .025-.05 < .1, Reject the null

Step 5: There is sufficient evidence at alpha equals .1 to support the alternative

Step 3: $\alpha = .1 \quad df = 31 \quad C.V. = 1.309$

Step 4: There is statistically significant evidence to support the alternative



References

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

SIT Session Lesson Plan

Week/Chapter: Week 13 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on today? Exam 3 Review

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
Opener:	Have everyone say there name and most challenging concept for test three	10 Minutes
Review	After students give their most challenging aspect of test three (I will write them down as they list them) I will ask each person if they fill they know any of the listed concept very well. I will write their names down by each that they are confident in. Each student will have a chance to explain a concept to the others. I will assist as needed to further explain concepts. If time is left I will have students discuss concepts and things they remember about all chapters covered this semester.	35 Minutes
Tutoring = Q & A		10 Minutes

After session comments/thoughts:

Review: Exam 3

1. What is the difference between null and alternative hypothesis?

2. Symbol for level of significance

3. For each of the following situations, write the alternative hypothesis (Utts & Heckard, 2012)
   a. The null hypothesis is Ho: p = .30 , and it is a two sided hypothesis test

   b. The null hypothesis is Ho: p = .45, and it is a one right sided test

   c. The null hypothesis is Ho: p = .45, and it is a one left sided test

4. Write the null and alternative hypothesis:
   a. An anthropologist knows that in a region she has studied, 15% of the population has a certain unique genetic trait. She plans to take a random sample of people in the new region and test them for that trait. If the percentage of the population with the trait in the new region differs from that in the known region, it would indicate that the two populations descended from different ancestors. (Utts & Heckard, 2012)

   b. A candy company's website claims that 30% of these candies are red, which is Paul's favorite. He doubts their claim and thinks that fewer than 30% are red. He is willing to assume that the candy is randomly placed into bags, and he plans to buy several bags of candy and count the proportion of red pieces. (Utts & Heckard, 2012)

5. In a recent survey, 61 out of 100 consumers reported that they preferred plastic bags instead of paper bags for their groceries. If there is no difference in the proportions who prefer each type in the population, the chance of such extreme results in a sample of this size is about .03. Because .03 is less than .05, we can conclude that there is a statistically significant difference in preference. (Utts & Heckard, 2012)
   a. P value?

   b. Level of significance?

   c. Sample proportion?

   d. Sample size?

   e. Null value ?

6. For each of the following outcomes, explain whether a Type 1 error, a Type 2 error, or a correct decision has been made. Draw the hypothesis testing error chart.

a. Null hypothesis is true, and the failed to reject null hypotheses

b. Alternative hypothesis is true, and failed to reject null hypotheses

c. Null hypothesis is true, and null hypothesis is rejected

d. Alternative hypothesis is true, and null hypothesis is rejected.

7. Medical researchers now believe there may be a link between baldness and heart attacks in men.

a. State the null hypothesis and the alternative hypothesis for a study used to investigate whether or not there is such a relationship. (Utts & Heckard, 2012)

b. Discuss what would constitute a Type 1 error in this study. (Utts & Heckard, 2012)

c. Discuss what would constitute a Type 2 error in this study. (Utts & Heckard, 2012)

8. A politician is trying to decide whether to vote for a new tax bill that calls for substantial reforms. A random sample of voters in his district leads him to believe the alternative hypothesis, Ha: p> .5, where p is the proportion of all voters in his district who support the bill as a consequence, he decides to vote for the bill. (Utts & Heckard, 2012)

a. What would a Type 1 error be in this situation, and what would be the consequences for the politician?

b. What would a Type 2 error be in this situation, and what would be the consequences for the politician?

9. Which hypotheses are two tailed?

10. Which hypotheses are one tailed?

11. What is the null standard error formula?

12. What is the standard error formula for sample proportion?

13. What is the z score formula for a proportion?

14. Conditions for conducting the Z test?

15. Conditions for conducting the T test?

16. Find the p value. Z statistic = -1.09; H_0: p = .5; H_a: p < .5. Specify the rejection region for alpha = .05, and then reach a conclusion for the test using the rejection region rule. (Utts & Heckard, 2012)

17. Find the p value. Z statistic = 2.10; H_0: p = .10; H_a: p ≠.10. Specify the rejection region for alpha = .05, and then reach a conclusion for the test using the rejection region rule (Utts & Heckard, 2012)

18. In each of the following, determine whether the conditions for conducting a z test for proportion are met. If not, explain why not.

a. Twenty students are randomly selected from the list of all sorority and fraternity members at a university to determine if a majority of sorority and fraternity students favor a new policy on alcohol on campus. The hypotheses are as follows: H_0: p = .50; H_a: p >.50 (Utts & Heckard, 2012)

b. Twenty employees of a large company are randomly selected to determine whether the proportion of company employees who are left handed exceeds the national proportion of 10% who are left handed. The hypotheses are as follows: H_0: p = .10; H_a: p >.10 (Utts & Heckard, 2012)

19. In a survey of 240 students in an elementary statistics class at the University of California – Davis, 20 said that they were left handed and 200 said that they were right handed. Assume that the students are representative of all students at the school. Does this provide evidence that the proportion of UCD students who are left handed differs from the national proportion of 0.10? Carry out the 5 steps of the hypothesis test for this situation. (Utts & Heckard, 2012)

20. In each case, suppose the sample proportion is $\hat{p} = .75$ and compute the test statistic and the p value. Then make a conclusion using a .05 level of significance. Comment on the relationship between the sample size and the conclusion.
H_0: $p = .70$; H_a: $p \neq .70$ (Utts & Heckard, 2012)

a. n= 100

b. n= 500

21. Power increase when sample size _____

22. The power increase when the difference between the _____ and the _____ increases

23. The managers at Hunter Chemical claim that their major product contains on average 4.0 fluid ounces of caustic material per gallon. They further state that the distribution of caustic material is normal with a standard deviation of 1.3 fluid ounces. If there is too much caustic material, the product will be dangerous. A government inspector is brought in to test the product. She randomly selects a sample of 100, gallon size containers of the product and finds that their mean weight of caustic materials are 4.5 fluid ounces per gallon. Carry out the hypotheses test for this situation. (Utts & Heckard, 2012)

24. It has been hypothesized that the mean pulse rate for college students is about 72 beats per minute. A sample of Penn state students recorded their pulse rate. Assume that the sample is representative of all Penn State women for pulse rate measurements, where they surveyed 35 students, and the mean was 76.9, and the standard deviation was 11.6. What is the hypothesis test conclusion for this scenario? (Utts & Heckard, 2012)

25. Perry wanted to test the folklore that women have not been given information about the sex of their unborn child can guess it at better than chance levels. They asked a sample of 104 pregnant women to guess the sex of their babies, and 57 guessed correctly. Assuming chance guessing would result in 50% correct guesses, test the hypothesis that women can guess at a better than chance level. Carry out the test using $\alpha = 0.05$, and make sure to state a conclusion. (Utts & Heckard, 2012)

26. A university is concerned that it is taking students too long to complete their requirements and graduate; the average time for all students is 4.7 years. The dean of the campus honors program claims that students who participated in that program in their first year have had a lower mean time to graduation. A random sample of 30 students who had participated is taken, and the mean and standard deviation for the time to completion for those students are 4.5 years and 0.5 year, respectively. Carry out the 5 steps to test the hypotheses of interest in this situation. Use $\alpha = 0.05$. (Utts & Heckard, 2012)

27. Gasoline prices have remained higher for state along the West Coast, particularly in Alaska and California. Let's say we drew a random sample of 100 California gas stations, and the mean price was $3.11. We also know that nationally the mean price of a gallon was $2.86, with a standard deviation of 0.17. We expect to find that the average price of gas from California gas stations will be higher than before. Carry out the steps to test the hypotheses of interest. Use $\alpha = .05$. (Utts & Heckard, 2012)

28. Compute the effect size for each of the following situations, and state whether it would be considered closer to a small, medium, or large effect: (Utts & Heckard, 2012)

a. In a paired difference test with $n = 100$ pairs, the test statistics is $t = 2.24$

b. In a paired difference test with $n = 50$ pairs, the test statistics is $t = -2.83$

29. What is the formula to compute the effect size for one proportion?

30. Identify whether a scatterplot would or would not be an appropriate visual summary of the relationship between the following variables. In each case, explain your reasoning. (Utts & Heckard, 2012)

a. Blood pressure and age

b. Region of country and opinion about stronger fun control laws.

31. The website gives age in years and body temp in degrees Fahrenheit for 100 blood donors ranging in age 17 to 84 years old. The regression equation is $\hat{y} = 98.6 - 0.0138x$. (Utts & Heckard, 2012)

a. In the regression relationship shown, which variable is the response and which is the explanatory variable?

b. What is the predicted body temp for someone who is 50 years old?

c. One of the donors was 50 years old and had a body temp of 97.6. What is the residual for this person? Explain what the residual tells you about this person's body temp in comparison to the average body temp for someone his age.

