Class Prep Quiz #1

1. A social psychology study administers questionnaires to 150 students from a large public university that ask about behaviors and social norms around texting. What is the likely population for the study?
   All possible college-age individuals in the US

2. Beta-carotene supplements have been thought to protect against cancer. A study published in the Journal of the National Cancer Institute randomly assigned 200 adults to receive a beta-carotene supplement or a placebo, and their health was studied over their lifetime. Which level of measurement describes the treatment assignment variable (beta-carotene supplement vs. placebo)?
   Nominal

3. A social psychology study administers questionnaires to 150 students from a large public university that ask about behaviors and social norms around texting. One of the questions asks whether it is appropriate to carry on a text conversation while on a date. The question uses a Likert scale with options 1 = "strongly disagree", 2 = "disagree", 3= "neither agree nor disagree", 4 = "agree", and 5 = "strongly agree". Which of the following is true of this variable? Check all that apply.
   In practice, this type of scale is often close enough to interval that researchers treat it like any other numeric scale
   For this to be an interval scale, the amount of agreement between "strongly disagree" and "disagree" must be the same amount between "agree" and "strongly agreeing"

4. Which of the following variables is measured on an ordinal scale?
   Highest level of education (high school, some college, college, post BA)

5. Which of the following variables is measured on an interval scale?
   Number of days per month having at least one alcoholic drink

6. Beta-carotene supplements have been thought to protect against cancer. A study published in the Journal of the National Cancer Institute randomly assigned 200 adults to receive a beta-carotene supplement or a placebo, and their health was studied over their lifetime. Who is the likely population for this study?
   All possible adults in the US

7. An industrial organizational psychologist working for a large company summarizes data from employee performance ratings that range from "underperforming", "average", "above average", and "superior". What level of measurement does the performance rating variable use?
   Ordinal

8. Which of the following describes the relationship between a sample and a population?
   We analyze data from a sample of participants, with the goal of learning something about an entire population of potential participants

9. A social psychology study administers questionnaires to 150 students from a large public university that ask about behaviors and social norms around texting. The 150 students constitute what group?
   A convenience sample

10. Which of the following variables is measured on a nominal scale?
    A psychological diagnosis (antisocial personality disorder, schizophrenia, bipolar disorder)

Content Quiz #1

1. Beta-carotene supplements have been thought to protect against cancer. A study published in the Journal of the National Cancer Institute randomly assigned 200 adults to receive a beta-carotene supplement or placebo, and their health was studied over their lifetime. Who is the likely population for this study

   All possible adults in the US

2. A study examining college student texting behavior asks the following questions:

   > 1. How many times per hour do you send a text message? "0", "1-5", "6-10", "11-50", "50 or more".
   > 2. Texting a friend while having a phone conversation with another friend is acceptable behavior: "Strongly Disagree", "Disagree", "Neutral","Agree","Strongly Agree"

   Which question would psychologists say most closely approximates an interval (i.e., numeric, continuous) measurement scale?

   Question 2

3. A social psychology study administers questionnaires to 150 students from a large public university that ask about behaviors and social norms around texting. One of the questions asks whether it is appropriate to carry on a text conversation while on a date. The question uses a Likert scale with options 1 = "strongly disagree", 2 = "disagree", 3= "neither agree nor disagree", 4 = "agree", and 5 = "strongly agree". Which of the following is true of this variable? Check all that apply.

   For this to be an interval scale, the amount of agreement between "strongly disagree" and "disagree" must be the same amount between "agree" and "strongly agree"

   In practice, this type of scale is often close enough to interval that researchers treat it like any other numeric scale

4. A social psychology study administers questionnaires to 150 students from a large public university that ask about behaviors and social norms around texting. What is the likely population for the study?

   All possible college-age individuals in the US

5. A study examining the use of and engagement with Facebook asked participants to report their exa ct number of friends. Which level of measurement describes this variable?

   Interval(i.e., numeric, continuous)

6. Beta-carotene supplements have been thought to protect against cancer. A study published in the Journal of the National Cancer Institute randomly assigned women to receive a beta-carotene supplement or a placebo, and their health was studied over their lifetime. Which level of measurement describes the treatment variable (beta-carotene supplement vs. placebo)?

   Nominal

7. A study examining the use of and engagement with Facebook asked participants about weekly minutes spent logged on. Which level of measurement describes this variable?

   Interval(i.e., numeric, continuous)

8. A social psychology study administers questionnaires to 150 students from a large public university that ask about behaviors and social norms around texting. The 150 students constitute what group?

   A convenience sample

9. A study examining the use of and engagement with Facebook asked participants the following question: "I feel out of touch when I haven't logged onto Facebook for a while." Participants indicated their level of agreement or disagreement using a 5-point Likert-type rating scale ranging from 1 ("Strongly Disagree") to 5 ("Strongly Agree"). Which level of measurement describes this variable?

   Ordinal

10. A national survey of smoking behavior classifies respondents as non-smokers, non-daily smokers, and daily smokers. What is the level of measurement for this variable?

    Ordinal

11. Which of the following variables is measured on an ordinal scale?

12. A study examining the use of and engagement with Facebook asked the following two questions:

> 1. How many friends do you actively follow on Facebook? "10 or less", "11–50", "51-100", "100 or more".
>
> 2. I feel out of touch when I haven't logged onto Facebook for a while: "Strongly Disagree", "Disagree", "Neutral", "Agree", "Strongly Agree"

Which question would psychologists say most closely approximates an interval (i.e., numeric, continuous) measurement scale?