32. A regression equation for y = handspun (cm) and x = height (inches) . if the roles of the variables are revered and only women are considered, the regression equation is

Average height = 51.1 + 0.7 (handspun) (Utts & Heckard, 2012)

a. Interpret the slope of 0.7 in terms of how height changes as handspan increase.

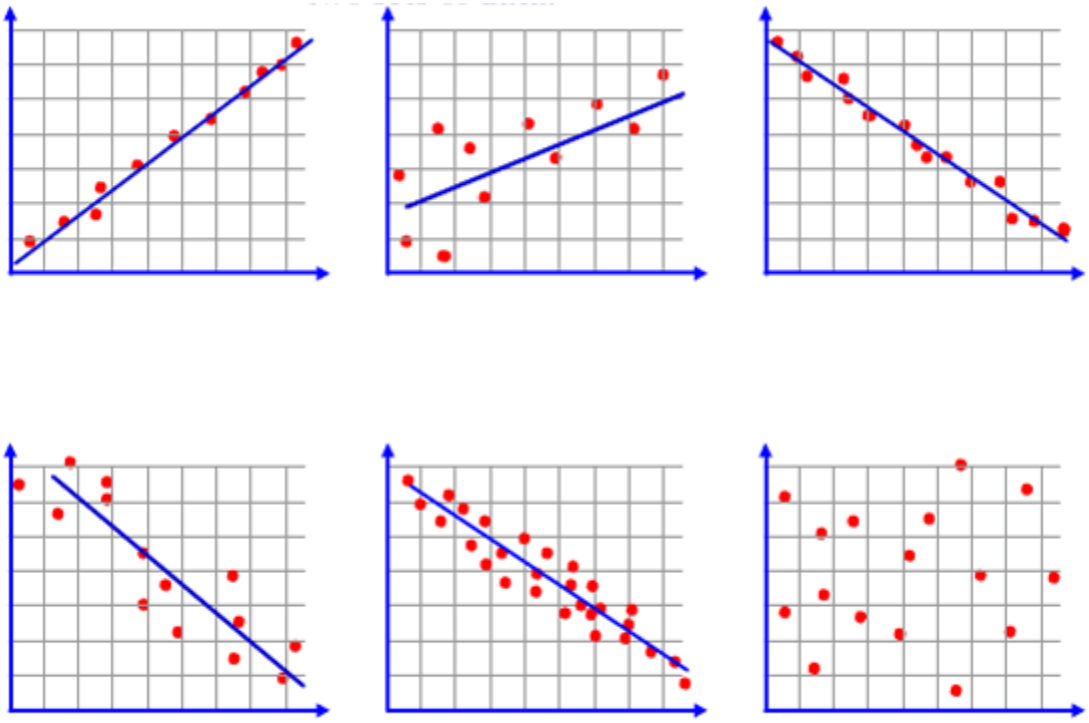
b. What is the estimated average height of women with a handspan of 20 cm?

c. Molly has a handspan of 20 cm and is 66.5 inches tall. What is the prediction error for Molly?

33. Which implies a stronger linear relationship: a correlation of + 0.4 of a correlation of -0.6? Briefly explain.

34. The correlation between latitude and average August temp is -0.78 for the 20 cities. Calculate coefficient of determination and write a sentence that interprets it in the context of this situation. (Utts & Heckard, 2012)

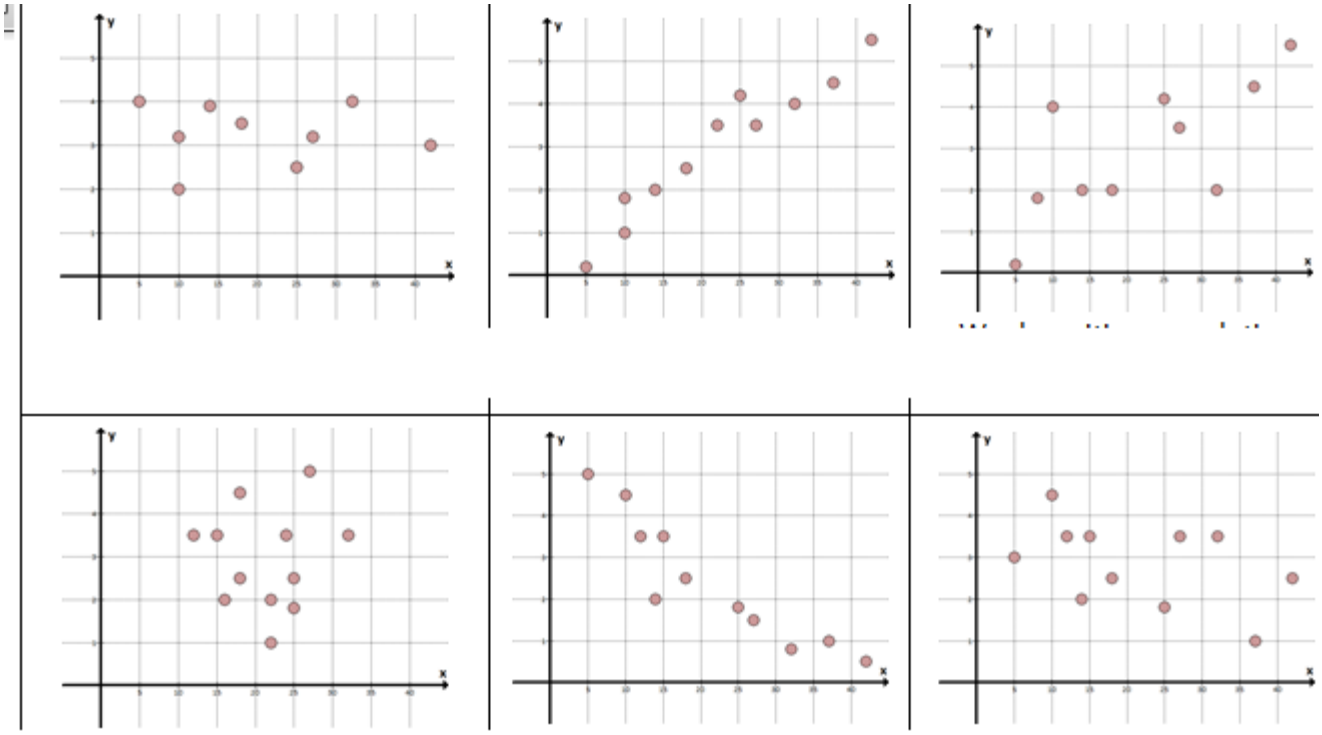
35. State whether each graph has a strong, weak or positive, negative correlation.



(Utts & Heckard, 2012)

36. Match the coefficient correlation with the appropriate graph (Utts & Heckard, 2012)

a. -0.05 b. 0.71 c. -0.45 d. 0.96 e. 0.04 f. -0.92



37. What is the correlation coefficient when $r^2 = 0.56$?

38. What is the formula for r^2?

Exam3 Review Answers

1) Null: statement to be tested; hoping to reject or disapprove
Alternative: statement that we are trying to find evidence to support

2) α

3) **a.** $H_a: p \neq .30.$
b. $H_a: p > .45.$
c. $H_a: p < .60.$
(Utts & Heckard, 2012)

4) a. Ho: p = .15
Ha: p ≠ .15

c. Ho: p = .30
Ha: p < .30 (Utts & Heckard, 2012)

---

5) **a.** The p-value is .03.
b. The level of significance is $\alpha = 0.05$.
c. The sample proportion is $\hat{p} = 61/100 = .61$.
d. The sample size is $n = 100$.
e. The null value is .5. If there were no difference in preference, then 50% of the population of consumers would prefer each type.
(Utts & Heckard, 2012)

---

6)

		DECISION	
		Reject H_o	*Fail to Reject H_o*
ACTUAL	*H_o True*	Type I Error ***Producer Risk*** **α-Risk** **False Positive**	**Correct Decision** **Confidence Interval = 1- α**
	H_a True	**Correct Decision** **Power = 1- β**	Type II Error ***Consumer Risk*** **β-Risk** **False Negative**

H_o: Null Hypothesis H_a: Alternative Hypothesis (Utts & Heckard, 2012)

---

7) **12.31** **a.** The correct decision has been made. The null is true and it is not rejected.
b. A type 2 error has occurred. The alternative is true, but the null is not rejected.
c. A type 1 error has occurred. The null is true, but it has been rejected.
d. The correct decision has been made. The alternative is true and the decision is in favor of the alternative.

12.34 **a.** *Null*: No relationship between baldness and heart attacks.
Alternative: There is a relationship between baldness and heart attacks.
b. A type 1 error would occur if it is decided there is a relationship between baldness and heart attacks when in reality there is no relationship.
c. A type 2 error would occur if there is not enough evidence to conclude that there is a relationship between baldness and heart attacks when in reality there is a relationship.

(Utts & Heckard, 2012)

a. A type 1 error occurs if the politician believes that more than one-half of voters in his district support the new tax bill when the proportion really is not more than one-half. The consequence is that he would vote for a bill that is not supported by a majority of the voters in his district.

b. A type 2 error occurs if the politician believes the proportion of voters in the district supporting the tax bill is not a majority when really it is a majority. The consequence is that that he would not vote for a bill that is supported by a majority of the voters.

8) (Utts & Heckard, 2012)

9) $P \neq P_0$

10) $P > P_0$ AND $P < P_0$

11) $= \sqrt{\frac{p_0(1-p_0)}{n}}$, (Utts & Heckard, 2012)

12) $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ (Utts & Heckard, 2012)

13) $$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$ (Utts & Heckard, 2012)

1. The sample is a simple random sample.
2. np_0 and $np_0(1-p_0)$ should be at least 10.

14) (Utts & Heckard, 2012)

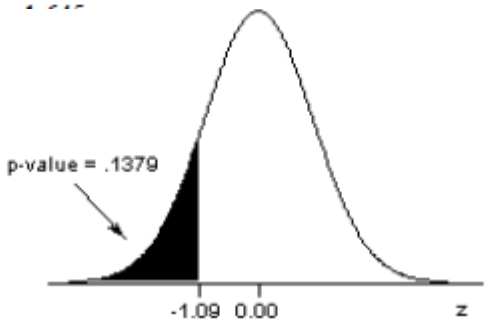
15) The sample is SRS, and N $\geq$ 30(Utts & Heckard, 2012)

c. p-value = .1379. This is the probability (area) to the left of $z = -1.09$. Table A.1 can be used to find $P(Z < -1.09) = .1379$.

c. Rejection region is $z < -1.645$.

Cannot reject the null hypothesis. z-statistic = –1.09 is not less than –1.645.

16)



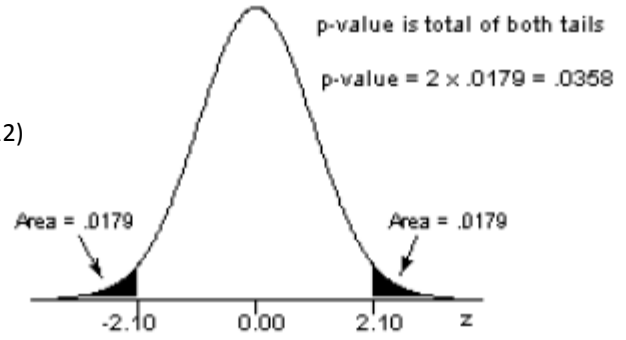
(Utts & Heckard, 2012)

12.41 **a.** p-value = .0358. This is the combined probability to the right of $z = 2.10$ and left of $z = -2.10$. Table A.1 can be used to find that $P(Z < -2.10) = .0179$. By symmetry, the area to the right of $z = 2.10$ is also .0179. So, the p-value = $2 \times .0179 = .0358$. (Utts & Heckard, 2012)

a. Rejection region is $|z| > 1.96$.

Reject the null hypothesis. z-statistic = 2.10 is greater than 1.96.

(Utts & Heckard, 2012)



17)

12.43 **a.** Yes. The sample is a random sample, and the sample size is large enough because $np_0 = (20)(.50) = 10$, and $n(1-p_0) = (20)(1-.50) = 10$. (Utts & Heckard, 2012)

b. No. The sample size is not large enough because $np_0 = (20)(.10) = 2$ is smaller than 10.

18)

12.91 Step 1: H_0: $p = .1$ (proportion left-handed same as in national population)

H_a: $p \neq .1$ (proportion left-handed not same as in national population.

Step 2: The necessary conditions for using the z-statistic are present. The sample is assumed to be representative of the larger population and the sample size is large enough so that both np_0 and $n(1-p_0)$ are greater than 10. Here, $n = 240$ and $p_0 = .1$.

The test statistic is $z = \dfrac{\text{Sample estimate - Null value}}{\text{Null standard error}} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$.

Sample estimate = $\hat{p} = 20/240 = .0833$

$$z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{.0833 - .1}{\sqrt{\dfrac{.1(1-.1)}{240}}} = \frac{-.0167}{.01936} = -0.86$$

Step 3: p-value ≈ .39. This is the combined probability (area) to the left of $z = -0.86$ and to the right of $z = 0.86$. Table A.1 gives $P(z < -0.86) = .1949$. Because this is a two-sided test, p-value = $2 \times .1949 = .3898$.

Note: An exact p-value based on the binomial distribution is given by Minitab as .395.

Step 4: Cannot reject the null hypothesis. The result is not statistically significant because the p-value is not smaller than .05, the usual standard for significance.

Step 5: We cannot conclude that the proportion of UCD students who are left-handed differs from the national proportion.

19)

(Utts & Heckard, 2012)

b. $\sqrt{\dfrac{.70(1-.70)}{100}} = .0458$. **b.** $\dfrac{.75-.70}{.0458} = 1.091$; $p \approx 2 \times P(Z < -1.09) = (2)(.1379) = .2758$, so fail to reject null hypothesis.

c. $\sqrt{\dfrac{.70(1-.70)}{500}} = .0205$. **c.** $\dfrac{.75-.70}{.0205} = 2.439$; $p \approx 2 \times P(Z < -2.44) = (2)(.0073) = .0146$, so reject the null hypothesis.

(Utts & Heckard, 2012)

20)

21) Power increases when the sample size increases

22) The power increase when the difference between the true population value and the null hypothesis value increases

23) $P(\bar{X} \geq 4.5 \mid \mu = 4) = P(\frac{X-\mu}{\sigma/\sqrt{n}} \geq \frac{4.5-4}{1.3/\sqrt{100}}) = P(Z \geq 3.85) = .0001$ (Utts & Heckard, 2012)

H0: μ = 0.4 Ha: μ > 0.4

REJECT THE NULL HYPOTHESES

---

b. Step 1: H_0:μ = 72 versus H_a:μ ≠ 72.

Step 2: Sample size (35) is sufficiently large to proceed. We assume the sample represents a random sample of Penn State women.

μ = mean pulse rate for population of Penn State women

Test statistic is $t = \frac{\text{Sample statistic - Null value}}{\text{Null standard error}} = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{76.9-72}{\frac{11.6}{\sqrt{35}}} = \frac{4.9}{1.961} = 2.50$

Steps 3,4, and 5: df = $n - 1 = 35 - 1 = 34$. The p-value is $2 \times P(t > 2.50)$. Using Table A.3, the p-value is between 2(.008) = .016 and 2(.013) = .026. We can reject the null hypothesis and conclude that mean female pulse rate is significantly different from 72. With software like Excel or Minitab, or a suitable calculator, it can be found that the p-value is 2(.0087) = .0174.

c. The data provide evidence that the mean pulse rate of Penn State women is different from 72 beats per minute, but do not give evidence that the mean pulse rate for Penn State men is significantly different from 72. (Utts & Heckard, 2012)

24) **12.111** Step 1: H_0: $p \leq .50$ (not better than chance level)

H_a: $p > .50$ (predict better than chance level)

p = proportion of population of pregnant women that can predict the sex of their babies

Step 2: We must assume the sample represents a random sample from the population of pregnant women. The sample size is sufficiently large so that $n\hat{p}$ and $n(1-\hat{p})$ are both greater than 10.

Sample proportion correct guesses is $\hat{p} = \frac{57}{104} = .548$

Test statistic is $z = \frac{\text{Sample statistic - Null value}}{\text{Null standard error}} = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{.548-.50}{\sqrt{\frac{.50(1-.50)}{104}}} = \frac{.048}{.049} = 0.98.$

Step 3: p-value = .1635. It is the area (probability) to the right of $z = 0.98$.

$P(z > 0.98) = P(z \leq -0.98) = .1635$. Equivalently, $P(z > 0.98) = 1 - P(z \leq 0.98) = 1 - .8365 = .1635$.

Steps 4 and 5: We do not reject the null hypothesis for $\alpha = .05$. There is not enough evidence to conclude that in the population of pregnant women represented by the sample, the proportion able to predict the sex of their babies is higher than .50 (the chance level).

Note: Minitab could be used to do this exercise (see Minitab Tip on page 485). The program reports the exact I-value based on binomial distribution probabilities as .189. (Utts & Heckard, 2012)

25) **13.29** Step 1: H_0:μ = 4.7, H_a:μ < 4.7, where μ = mean time to graduate (in years) for students who participated in the honors program in their first year of college.

Step 2: Sample size (30) is sufficiently large to proceed. It is stated in the problem that we can assume the sample represents a random sample of the population of students who participated in the honors program in their first year of college.