Question 2

13. Which of the following variables is measured on a nominal scale?
    Political Affiliation/ethnicity

14. A depression questionnaire asks the question, "I felt that I had nothing to look forward to." Respondents answered on a scale ranging from 1 (Strongly disagree) to 5 (Strongly agree). Which level of measurement describes this variable?
    The correct answers are: Ordinal, Quasi-interval

15. Which of the following variables is measured on an ordinal scale?
    Self-reported health (poor, fair, good, excellent)

16. A study examining the use of and engagement with Facebook asked participants to report the number of friends. Respondents chose between nine options of custom ranges, e.g., "10 or less", "11–50", "51-100", etc. Which level of measurement describes this variable?
    Ordinal

17. An industrial organizational psychologist working for a large company summarizes data from employee performance ratings that range from "underperforming", "average", "above average", and "superior". What level of measurement does the performance rating variable use?
    Ordinal

18. Which of the following variables is measured on an interval scale?
    Number of days per month having at least one alcoholic drink

19. Which of the following describes the relationship between a sample and a population?
    We analyze data from a sample of participants, with the goal of learning something about an entire population of potential participants

20. Which of the following variables is measured on an interval (i.e., numeric, continuous) scale?
    IQ

21. A computer-administered measure of prejudice presents participants with a statement about a particular ethic group and then records the amount of time it takes participants to respond to the prompt. What is the level of measurement for the reaction time variable?
    Interval (i.e., numerical, continuous)

22. Which of the following variables is measured on an interval (i.e., numeric, continuous) scale?
    Time (in minutes) spent on social media per day

23. A national survey of smoking behavior asks respondents to report their family's total household income to the nearest dollar. What is the level of measurement for the income variable?
    Interval

24. A research studying mathematics anxiety asks a sample of students the following question: "I feel that my anxiety toward math is a barrier for scoring well on the exam." Participants indicated their level of agreement or disagreement using a 5-point Likert-type rating scale ranging from 1 ("Strongly Disagree") to 5 ("Strongly Agree"). Which level of measurement describes this variable?
    Quasi-Interval or Ordinal

25. Beta-carotene supplements have been thought to protect against cancer. A study published in the Journal of the National Cancer Institute randomly assigned 200 adults to receive a beta-carotene supplement or a placebo, and their health was studied over their lifetime. Which level of measurement describes the treatment assignment variable (beta-carotene supplement vs. placebo)?
Nominal

26. A national survey of smoking behavior classifies respondents into four ethnic groups: white non-hispanic, black non-hispanic, hispanic, and other non-hispanic. What is the level of measurement for this variable?
Nominal

2. Summarizing Data

Class prep quiz #2

1. A survey asks respondents to rate the degree to which they are favorable or unfavorable toward some type of nationalized healthcare service (e.g., Medicare for all, Obamacare, or some comparable plan). Which of the following distribution shapes displays a set of results where most respondents are very favorable toward such a plan?


(A)
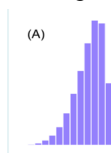
(A)

2. Which of the following is the best definition of the term "distribution"?
A table or graph that displays all possible values of a variable in a sample of data as well as how often each score occurs

3. A chemistry midterm test is very difficult, with relatively few students passing. Which of the following distributions most likely describes the test scores?

(A)   (B)   (C)
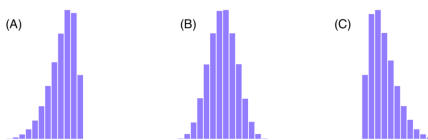
(C)

4. A study of smoking trends in the US uses a variable called smoking intensity that combines frequency (e.g., how many days) and amount (e.g., cigarettes per day) into a single numeric value. The distribution is shown below. What type of graph is being displayed?
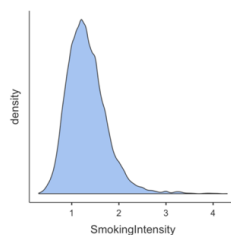
Kernel Density

5. A questionnaire item asks 1500 respondents for their highest level of education. A frequency distribution table of the responses is shown below. What value b my no no

6. elongs in place of the red question mark?

187

How was this value found? I tried to make ratios of the percentage 12.5% with the missing value and the other values with the remaining 87.5% and I got 187.something pls HELP
Yeah there's 1500 respondents so I multiplied 1500 x 0.125 and got 187.5 but there can't be .5 of a person so it's just 187 I think
I thought we just added the counts then do 1500-the counts but idk lol thats how i got 187
Add the values then subtract it from 1500

7. The standard deviation – a measure of "spread" or "dispersion" – is a statistic we will use throughout the class. What information does this value provide?

How similar or different scores are for a group of participants

8. A study of smoking trends in the US uses a variable called smoking intensity that combines frequency (e.g., how many days) and amount (e.g., cigarettes per day) into a single numeric value. The distribution is shown below. Which of the following is true?
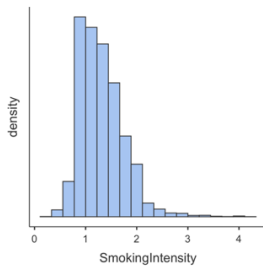
Most respondents had fairly low smoking intensity scores, and high intensity scores are less common

9. What is meant by the "center" of a distribution?

A typical or common score that is close to the peak of the distribution

10. A study of smoking trends in the US uses a variable called smoking intensity that combines frequency (e.g., how many days) and amount (e.g., cigarettes per day) into a single numeric value. The distribution is shown below. Which of the following is true?
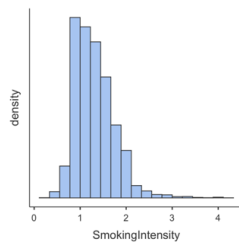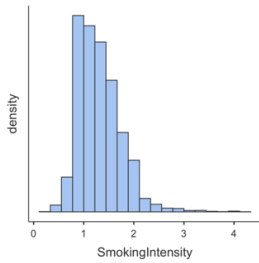

The distribution is positively (right) skewed

11. A study of smoking trends in the US uses a variable called smoking intensity that combines frequency (e.g., how many days) and amount (e.g., cigarettes per day) into a single numeric value. The distribution is shown below. What type of graph is being displayed?