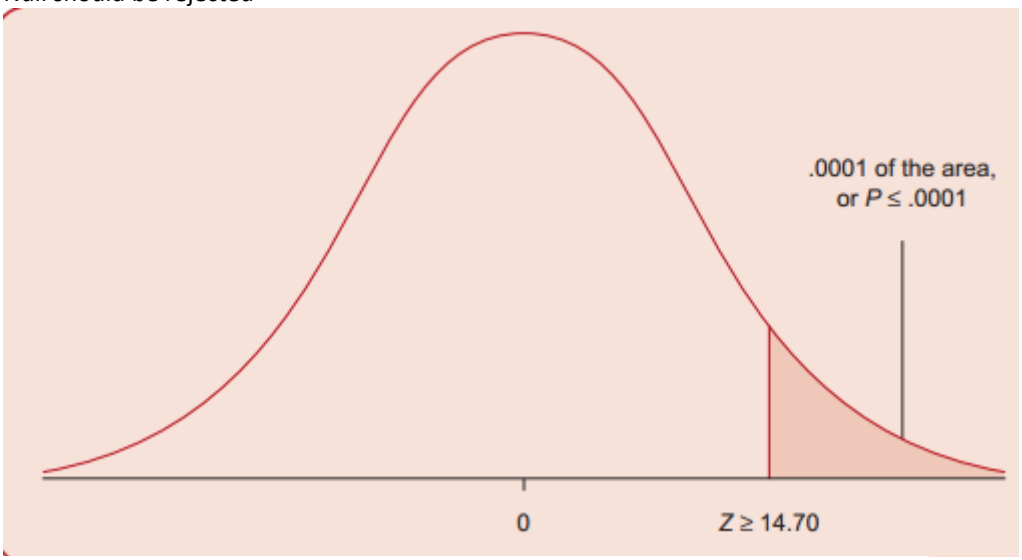
Test statistic is $t = \frac{\text{Sample statistic - Null value}}{\text{Null standard error}} = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{4.5-4.7}{\frac{0.5}{\sqrt{30}}} = \frac{-0.2}{0.0913} = -2.19$

Steps 3,4, and 5: df = $n - 1 = 30 - 1 = 29$. The p-value is $P(t < -2.19)$. Using Table A.3, the p-value is between .013 and .027. With software like Excel or Minitab, or a suitable calculator, it can be found that the p-value is .018. Using the standard $\alpha = .05$, we can reject the null hypothesis and conclude that the mean time to graduate is lower for the population of students who participated in the honors program than it is for the general population of students. (Utts & Heckard, 2012)

26)

$Z = \frac{3.11 - 2.86}{0.17/\sqrt{100}} = 14.70$ (Utts & Heckard, 2012)

27) Null should be rejected $H_0: \mu_Y = \$2.86$ $H_1: \mu_Y > \$2.86$



13.77 **a.** Effect size is $\frac{2.24}{\sqrt{100}} = 0.224$; small.

b. Effect size is $\frac{-2.83}{\sqrt{50}} = -0.4$; closer to medium. (Utts & Heckard, 2012)

28) d(hat) = t/ n(1/2)

$$ES = \frac{|p_1 - p_0|}{\sqrt{p_0(1-p_0)}}$$ (Utts & Heckard, 2012)

29)

3.5 **a.** Appropriate because both variables are quantitative.

b. Not appropriate because both variables are categorical. (Utts & Heckard, 2012)

30)

3.75 **a.** Body temperature is the response variable and age is the explanatory variable.

b. $\hat{y} = 98.6 - 0.0138x = 98.6 - 0.0138(50) = 97.91$ degrees. (Utts & Heckard, 2012)

c. The predicted temperature for a person whose age = 50 was found in part (b) to be 97.91. Residual = actual – predicted = 97.6 – 97.91 = –0.31. His body temperature is 0.31 degrees lower than would be predicted based on his age.

31)

3.19 **a.** For each one-centimeter increase in handspan, average height increases 0.7 inches.

b. $\hat{y} = 51.1 + 0.7(20) = 65.1$ in.

c. $66.5 - 65.1 = 1.4$ in. (Utts & Heckard, 2012)

32)

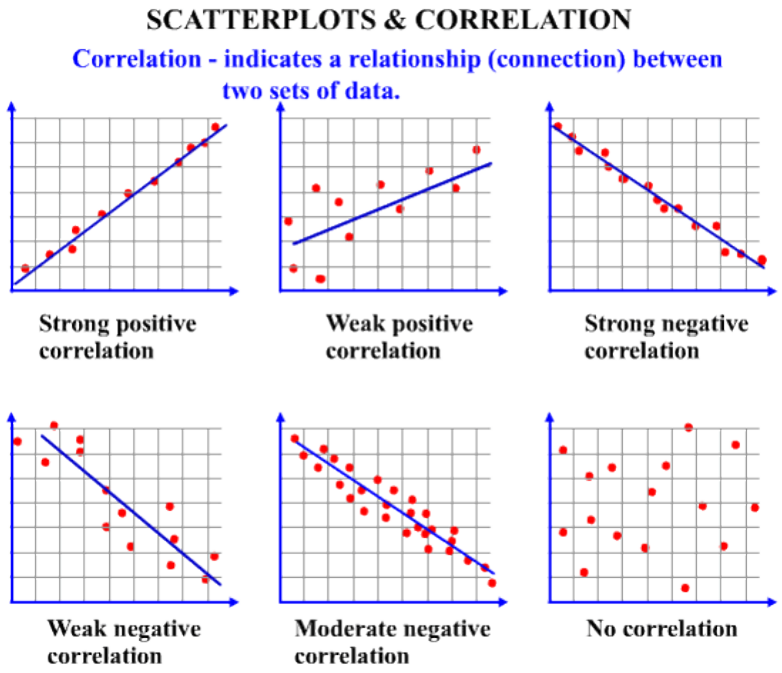
3.32 A correlation of –0.6 implies a stronger linear relationship than a correlation of +0.4 because in absolute magnitude it is closer to 1. To judge the strength of a linear relationship, ignore the sign of a correlation coefficient. (Utts & Heckard, 2012)

33)

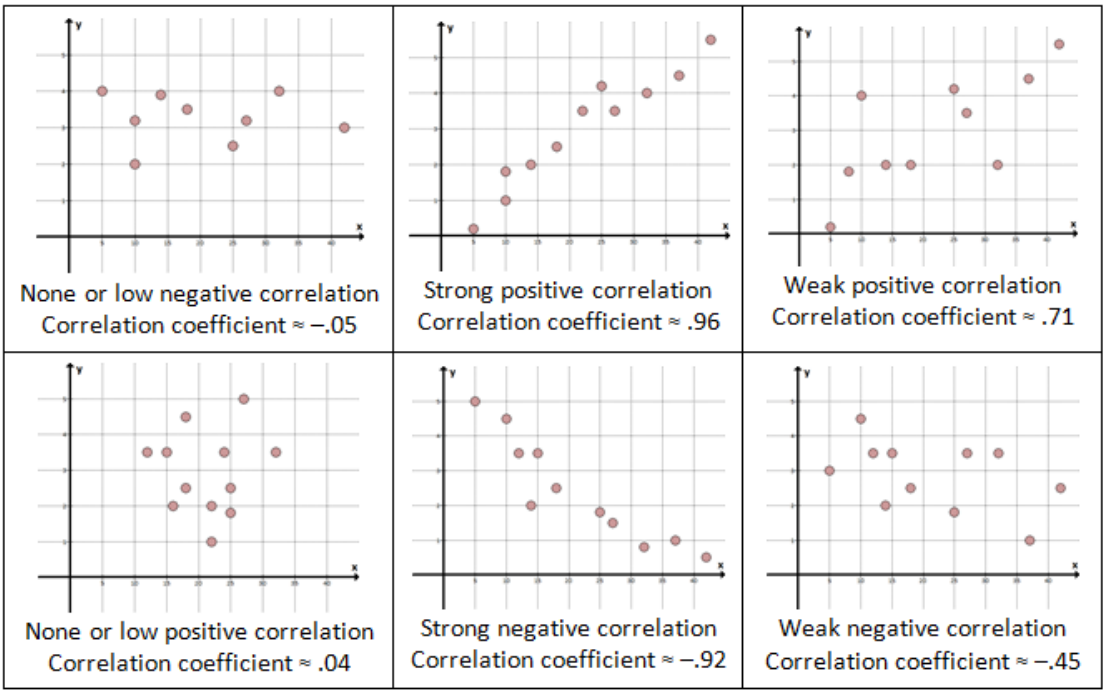
3.44 **a.** $r^2 = (-0.78)^2 = .6084$. This means that latitude explains about 60.8% of the variation in average August temperature. (Utts & Heckard, 2012)

34)

35)



(Utts & Heckard, 2012)



(Utts & Heckard, 2012)

36)

37) R = .748

38) R^2= explained variation/ total variation = 1- (unexplained variation/ total variation)

Utts, J., & Heckard, R. (2012). *Mind on statistics* (Fifth ed., p. 739). Stamford: Cengage Learning.

References:

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

Review: Exam 3

1. What is the difference between null and alternative hypothesis?

2. Symbol for level of significance

3. For each of the following situations, write the alternative hypothesis (Utts & Heckard, 2012)
   a. The null hypothesis is Ho: p = .30 , and it is a two sided hypothesis test

   b. The null hypothesis is Ho: p = .45, and it is a one right sided test

   c. The null hypothesis is Ho: p = .45, and it is a one left sided test

4. Write the null and alternative hypothesis:
   a. An anthropologist knows that in a region she has studied, 15% of the population has a certain unique genetic trait. She plans to take a random sample of people in the new region and test them for that trait. If the percentage of the population with the trait in the new region differs from that in the known region, it would indicate that the two populations descended from different ancestors. (Utts & Heckard, 2012)

   b. A candy company's website claims that 30% of these candies are red, which is Paul's favorite. He doubts their claim and thinks that fewer than 30% are red. He is willing to assume that the candy is randomly placed into bags, and he plans to buy several bags of candy and count the proportion of red pieces. (Utts & Heckard, 2012)

5. In a recent survey, 61 out of 100 consumers reported that they preferred plastic bags instead of paper bags for their groceries. If there is no difference in the proportions who prefer each type in the population, the chance of such extreme results in a sample of this size is about .03. Because .03 is less than .05, we can conclude that there is a statistically significant difference in preference. (Utts & Heckard, 2012)
   a. P value?

   b. Level of significance?

   c. Sample proportion?

   d. Sample size?

   e. Null value ?

6. For each of the following outcomes, explain whether a Type 1 error, a Type 2 error, or a correct decision has been made. Draw the hypothesis testing error chart.

a. Null hypothesis is true, and the failed to reject null hypotheses

b. Alternative hypothesis is true, and failed to reject null hypotheses

c. Null hypothesis is true, and null hypothesis is rejected

d. Alternative hypothesis is true, and null hypothesis is rejected.

7. Medical researchers now believe there may be a link between baldness and heart attacks in men.

a. State the null hypothesis and the alternative hypothesis for a study used to investigate whether or not there is such a relationship. (Utts & Heckard, 2012)

b. Discuss what would constitute a Type 1 error in this study. (Utts & Heckard, 2012)

c. Discuss what would constitute a Type 2 error in this study. (Utts & Heckard, 2012)

8. A politician is trying to decide whether to vote for a new tax bill that calls for substantial reforms. A random sample of voters in his district leads him to believe the alternative hypothesis, Ha: p> .5, where p is the proportion of all voters in his district who support the bill as a consequence, he decides to vote for the bill. (Utts & Heckard, 2012)

a. What would a Type 1 error be in this situation, and what would be the consequences for the politician?

b. What would a Type 2 error be in this situation, and what would be the consequences for the politician?

9. Which hypotheses are two tailed?

10. Which hypotheses are one tailed?

11. What is the null standard error formula?

12. What is the standard error formula for sample proportion?

13. What is the z score formula for a proportion?

14. Conditions for conducting the Z test?

15. Conditions for conducting the T test?

16. Find the p value. Z statistic = -1.09; H_0: p = .5; H_a: p < .5. Specify the rejection region for alpha = .05, and then reach a conclusion for the test using the rejection region rule. (Utts & Heckard, 2012)

17. Find the p value. Z statistic = 2.10; H_0: p = .10; H_a: p ≠.10. Specify the rejection region for alpha = .05, and then reach a conclusion for the test using the rejection region rule (Utts & Heckard, 2012)

18. In each of the following, determine whether the conditions for conducting a z test for proportion are met. If not, explain why not.

a. Twenty students are randomly selected from the list of all sorority and fraternity members at a university to determine if a majority of sorority and fraternity students favor a new policy on alcohol on campus. The hypotheses are as follows: H_0: p = .50; H_a: p >.50 (Utts & Heckard, 2012)

b. Twenty employees of a large company are randomly selected to determine whether the proportion of company employees who are left handed exceeds the national proportion of 10% who are left handed. The hypotheses are as follows: H_0: p = .10; H_a: p >.10 (Utts & Heckard, 2012)

19. In a survey of 240 students in an elementary statistics class at the University of California – Davis, 20 said that they were left handed and 200 said that they were right handed. Assume that the students are representative of all students at the school. Does this provide evidence that the proportion of UCD students who are left handed differs from the national proportion of 0.10? Carry out the 5 steps of the hypothesis test for this situation. (Utts & Heckard, 2012)

20. In each case, suppose the sample proportion is $\hat{p} = .75$ and compute the test statistic and the p value. Then make a conclusion using a .05 level of significance. Comment on the relationship between the sample size and the conclusion.
H_0: $p = .70$; H_a: $p \neq .70$ (Utts & Heckard, 2012)

a. n= 100

b. n= 500

21. Power increase when sample size _____

22. The power increase when the difference between the _____ and the _____ increases

23. The managers at Hunter Chemical claim that their major product contains on average 4.0 fluid ounces of caustic material per gallon. They further state that the distribution of caustic material is normal with a standard deviation of 1.3 fluid ounces. If there is too much caustic material, the product will be dangerous. A government inspector is brought in to test the product. She randomly selects a sample of 100, gallon size containers of the product and finds that their mean weight of caustic materials are 4.5 fluid ounces per gallon. Carry out the hypotheses test for this situation. (Utts & Heckard, 2012)

24. It has been hypothesized that the mean pulse rate for college students is about 72 beats per minute. A sample of Penn state students recorded their pulse rate. Assume that the sample is representative of all Penn State women for pulse rate measurements, where they surveyed 35 students, and the mean was 76.9, and the standard deviation was 11.6. What is the hypothesis test conclusion for this scenario? (Utts & Heckard, 2012)

25. Perry wanted to test the folklore that women have not been given information about the sex of their unborn child can guess it at better than chance levels. They asked a sample of 104 pregnant women to guess the sex of their babies, and 57 guessed correctly. Assuming chance guessing would result in 50% correct guesses, test the hypothesis that women can guess at a better than chance level. Carry out the test using $\alpha = 0.05$, and make sure to state a conclusion. (Utts & Heckard, 2012)

26. A university is concerned that it is taking students too long to complete their requirements and graduate; the average time for all students is 4.7 years. The dean of the campus honors program claims that students who participated in that program in their first year have had a lower mean time to graduation. A random sample of 30 students who had participated is taken, and the mean and standard deviation for the time to completion for those students are 4.5 years and 0.5 year, respectively. Carry out the 5 steps to test the hypotheses of interest in this situation. Use $\alpha = 0.05$. (Utts & Heckard, 2012)

27. Gasoline prices have remained higher for state along the West Coast, particularly in Alaska and California. Let's say we drew a random sample of 100 California gas stations, and the mean price was $3.11. We also know that nationally the mean price of a gallon was $2.86, with a standard deviation of 0.17. We expect to find that the average price of gas from California gas stations will be higher than before. Carry out the steps to test the hypotheses of interest. Use $\alpha = .05$. (Utts & Heckard, 2012)

28. Compute the effect size for each of the following situations, and state whether it would be considered closer to a small, medium, or large effect: (Utts & Heckard, 2012)

a. In a paired difference test with $n = 100$ pairs, the test statistics is $t = 2.24$

b. In a paired difference test with $n = 50$ pairs, the test statistics is $t = -2.83$

29. What is the formula to compute the effect size for one proportion?

30. Identify whether a scatterplot would or would not be an appropriate visual summary of the relationship between the following variables. In each case, explain your reasoning. (Utts & Heckard, 2012)

a. Blood pressure and age

b. Region of country and opinion about stronger fun control laws.

31. The website gives age in years and body temp in degrees Fahrenheit for 100 blood donors ranging in age 17 to 84 years old. The regression equation is $\hat{y} = 98.6 - 0.0138x$. (Utts & Heckard, 2012)

a. In the regression relationship shown, which variable is the response and which is the explanatory variable?

b. What is the predicted body temp for someone who is 50 years old?

c. One of the donors was 50 years old and had a body temp of 97.6. What is the residual for this person? Explain what the residual tells you about this person's body temp in comparison to the average body temp for someone his age.

32. A regression equation for y = handspun (cm) and x = height (inches) . if the roles of the variables are revered and only women are considered, the regression equation is

Average height = 51.1 + 0.7 (handspun) (Utts & Heckard, 2012)

a. Interpret the slope of 0.7 in terms of how height changes as handspan increase.