density

SmokingIntensity

Histogram

Content Quiz #2

1. A distribution of midterm exam scores has a strong negative (left) skew. Which of the following is true?
   Few people have very low scores

2. Which of the following graphs best corresponds to a positively (right) skewed distribution?

Answer: SD   D   A   SA
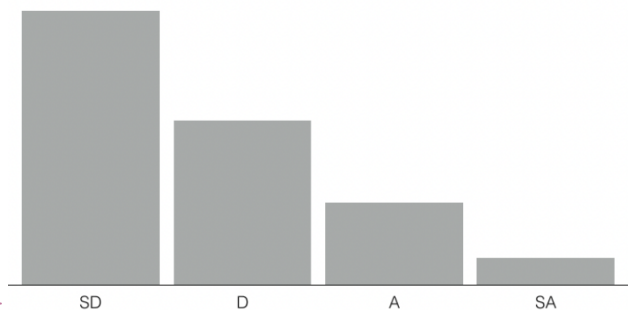
3. Consider the following frequency distribution for a survey question asking about agreement on a political ballot initiative (the Frequency column is the same as the Count). Assuming that the total number of respondents is 250, what is the value for the frequency in the cell labeled X?

Response	Frequency	Percent	Cumulative Percent
Strongly agree	50	20	100
Agree	50	20	80
Disagree	X	10	60
Strongly disagree	125	50	50

25

4. In a frequency distribution, which quantity describes the number of respondents providing a particular response?
   Count (Frequency)

5. A depression questionnaire asks the following three items, each of which is measured on a 1 (Strongly disagree) to 5 (Strongly agree) scale.

1. I felt that I had nothing to look forward to
2. I felt down-hearted and blue.
3. I felt that life was meaningless.

Consider the following score profile: 1, 2, 3. What is the value of a scale score computed as the average level of agreement across the questions?

2

6. A survey asks respondents to rate the degree to which they are favorable or unfavorable toward some type of nationalized healthcare service (e.g., Medicare for all, Obamacare, or some comparable plan). Which of the following distribution shapes displays a set of results where most respondents are very favorable toward such a plan? Note that the horizontal axis goes from low or negative (left) to high or positive (right).

A

7. Which of the following is the best definition of the term "distribution"?

A table or graph that displays all possible values of a variable in a sample of data as well as how often each score occurs

8. A study of post-partum depression found a relatively even number of mothers with low, moderate, and high depression levels. What distribution shape best describes the depression variable?
Uniform

9. A study of post-partum depression includes a questionnaire item asking mothers to report the number of children in the family. Most mothers reported having only one child, and the number of mothers reporting 2, 3, and 4 children steadily decreased. What distribution shape best describes this variable?
Positively (right) skewed

Response	Frequency	Percent	Cumulative Percent
Strongly agree	120	24	100
Agree	125	25	76
Disagree	125	25	51
Strongly disagree	130	26	26

10. Consider the following frequency distribution for a survey question asking about agreement on a political ballot initiative (the Frequency column is the same as the Count). What distribution shape best describes this question? In

Uniform

11. The distribution of income in the U.S. has a strong positive (right) skew. Which of the following is true?
Few people have very high incomes

12. A questionnaire item asks 1500 respondents for their highest level of education. A frequency distribution table of the responses is shown below. What value belongs in place of the red question mark?

187

Frequencies of Education

Levels	Counts	% of Total
Less than High School	373	24.9%
High School	554	36.9%
Some College	386	25.7%
College	???	12.5%

13. A study of post-partum depression found that most mothers reported relatively low levels of depression, and higher levels of depression were relatively uncommon. What distribution shape best describes this variable?
Positively (right) skewed

14. Consider the following frequency distributions from an attitudinal questionnaire. Which of the following distributions represents a positively (right) skewed distribution?
Strongly Agree 25
Agree 75

Disagree 150
Strongly Disagree 250

15. Consider the following frequency distribution for a survey question asking about highest level of education (the Frequency column is the same as the Count). What is the value for the percent in the cell labeled X?
    10

16. A depression questionnaire asks the following three ittems, each of which is measured on a 1 (Strongly disagree) to 5 (Strongly agree) scale.

    1. I felt that I had nothing to look forward to
    2. I felt down-hearted and blue.
    3. I felt that life was meaningless.

Consider the following score profile: 3, 4, 5. What is the value of a scale score computed as the average level of agreement across the questions?
    4

17. Consider the following frequency distribution for a survey question asking about agreement on a political ballot initiative (the Frequency column is the same as the Count). What distribution shape best describes these responses?
    Positively right skewed

18. A researcher forms a scale score by averaging responses to the three items. What is the minimum value for the scale score?
    1

19. Researchers are interested in studying factors related to teenage drinking frequency (higher scores = more frequent drinking). In particular, the researchers want to determine whether growing up with an alcoholic parent versus not has an influence on the frequency of drinking. Which of the following results would indicate that children of alcoholic parents drink more frequently?
    The centers of the two distributions differ, with children of alcoholics centered at a higher value

20. A study of post-partum depression found that most mothers reported moderate levels of depression, and very low and very high depression scores were relatively rare. What distribution shape best describes this variable? Normal

21. A depression questionnaire asks the following three items, each of which is measured on a 1 (Strongly disagree) to 5 (Strongly agree) scale.