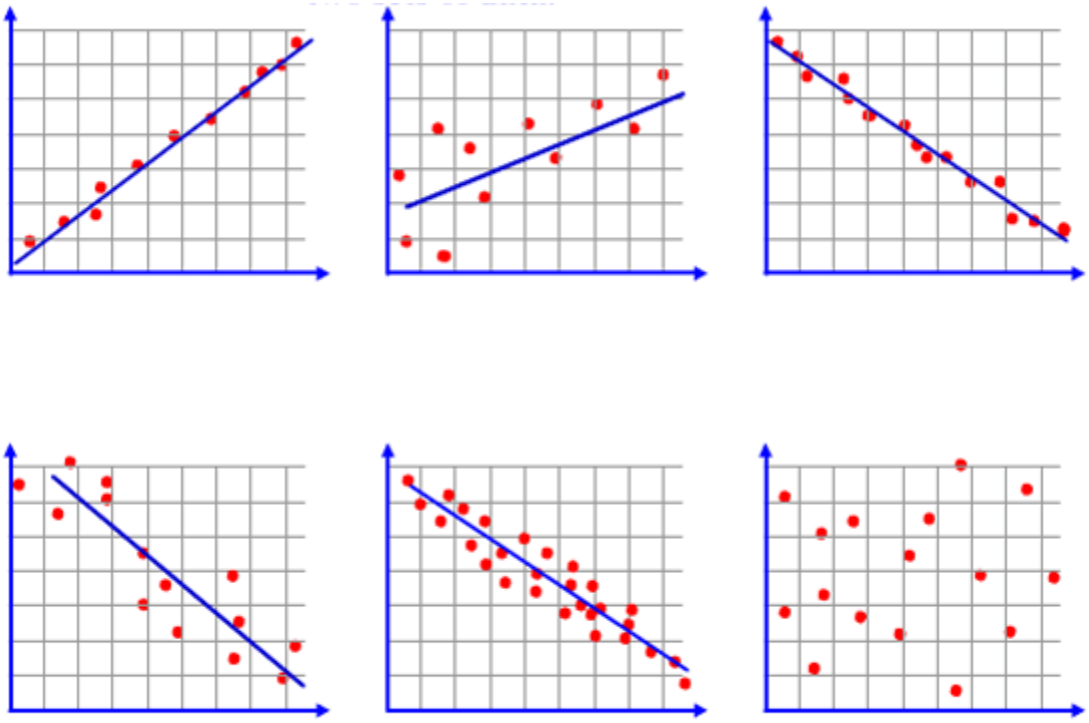
b. What is the estimated average height of women with a handspan of 20 cm?

c. Molly has a handspan of 20 cm and is 66.5 inches tall. What is the prediction error for Molly?

33. Which implies a stronger linear relationship: a correlation of + 0.4 of a correlation of -0.6? Briefly explain.

34. The correlation between latitude and average August temp is -0.78 for the 20 cities. Calculate coefficient of determination and write a sentence that interprets it in the context of this situation. (Utts & Heckard, 2012)

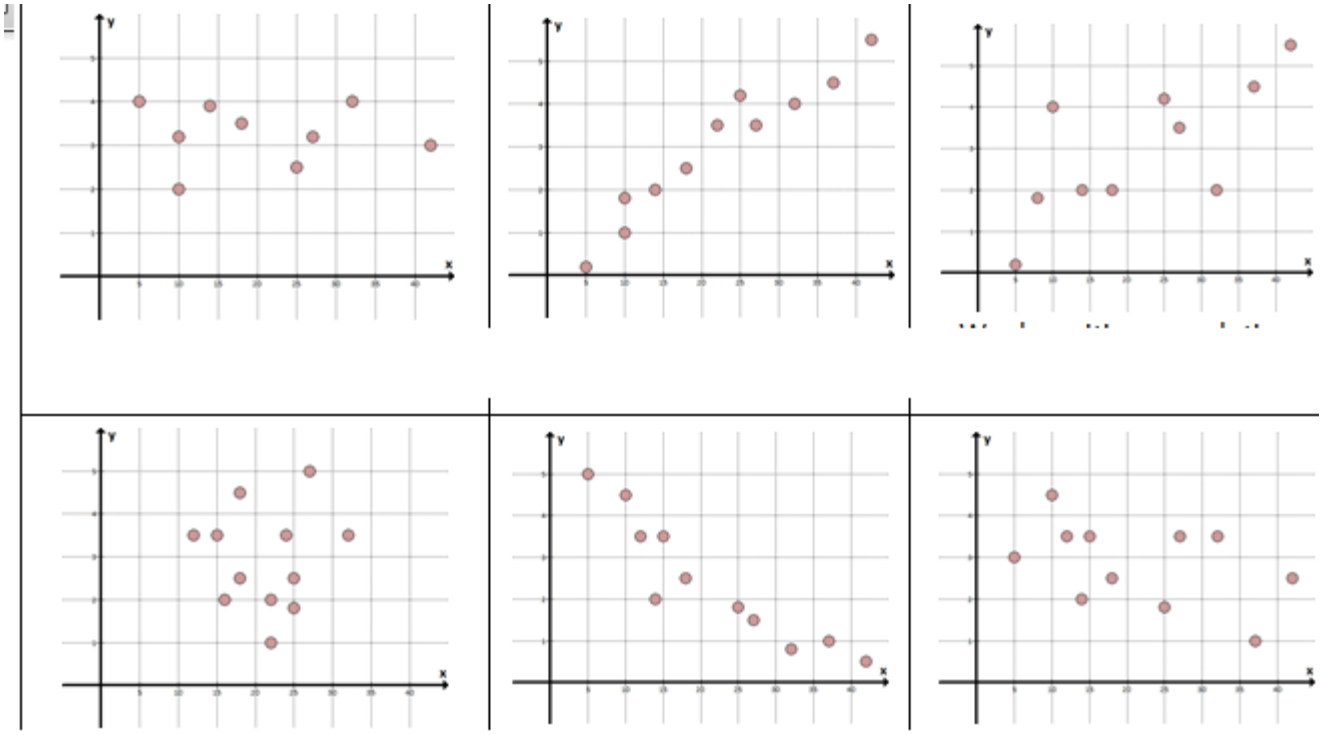
35. State whether each graph has a strong, weak or positive, negative correlation.



(Utts & Heckard, 2012)

36. Match the coefficient correlation with the appropriate graph (Utts & Heckard, 2012)

a. -0.05 b. 0.71 c. -0.45 d. 0.96 e. 0.04 f. -0.92



37. What is the correlation coefficient when $r^2 = 0.56$?

38. What is the formula for r^2?

Exam3 Review Answers

1) Null: statement to be tested; hoping to reject or disapprove
Alternative: statement that we are trying to find evidence to support

2) α

3) **a.** $H_a: p \neq .30.$
b. $H_a: p > .45.$
c. $H_a: p < .60.$
(Utts & Heckard, 2012)

4) a. Ho: p = .15
Ha: p ≠ .15

c. Ho: p = .30
Ha: p < .30 (Utts & Heckard, 2012)

---

5) **a.** The p-value is .03.
b. The level of significance is $\alpha = 0.05$.
c. The sample proportion is $\hat{p} = 61/100 = .61$.
d. The sample size is $n = 100$.
e. The null value is .5. If there were no difference in preference, then 50% of the population of consumers would prefer each type.
(Utts & Heckard, 2012)

---

6)

		DECISION	
		Reject H_o	*Fail to Reject H_o*
ACTUAL	*H_o True*	Type I Error ***Producer Risk*** **α-Risk** **False Positive**	**Correct Decision** **Confidence Interval = 1- α**
	H_a True	**Correct Decision** **Power = 1- β**	Type II Error ***Consumer Risk*** **β-Risk** **False Negative**

H_o: Null Hypothesis H_a: Alternative Hypothesis (Utts & Heckard, 2012)

---

7) **12.31** **a.** The correct decision has been made. The null is true and it is not rejected.
b. A type 2 error has occurred. The alternative is true, but the null is not rejected.
c. A type 1 error has occurred. The null is true, but it has been rejected.
d. The correct decision has been made. The alternative is true and the decision is in favor of the alternative.

12.34 **a.** *Null*: No relationship between baldness and heart attacks.
Alternative: There is a relationship between baldness and heart attacks.
b. A type 1 error would occur if it is decided there is a relationship between baldness and heart attacks when in reality there is no relationship.
c. A type 2 error would occur if there is not enough evidence to conclude that there is a relationship between baldness and heart attacks when in reality there is a relationship.

(Utts & Heckard, 2012)

a. A type 1 error occurs if the politician believes that more than one-half of voters in his district support the new tax bill when the proportion really is not more than one-half. The consequence is that he would vote for a bill that is not supported by a majority of the voters in his district.

b. A type 2 error occurs if the politician believes the proportion of voters in the district supporting the tax bill is not a majority when really it is a majority. The consequence is that that he would not vote for a bill that is supported by a majority of the voters.

8) (Utts & Heckard, 2012)

9) $P \neq P_0$

10) $P > P_0$ AND $P < P_0$

11) $= \sqrt{\frac{p_0(1-p_0)}{n}}$, (Utts & Heckard, 2012)

12) $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ (Utts & Heckard, 2012)

13) $$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$ (Utts & Heckard, 2012)

1. The sample is a simple random sample.
2. np_0 and $np_0(1-p_0)$ should be at least 10.

14) (Utts & Heckard, 2012)

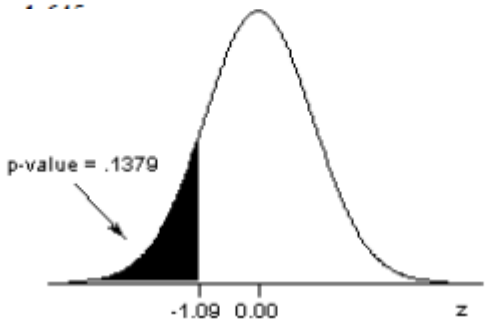
15) The sample is SRS, and N $\geq$ 30(Utts & Heckard, 2012)

c. p-value = .1379. This is the probability (area) to the left of $z = -1.09$. Table A.1 can be used to find $P(Z < -1.09) = .1379$.

c. Rejection region is $z < -1.645$.

Cannot reject the null hypothesis. z-statistic = –1.09 is not less than –1.645.

16)



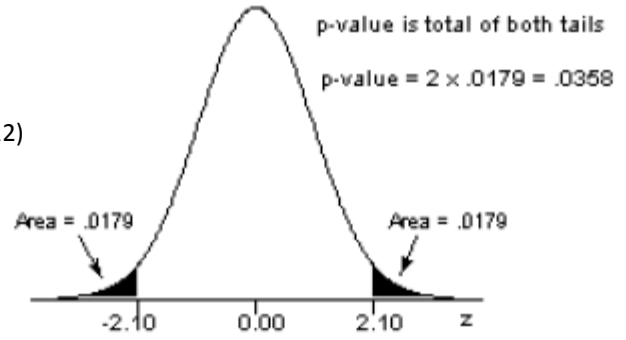
(Utts & Heckard, 2012)

12.41 **a.** p-value = .0358. This is the combined probability to the right of $z = 2.10$ and left of $z = -2.10$. Table A.1 can be used to find that $P(Z < -2.10) = .0179$. By symmetry, the area to the right of $z = 2.10$ is also .0179. So, the p-value = $2 \times .0179 = .0358$. (Utts & Heckard, 2012)

a. Rejection region is $|z| > 1.96$.

Reject the null hypothesis. z-statistic = 2.10 is greater than 1.96.

(Utts & Heckard, 2012)



17)

12.43 **a.** Yes. The sample is a random sample, and the sample size is large enough because $np_0 = (20)(.50) = 10$, and $n(1-p_0) = (20)(1-.50) = 10$. (Utts & Heckard, 2012)

b. No. The sample size is not large enough because $np_0 = (20)(.10) = 2$ is smaller than 10.

18)

12.91 Step 1: H_0: $p = .1$ (proportion left-handed same as in national population)

H_a: $p \neq .1$ (proportion left-handed not same as in national population.

Step 2: The necessary conditions for using the z-statistic are present. The sample is assumed to be representative of the larger population and the sample size is large enough so that both np_0 and $n(1-p_0)$ are greater than 10. Here, $n = 240$ and $p_0 = .1$.

The test statistic is $z = \dfrac{\text{Sample estimate - Null value}}{\text{Null standard error}} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$.

Sample estimate = $\hat{p} = 20/240 = .0833$

$$z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{.0833 - .1}{\sqrt{\dfrac{.1(1-.1)}{240}}} = \frac{-.0167}{.01936} = -0.86$$

Step 3: p-value ≈ .39. This is the combined probability (area) to the left of $z = -0.86$ and to the right of $z = 0.86$. Table A.1 gives $P(z < -0.86) = .1949$. Because this is a two-sided test, p-value = $2 \times .1949 = .3898$.

Note: An exact p-value based on the binomial distribution is given by Minitab as .395.

Step 4: Cannot reject the null hypothesis. The result is not statistically significant because the p-value is not smaller than .05, the usual standard for significance.

Step 5: We cannot conclude that the proportion of UCD students who are left-handed differs from the national proportion.

19)

(Utts & Heckard, 2012)

b. $\sqrt{\dfrac{.70(1-.70)}{100}} = .0458$. **b.** $\dfrac{.75-.70}{.0458} = 1.091$; $p \approx 2 \times P(Z < -1.09) = (2)(.1379) = .2758$, so fail to reject null hypothesis.

c. $\sqrt{\dfrac{.70(1-.70)}{500}} = .0205$. **c.** $\dfrac{.75-.70}{.0205} = 2.439$; $p \approx 2 \times P(Z < -2.44) = (2)(.0073) = .0146$, so reject the null hypothesis.

(Utts & Heckard, 2012)

20)

21) Power increases when the sample size increases

22) The power increase when the difference between the true population value and the null hypothesis value increases

23) $P(\bar{X} \geq 4.5 \mid \mu = 4) = P(\frac{X-\mu}{\sigma/\sqrt{n}} \geq \frac{4.5-4}{1.3/\sqrt{100}}) = P(Z \geq 3.85) = .0001$ (Utts & Heckard, 2012)

H0: μ = 0.4 Ha: μ > 0.4

REJECT THE NULL HYPOTHESES

---

b. Step 1: H_0:μ = 72 versus H_a:μ ≠ 72.

Step 2: Sample size (35) is sufficiently large to proceed. We assume the sample represents a random sample of Penn State women.

μ = mean pulse rate for population of Penn State women

Test statistic is $t = \frac{\text{Sample statistic - Null value}}{\text{Null standard error}} = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{76.9-72}{\frac{11.6}{\sqrt{35}}} = \frac{4.9}{1.961} = 2.50$

Steps 3,4, and 5: df = $n - 1 = 35 - 1 = 34$. The p-value is $2 \times P(t > 2.50)$. Using Table A.3, the p-value is between 2(.008) = .016 and 2(.013) = .026. We can reject the null hypothesis and conclude that mean female pulse rate is significantly different from 72. With software like Excel or Minitab, or a suitable calculator, it can be found that the p-value is 2(.0087) = .0174.

c. The data provide evidence that the mean pulse rate of Penn State women is different from 72 beats per minute, but do not give evidence that the mean pulse rate for Penn State men is significantly different from 72. (Utts & Heckard, 2012)

24) **12.111** Step 1: H_0: $p \leq .50$ (not better than chance level)

H_a: $p > .50$ (predict better than chance level)

p = proportion of population of pregnant women that can predict the sex of their babies

Step 2: We must assume the sample represents a random sample from the population of pregnant women. The sample size is sufficiently large so that $n\hat{p}$ and $n(1-\hat{p})$ are both greater than 10.

Sample proportion correct guesses is $\hat{p} = \frac{57}{104} = .548$

Test statistic is $z = \frac{\text{Sample statistic - Null value}}{\text{Null standard error}} = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{.548-.50}{\sqrt{\frac{.50(1-.50)}{104}}} = \frac{.048}{.049} = 0.98.$

Step 3: p-value = .1635. It is the area (probability) to the right of $z = 0.98$.

$P(z > 0.98) = P(z \leq -0.98) = .1635$. Equivalently, $P(z > 0.98) = 1 - P(z \leq 0.98) = 1 - .8365 = .1635$.

Steps 4 and 5: We do not reject the null hypothesis for $\alpha = .05$. There is not enough evidence to conclude that in the population of pregnant women represented by the sample, the proportion able to predict the sex of their babies is higher than .50 (the chance level).

Note: Minitab could be used to do this exercise (see Minitab Tip on page 485). The program reports the exact I-value based on binomial distribution probabilities as .189. (Utts & Heckard, 2012)

25) **13.29** Step 1: H_0:μ = 4.7, H_a:μ < 4.7, where μ = mean time to graduate (in years) for students who participated in the honors program in their first year of college.

Step 2: Sample size (30) is sufficiently large to proceed. It is stated in the problem that we can assume the sample represents a random sample of the population of students who participated in the honors program in their first year of college.

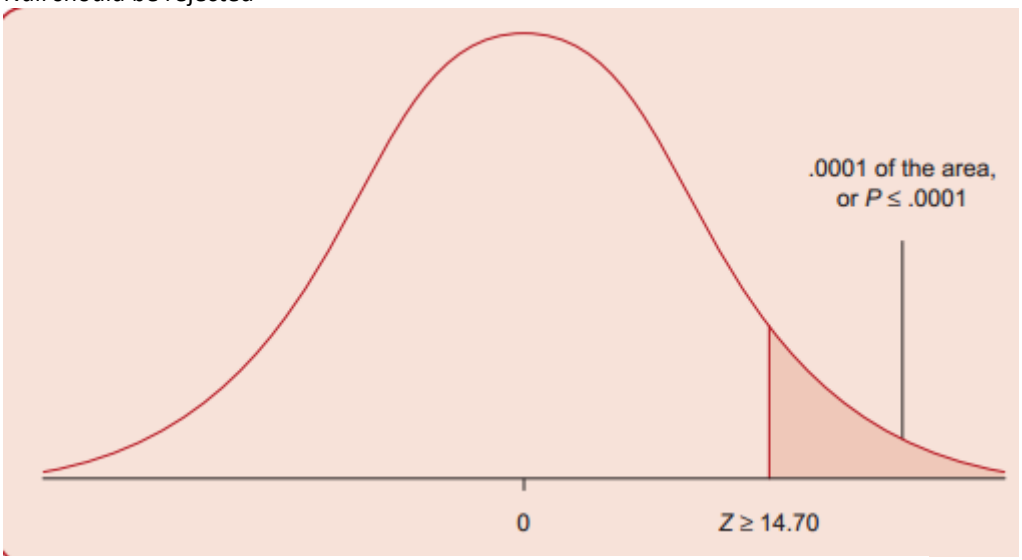
Test statistic is $t = \frac{\text{Sample statistic - Null value}}{\text{Null standard error}} = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{4.5-4.7}{\frac{0.5}{\sqrt{30}}} = \frac{-0.2}{0.0913} = -2.19$