    1. I felt that I had nothing to look forward to
    2. I felt down-hearted and blue.
    3. I felt that life was meaningless
    A researcher forms a scale score by averaging responses to the three items. What is the maximum value for the scale score? 5

22. A study of post-partum depression includes a questionnaire item asking mothers to report the number of children in the family. The highest frequency was associated with having only one child, and the frequency of mothers reporting 2, 3, and 4 children steadily decreased. What distribution shape best describes this variable?
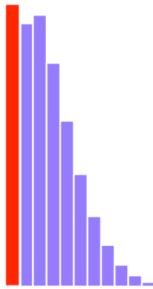   i.   Positively (right) skewed

Class Prep Quiz #3

1. The following graph displays the distribution of chronic pain ratings for a sample of study participants. The red bar represents the location of which measure of central tendency?
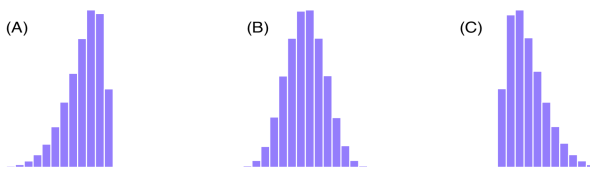

Mode

2. A researcher wants to use a measure of central tendency to summarize the distribution of depression scores (a numeric or continuous variable). To do so, she orders the scores from low to high and finds the middle score. Which measure of central tendency did she use?
Median

3. A researcher studying psychological characteristics influenced by chronic pain measures pain severity on a 1 to 10 scale. Consider the following subsample of scores. 2,3,1,7,7. What is the median of the sample?
3

4. The following graph shows three distributions of test scores. For which distribution is the mean the most appropriate measure of central tendency?

(B)

5. A study of chronic patients asks respondents to rate their level of consistent pain. What is the mode of the distribution?

Frequencies of PainSeverity

Levels	Counts	% of Total
None or Little Pain	58	19.3%
Moderate Pain	138	46.0%
Severe Pain	104	34.7%

Moderate pain

6. A researcher wants to use a measure of central tendency to summarize the ethnicity breakdown of his sample. Which measures of central tendency can be computed for this variable? Note that the question is NOT asking which one you should use, but which one you could use. Check all that apply.
Mode

7. Suppose the following graphs display midterm test scores from three different classes. In which graph is the numeric value of the mean larger (i.e., further to the right on the horizontal axis) than the numeric value of the median?

(A)   (B)   (C)
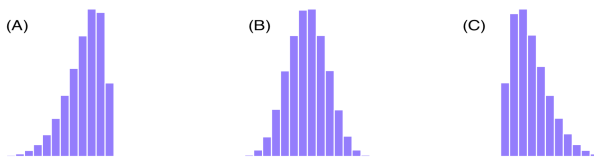
(C) ✓
(A) I got C for this one because I thought a positively skewed distribution would cause the mean to be larger (further to the right) bc of the outliers? Is this incorrect
- Yea i also think C is the correct answer

8. The following text appeared in a news story on CNBC's website: "According to a report from the Economic Policy Institute, the mean retirement savings of a family between 44 and 49 years old is $81,347. But that number doesn't tell the whole story. Since so many families have zero savings and since super-savers can pull up the average, the median savings, or those at the 50th percentile, may be a better gauge. The median for families between 44 and 49 is only $6,200." Based on this information, which of the following distribution shapes corresponds to the distribution of retirement savings?

(A)   (B)   (C)
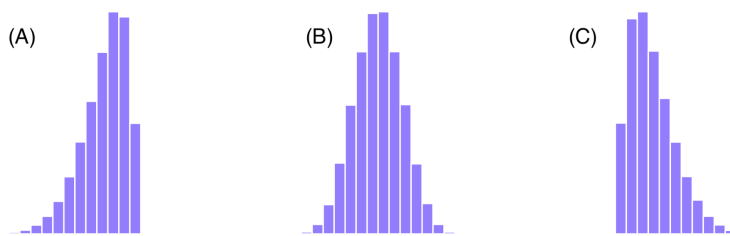
(C)

9. A researcher studying psychological characteristics influenced by chronic pain measures pain severity on a 1 to 10 scale. Consider the following subsample of scores. 2,3,1,7,7. What is the mean of the sample?
4

10. Suppose the following graphs display midterm test scores from three different classes. In which graph are the numeric values of the mean and median about the same (i.e., at about the same location on the horizontal axis)?

(A)   (B)   (C)
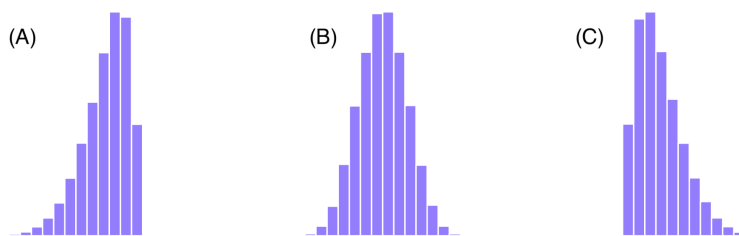
(B)

11. A researcher asks respondents to report their age in a series of bins or categories: "18-24", "25-29", "30-34", "35-39", and "40 or above". Which measures of central tendency can be computed for this variable? Note that the question is NOT asking which one you should use, but which one you could use. Check all that apply.

A researcher studying post-partum depression wants to summarize the scores with a measure of central tendency. Which of the following best describes her goal?
- **Describe the typical score**

Content Quiz #3:

1. Considering the frequency distribution below, what response category corresponds to the mode?

| Response	Frequency	Percent	Cumulative Percent
Strongly agree	250	50	100
Agree	150	30	50
Disagree	75	15	20
Strongly disagree	25	5	5

   a. Strongly agree
2. In a normal distribution, which of the following is true?
   a. The mean is the same as the median
3. Which measure of central tendency can take on more than one value?
   a. Mode
4. A researcher asks respondents to report their income in a series of bins or categories: "Less than 20K", 20K-39K", "40K-59K", "60K-79K", "80K-99K", and "100K or higher". Which measures of central tendency can be computed for this variable? Note that the question is NOT asking which one you should use, but which one you could use. Check all that apply.
   a. Mode and Median
5. Consider the following set of depression scores: 11, 12, 12, 13, 15, 16, 17, 19, 22, 23. Suppose that the value of 23 changed to a 30. Which measure of central tendency would not be affected by this change? Check all that apply.
   a. median
   b. Mode
6. Consider the following set of depression scores: 1, 2, 2, 3, 5, 6, 7, 9, 12, 13. What is the sample mean?
   a. 6
7. Consider the following set of depression scores: 11, 12, 12, 13, 15, 16, 17, 19, 22, 23. What is the sample median?
   a. 15.5
   b. I had the same question but without 23, at the end, so answer is 15
8. A researcher wants to use a measure of central tendency to summarize the ethnicity breakdown of his sample. Which measures of central tendency can be computed for this variable? Note that the question is NOT asking which one you should use, but which one you could use. Check all that apply.
   a. Mode
9. A researcher is examining the age distribution for her sample and finds the following values: mean = 29.8, median = 29.6, mode = 30. Which of the following is most likely true about the shape of the age distribution?
   a. Normally distributed

10. A researcher reports the mean income for her sample. Which level of measurement likely describes the income variable?

    a. interval/ratio/approximately continuous

11. Consider the following set of depression scores: 11, 12, 12, 13, 15, 16, 17, 19, 22, 23. What is the sample mode?

    a. 12

12. The following text appeared in a news story on CNBC's website: "According to a report from the Economic Policy Institute, the mean retirement savings of a family between 44 and 49 years old is $81,347. But that number doesn't tell the whole story. Since so many families have zero savings and since super-savers can pull up the average, the median savings, or those at the
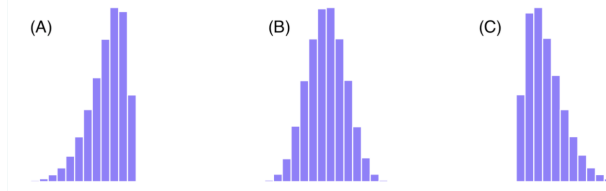
50th percentile, may be a better gauge. The median for families between 44 and 49 is only $6,200." Based on this information, which of the following distribution shapes corresponds to the distribution of retirement savings?

    a. C

13. In 2011, the mean salary of an NFL player was $1.9 million, and the median salary was $770,000. Which of the following is most likely true about distribution of salaries?

    a. Positively (right) skewed

14. A researcher studying post-partum depression wants to report the central tendency for her depression scale score. In most situations, which measure of central tendency is most desirable?

    a. Mean

15. Consider the following set of depression scores: 1, 2, 2, 3, 5, 6, 7, 9, 12, 13. What is the sample median?

    a. 5.5

16. Consider the following set of depression scores: 1, 2, 2, 3, 5, 6, 7, 9, 12, 13. What is the sample mode?

    a. 2

17. A researcher studying aggressive behavior in children decides to report the median number of aggressive behaviors that children engage in during the observation period. Which of the following most likely describes the reason for this choice?

    a. The distribution of scores is skewed

18. A researcher wants to use a measure of central tendency to summarize the distribution of depression scores (a numeric or continuous variable). To do so, she orders the scores from low to high and finds the middle score. Which measure of central tendency did she use?
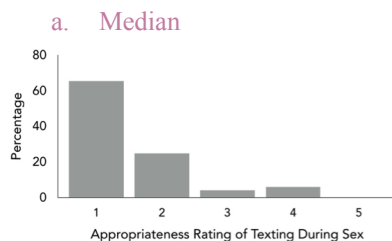
    a. Median

19. Consider the bar graph below, which depicts the distribution of appropriateness ratings for texting during sex. Which of the following is most likely true?

    a. The mean is higher than the median

1. Participants were randomly assigned to participate in either a treatment or a control condition. The TREATMENT variable (0 = control, 1 = treatment) in the data set codes which conditions each person was assigned to. What level of measurement describes the TREATMENT variable?
   a. A
2. The AGE variable in the data set reflects age **to the nearest year.** What level of measurement describes the AGE variable?
   a. C
3. Consider the first depression item, DEPRESS1, which ranges from 1 (Strongly Disagree) to 4 (Strongly Agree). What level of measurement describes the DEPRESS1 variable?
   a. B
4. What is the DEPRESS scale score value for the individual in the first row of the data set? You will need to click the DATA tab to see the new variable you created.
   a. 3.857
5. Suppose that a respondent obtained the lowest possible value of the DEPRESS scale score. What would this person's score be?
   a. 1
6. Suppose that a respondent obtained the highest possible value of the DEPRESS scale score. What would this person's score be?
   a. 4
7. What level of measurement describes the DEPRESS scale score variable?
   a. C
8. What is the INTERFERE scale score value for the individual in the first row of the data set? You will need to click the DATA tab to see the new variable you created.
   a. 5.167
9. Consider a participant with an INTERFERE scale score of 1.5. Based on this person's location in the distribution, how would you characterize the extent to which pain interferes with this person's day-to-day activities?
   a. A
10. Use jamova to obtain a histogram with kernel density plot for the DEPRESS scale score you created. See the lecture slides or the lab video for instructions if needed. Based on the distribution, what is your conclusion about the overall magnitude of depressive symptoms in this sample?
    a. C
11. Considering the graph, how would you characterize the shape of the depression distribution?
    a. A
12. Use jamovi to obtain a kernel density plot for the INTERFERE scale scores. Based on the graph, which of the following best describes the distribution shape of this variable?
    a. B
13. Given the shape of the distribution, which of the following interpretations is correct regarding the overall degree to which pain interferes with daily life in this sample?
    a. A
14. Considering the two kernel density plots and paying close attention to the center of the distributions (you can also look at the means, which jamovi prints in a table above the graphs), which of the following statements is best supported by the data? Note that lower scores on INTERFERE are better because a lower score indicates that participants are experiencing fewer pain-related barriers in their day to day lives.
    a. B
15. Considering the two kernel density plots and paying close attention to the center of the distributions, which of the following statements is best supported by the data

a.   C

4. Variability

Class Prep Quiz #4

1.   When computing the standard deviation or variance from a sample of data, we compute each person's distance from the mean, then we average those distances by dividing by the number of scores MINUS ONE. What is the reason for this operation instead of just averaging by dividing by the number of scores?
     The resulting standard deviation or variance from the data is a better guess about the spread in the entire population