Steps 3,4, and 5: df = $n - 1 = 30 - 1 = 29$. The p-value is $P(t < -2.19)$. Using Table A.3, the p-value is between .013 and .027. With software like Excel or Minitab, or a suitable calculator, it can be found that the p-value is .018. Using the standard $\alpha = .05$, we can reject the null hypothesis and conclude that the mean time to graduate is lower for the population of students who participated in the honors program than it is for the general population of students. (Utts & Heckard, 2012)

26)

$Z = \frac{3.11 - 2.86}{0.17/\sqrt{100}} = 14.70$ (Utts & Heckard, 2012)

27) Null should be rejected $H_0: \mu_Y = \$2.86$ $H_1: \mu_Y > \$2.86$



13.77 **a.** Effect size is $\frac{2.24}{\sqrt{100}} = 0.224$; small.

b. Effect size is $\frac{-2.83}{\sqrt{50}} = -0.4$; closer to medium. (Utts & Heckard, 2012)

28) d(hat) = t/ n(1/2)

$$ES = \frac{|p_1 - p_0|}{\sqrt{p_0(1-p_0)}}$$ (Utts & Heckard, 2012)

29)

3.5 **a.** Appropriate because both variables are quantitative.

b. Not appropriate because both variables are categorical. (Utts & Heckard, 2012)

30)

3.75 **a.** Body temperature is the response variable and age is the explanatory variable.

b. $\hat{y} = 98.6 - 0.0138x = 98.6 - 0.0138(50) = 97.91$ degrees. (Utts & Heckard, 2012)

c. The predicted temperature for a person whose age = 50 was found in part (b) to be 97.91. Residual = actual – predicted = 97.6 – 97.91 = –0.31. His body temperature is 0.31 degrees lower than would be predicted based on his age.

31)

3.19 **a.** For each one-centimeter increase in handspan, average height increases 0.7 inches.

b. $\hat{y} = 51.1 + 0.7(20) = 65.1$ in.

c. $66.5 - 65.1 = 1.4$ in. (Utts & Heckard, 2012)

32)

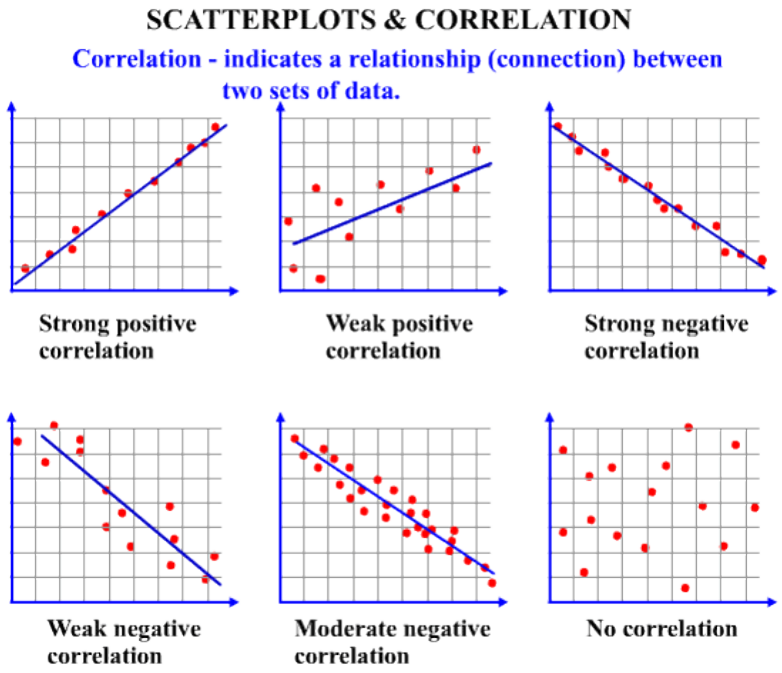
3.32 A correlation of –0.6 implies a stronger linear relationship than a correlation of +0.4 because in absolute magnitude it is closer to 1. To judge the strength of a linear relationship, ignore the sign of a correlation coefficient. (Utts & Heckard, 2012)

33)

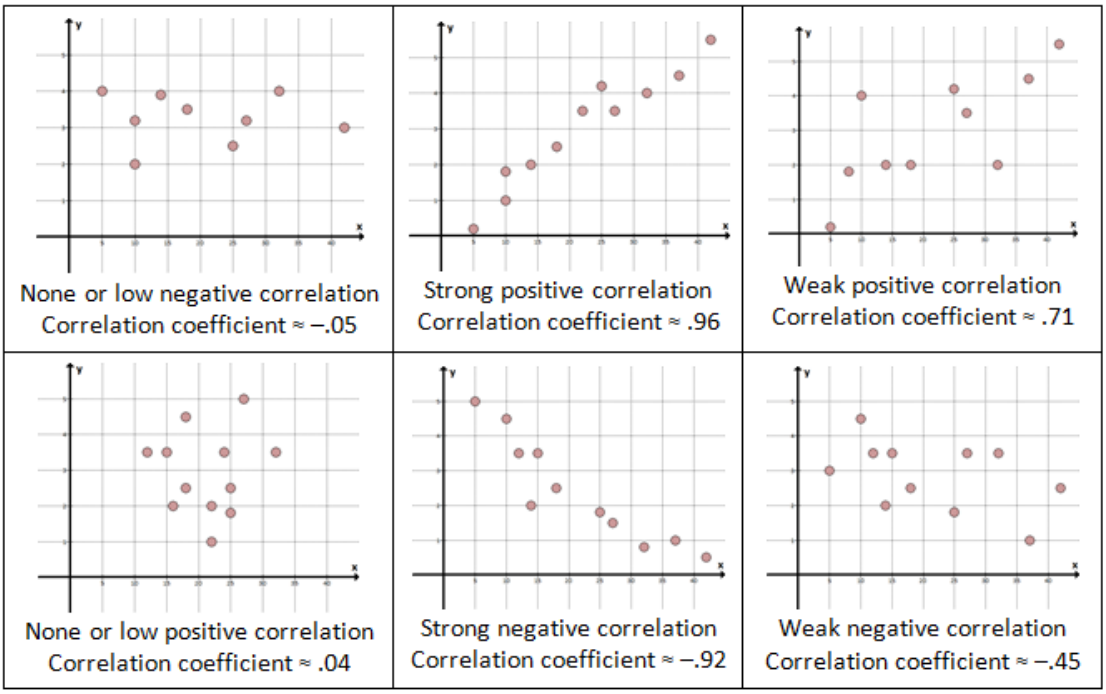
3.44 **a.** $r^2 = (-0.78)^2 = .6084$. This means that latitude explains about 60.8% of the variation in average August temperature. (Utts & Heckard, 2012)

34)

35)



(Utts & Heckard, 2012)



(Utts & Heckard, 2012)

36)

37) R = .748

38) R^2= explained variation/ total variation = 1- (unexplained variation/ total variation)

Utts, J., & Heckard, R. (2012). *Mind on statistics* (Fifth ed., p. 739). Stamford: Cengage Learning.

References:

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.

SIT Session Lesson Plan

Week/Chapter: Week 14 Course Assistant: _____

Course: Introduction to Statistics Instructor: _____

Objective: What are the one or two most difficult concepts that the students need to work on today? Post Exam 3 and Final Exam Review

Beginning reminders:

1. Arrange seats in a circle
2. Make sure everyone has signed in
3. Review lesson plan with group
4. Remember to relax and be flexible!

Content to Cover:	Processes to Use*:	Time
Opener:	Have everyone say there name and most challenging concept for test three	5 Minutes
Post Exam 3 Review	After students give their most challenging aspect of test three (I will write them down as they list them) I will ask each person if they fill they know any of the listed concept very well. I will write their names down by each that they are confident in. Each student will have a chance to explain a concept to the others.	20 Minutes
Final Exam Review	Review for final exam. Review every chapter and point out practice problems in the book. We will take turns being scribe and using the board, and work together to solve each problem.	20 Minutes
Tutoring = Q & A		5-10 Minutes

After session comments/thoughts:

References:

Utts, Jessica M. and Robert F. Heckard. (2012). *Mind on Statistics* (5th ed.). Stamford, CT: Cengage Learning.

*Most activities are adaptations of the SI Strategy Cards from the International Center for Supplemental Instruction located at UMKC and may be found in:

Leader Resource Manual for Supplemental Instruction (SI). (2004). Amanda McDaniel, ed. University of Missouri – Kansas City.