2.   A researcher studying psychological characteristics influenced by chronic pain measures pain severity on a 1 to 10 scale. Consider the following subsample of scores. 2,3,1,7,7. What is the standard deviation of the sample?
     2.53 ?? or 2.83 ?? ← standard deviation for samples means you have to do n-l, so I think it would be 2.83
     I got 2.53
     I got 2.828
     2.83; it's a subsample of sample, if it's a sample u have to do n-1

3.   Measures of spread or dispersion include the standard deviation and range. What is the primary purpose of these statistics? What information do they convey?
     The degree to which scores in a distribution are similar or different

4.   A researcher studying psychological characteristics influenced by chronic pain measures pain severity on a 1 to 10 scale. Consider the following subsample of scores. 2,3,1,7,7. What is the range of the sample?
     6

5.   Consider a sample of midterm test scores. In which of the following situations will the standard deviation be the smallest?
     Every student has exactly the same score

6.   Consider the following two distributions. Which distribution has the largest standard deviation?
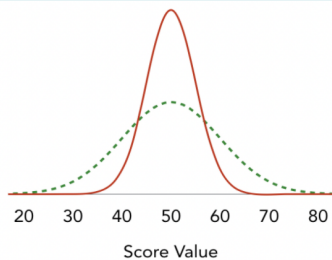
The green distribution with the dashed line

7.   The average math score in a calculus class is 77 with a standard deviation of 10. Which of the following is true?
     The average distance between a student's score and the mean is 10 points

8.   A researcher studying psychological characteristics influenced by chronic pain measures pain severity on a 1 to 10 scale. The researcher summarizes the results as follows: "On average, individual pain ratings differed from the sample mean by 2 points." What statistic is the researcher using to summarize the data?
     The standard deviation

9.   A social psychologist studying social norms related to texting collects data on the number of texts sent in a typical day. A histogram of the responses is shown below. The distribution has a mean of 16 and standard deviation of 25, but there is one extreme (outlier) score of 200 texts. What would happen to the standard

deviation if the researcher removed this one score from the data and recomputed the standard deviation without it?

The standard deviation would be smaller (std. dev. < 25)

10. Consider a sample of midterm test scores. Suppose the variance of the scores is 81. What is the value of the standard deviation?

9

Content Quiz #4

1. A sample of depression scores has a mean of 20 and a variance of 25. Which of the following interpretations is true about the standard deviation?
   Standard deviation = 5

2. Consider the following set of scores: 1, 2, 2, 7, 8. What is the sample standard deviation (round to two decimals)?
   3.24

3. A sample of depression scores has a mean of 20 and a standard deviation of 5. Which of the following interpretations is definitely true?
   The average difference between a score and the mean is 5 points

4. College students are asked to rate the appropriateness of texting during class and texting while on a date. The standard deviation for texting in class is 1.25, and the standard deviation for texting on a date is 1.00. Which question elicited responses that were most similar?
   Texting while on a date

5. In a survey of hospital patients, the average age is 50 with a standard deviation of 10. Which of the following ages would definitely be considered an outlier (i.e., an uncommon score located in the tail of the distribution)?
   20

6. IQ scores are normally distributed with a mean of 100 and a variance of 225. Which of the following is true?
   The average difference between an individual IQ and the mean is 225 points

7. In a survey of hospital patients, the average age is 50 with a standard deviation of 10. What is the age of a respondent who is zero standard deviations from the mean?
   50

8. In a survey of hospital patients, the average age is 50 with a standard deviation of 10. What is the age of a respondent who is two standard deviations below the mean?
   30

9. A researcher studying post-partum depression wants to summarize the scores with a measure of variability. Which of the following best describes her goal?
   Describe the degree to which scores differ

10. A sample of depression scores has a mean of 20 and a standard deviation of 5. The sample standard deviation was computed using a formula with N-1 (the degrees of freedom) in the denominator. Which of the following is true regarding the N-1 value in the standard deviation formula?
    The standard deviation = 5 value is a more accurate approximation of the standard deviation of depression scores in the full population

11. A researcher reports that the average squared distance from the scores to the mean is 9 points. Which measure of variability is he reporting?
    Variance

12. In which of the following situations will the standard deviation be the smallest?
    All scores are identical

13. A researcher reports that the average squared distance from the scores to the mean is 9 points. Which of the following is true about the standard deviation?
    Standard deviation < 9
14. A researcher studying depression examines scores for males and females. Males had a mean depression score of 18 with a standard deviation of 6, while females had a mean depression score of 20 with a standard deviation of 8. In which group were depression scores most different?
    Females
15. A researcher is summarizing a sample of ages and reports that the average difference between a respondent's age and the sample mean is 8 years. Which measure of variability is she reporting?
    Standard deviation
16. A sample of depression scores has a mean of 20 and a standard deviation of 5. Which of the following scores represents someone with unusually low depression (i.e., an outlier)?
    10
17. Consider two samples of SAT scores. Sample 1 has 490, 495, 500, 506, and 510. Sample 2 has 490, 492, 492, 497, and 500. Which sample has a smaller standard deviation (do not do the calculations)?
    Sample 2
18. Consider two samples of SAT scores. Sample 1 has 490, 492, 492, 497, and 500. Sample 2 has 490, 495, 500, 506, and 510. Which sample has a larger standard deviation (do not do the calculations)?
    Sample 2
19. The following sample of SAT scores has a standard deviation of 8.07: 490, 495, 500, 506, and 510. What would happen to the standard deviation value if the score of 510 changed to 550 (i.e., the scores were 490, 495, 500, 506, and 550)?
    Standard deviation > 8.07
20. A researcher studying psychological characteristics influenced by chronic pain measures pain severity on a 1 to 10 scale. The researcher summarizes the results as follows: "On average, individual pain ratings differed from the sample mean by 2 points." What statistic is the researcher using to summarize the data?
    The standard deviation
21. A sample of depression scores has a mean of 20 and a variance of 25. Which of the following interpretations is definitely true?
    The average squared difference between a score and the mean is 25 points
22. The population standard deviation formula has N in the denominator, whereas the sample standard deviation formula has N-1 in the denominator. What is the rationale for using N-1 rather than N when computing the average distance from the center of the data?
    Dividing by N-1 would make the sample standard deviation a more accurate estimate of the population standard deviation
23. A researcher wants to report a measure of variability for gender. Which measures of variability can be computed for this variable? Check all that apply.
    Measures of variability cannot be computed for gender
24. In a survey of hospital patients, the average age is 50 with a standard deviation of 10. What is the age of a respondent who is one standard deviation above the mean?
    60
25. A researcher calculates the standard deviation of age and obtains a value of zero. Which of the following is true?
    All ages were identical
26. In a survey of hospital patients, the average age is 50 with a standard deviation of 10. Which of the following ages would definitely be considered an outlier (ie., an uncommon score located in the fail of the distribution)?
    20

27. A researcher studying depression examines scores for males and females. Males had a mean of depression score of 18 with a standard deviation of 6. While females had a mean depression score of 20 with a standard deviation of 8. In which group were depression scores most different?

    Females

28. A researcher is summarizing a sample of ages and reports that the average difference between a respondent's age and the sample mean is 8 years. Which measure of variability is she reporting?

    Standard Deviation

29. When computing the standard deviation or variance from a sample of data, we compute each person's distance from the mean, then we average those distances by dividing by the number of scores MINUS ONE. What is the reason for this operation instead of just averaging by dividing by the number of scores?

    The resulting standard deviation or variance from the data is a better guess about the spread in the entire population

30. A sample of males has a depression mean of 20 and a standard deviation of 5, whereas a sample of females has a depression mean of 19 and a standard deviation of 4. In which group are depression scores most similar to one another?

    Females

31. In which of the following situations will the standard deviation be the largest?

    All scores are very different from the mean

32. IQ scores are normally distributed with a mean of 100 and a standard deviation of 15. Which of the following is true?

    The average difference between an individual IQ and the mean is 15 points

33. A sample of depression scores has a mean of 20 and a standard deviation of 5. Which of the following scores represents someone with unusually high depression (i.e., an outlier)?

    30

34. The population standard deviation formula has N in the denominator, whereas the sample standard deviation formula has N-1 (degrees of freedom) in the denominator. What is the rationale for using N-1 rather than N when computing the average distance from the center of the data?

    N-1 makes the sample standard deviation closer to the standard deviation of the full population of scores

35. Consider the following set of scores: 11, 12, 12, 17, 18. What is the sample standard deviation (round to two decimals)?

    3.24

5. Standardizing Variables

Class Prep Quiz #5

1. After converting depression scores to the z-score metric, a study participant's depression score is $z = -2.5$. Which of the following is definitely true about this person's depression?

    The individual is far below the mean

2. Two researchers are studying a new treatment for depression using a design involving a treatment and a control group. Researcher A finds a Cohen's d effect size of .20, and Researcher B finds a Cohen's d effect size of .50. In which study was the treatment group's mean most different from the control group's mean?

    Researcher B

3. SAT scores have a mean of 500 and standard deviation of 200. ACT scores have a mean of 20 and a standard deviation of 5. Suppose that student A has an SAT score of 600, and student B has an ACT score of 25. Using z-scores, which student had a better college entrance exam score?

4. To better understand scores, a researcher converts his sample of depression scores to z-values (high depression scores mean more depressive symptoms, low scores mean less depression). Which of the following z-scores most likely corresponds to a clinical depression diagnosis?
   $z = 2.0$

5. A researcher studying psychological characteristics influenced by chronic pain measures pain severity on a 1 to 10 scale. The sample mean was 5 and the standard deviation was 2. What is the z-score for someone with a pain rating of 6?
   0.5

6. IQ scores are normally distributed. What z-score would somebody need to obtain in order to score exactly at the mean of the distribution?
   $0 \leftarrow$ z score distributions mean the mean becomes standardized to 0

7. Suppose that your IQ corresponds to a z-score of +1.0. What is the interpretation of this value?
   Your IQ falls above the mean by a distance that equals the standard deviation

8. A researcher is studying a new treatment for depression using a design involving a treatment and a control group. She reports that the Cohen's d effect size is .50. Which of the following interpretations is consistent with this finding?
   The treatment and control group averages differ by an amount equal to one half of a standard deviation unit

9. A test preparation company reports that, after taking their GRE preparation class, student scores increased by a magnitude that is equivalent to .25 standard deviation units relative to a control group (i.e., Cohen's d = .25). Which of the following sets of descriptive statistics produced this effect size?
   I put GRE class mean = 525 (SD = 100), control mean = 500 (SD = 100) ← me too, same ! ← same here ← same!
   GRE class mean = 475 (SD = 100), control mean = 500 (SD = 100)

10. A researcher is studying a new treatment for depression using a design involving a treatment and a control group. The treatment group depression mean is 20 and the control group's mean is 23 (lower scores are better). The standard deviation is 6. What is the numeric value of Cohen's d effect size?
    0.5
    I got -0.5 because I did 20-23
    -- I also wrote -0.5, wrote -0.5 too!
    Is 'numeric value' stating only positive values? Or would it be -.5

    Google says "Cohen's d is a measure of the magnitude of effect and cannot be negative"

    Definitions of numerical value. noun. a real number regardless of its sign. synonyms: absolute value. numerical value - Dictionary Definition : Vocabulary.comhttps://www.vocabulary.com
     So it would be .5? I think so but do 2 attempts to be sure :)

Content Quiz #5

1. After converting SAT scores to the z-score metric, a college applicant's score is z = 2.25. Which of the following is true about this person's test performance?
   - The individual is far above the mean

2. A researcher uses an established behavioral observation measure to collect antisocial behavior scores for a group of elementary school children. The sample mean is 20 and the standard deviation is 3. On the z-score metric, what is the mean behavior score?

   - 0

3. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. He assigns 20 subjects to the treatment condition and 20 to the control, and he measures pain intensity on a 1-10 scale (low pain ratings are good, high pain ratings are bad). Which of the following sets of descriptive statistics would produce the largest standardized mean difference (i.e., Cohen's d effect size)?

- Treatment mean = 3 (SD = 2), control mean = 6 (SD = 2)

4. In a positively (right) skewed distribution, a student obtains a test score below the mean. What do we know about this student's z-score?

- The z-score is negative

5. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. He assigns 20 subjects to the treatment condition and 20 to the control, and he measures pain intensity on a 1-10 scale (low pain ratings are good, high pain ratings are bad). At the end of the study the researcher finds an improvement, the standardized mean difference (i.e., Cohen's d) effect size of which is 1.00. Which of the following sets of descriptive statistics produced this effect size?

- Treatment mean = 4 (SD = 2), control mean = 6 (SD = 2)

6. Two researchers are studying a new treatment for depression using a design involving a treatment and a control group. Researcher A finds a standardized mean difference effect size (i.e., Cohen's d) of d = des.20, and Researcher B finds an effect size of d = .50. Which study implemented the most effective treatment?

- Researcher B

7. A researcher uses an established behavioral observation measure to collect antisocial behavior scores for a group of elementary school children. The sample mean is 20 and the standard deviation is 3. On the z-score metric, what is the standard deviation of the behavior scores?

- 1

8. A psychologist is interested in determining what type of intervention most improves attitudes toward vaccination. The "vaccination risk" condition has participants view pictures of kids with preventable diseases. Comparing the pretest and posttest means, the researchers found a standardized mean difference (i.e., Cohen's d effect size) of d = .85. Based on widely accepted rules of thumb, how would you characterize the magnitude of this treatment effect.

- Large

9. A test preparation company reports that, after taking their GRE preparation class, student scores increased by a magnitude that is equivalent to .25 standard deviation units (i.e., the Cohen's d effect was was d = .25). Which of the following sets of descriptive statistics produced this effect size?

- The test prep class produced a relatively small improvement in scores

10. Grades on a chemistry midterm have a mean of 70 and a standard deviation of 10. What is the z-score for an exam grade of 81 (report to one decimal)?

- 1.1

11. Grades on a psychology midterm have a mean of 80 and a standard deviation of 5. On the z-score metric, the student's score was 0. Which interpretation is correct regarding the student's performance?

c.The student scored about average

12. Researchers are interested in determining whether a smoking cessation intervention reduces smoking frequency. They assign 50 subjects to the treatment condition and 50 to the control, and measure smoking frequency on a 1-5 scale. At the end of the study the researcher finds that the treatment subjects improved, with a standardized mean difference (i.e., Cohen's d) effect size of d = .51. How would you characterize the impact of the new class on smoking behavior?

- The intervention had a moderate positive impact on smoking behavior

13. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. He assigns 20 subjects to the treatment condition and 20 to the control, and he measures pain intensity on a 1-10 scale. Which of the following sets of descriptive statistics would produce the smallest standardized mean difference (i.e., Cohen's d effect size)?

- Treatment mean = 5 (SD = 2), control mean = 6 (SD = 2)

14. A university runs a new class designed to improve the Freshman experience. The hypothesis is that participation in the new class will improve attitudes toward school. To evaluate the hypothesis, the research team assigns a sample of 2000 incoming Freshman to the intervention class and 2000 to a standard Freshman class load (control). The researchers find a standardized mean difference (i.e., Cohen's d) effect size of d = .07. How would you characterize the impact of the new class on attitudes?
- The class had little to no impact on attitudes

15. A test preparation company reports that, after taking their GRE preparation class, student scores increased by a magnitude that is equivalent to .25 standard deviation units (i.e., the Cohen's d effect was was d = .25). Which of the following is true?
- The test prep class produced a relatively small improvement in scores

16. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. He assigns 20 subjects to the treatment condition and 20 to the control. At the end of the study the researcher finds a standardized mean difference (i.e., Cohen's d) effect size of d = .30. Based on widely accepted rules of thumb, how would you characterize the magnitude of this treatment effect.
- Small

17. A psychologist is interested in determining what type of intervention most improves attitudes toward vaccination. The "research literature" condition has participants read summaries of published studies, all of which show no linkage between vaccinations and autism. The "vaccination risk" condition has participants view pictures of kids with preventable diseases. Finally, a control condition does not participate in any intervention but simply reports their attitudes. Both interventions improve attitudes, such that the "research literature" condition has a standardized mean difference (i.e., Cohen's d effect size) of d = .20, and the "vaccination risk" condition has a standardized mean difference of d = .30. Which intervention was most effective at improving attitudes?
- Illustrating vaccination risks was more effective

18. Grades on a psychology midterm have a mean of 80 and a standard deviation of 5. What is the z-score for an exam grade of 75 (report to one decimal)?
- -1.0

19. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. He assigns 20 subjects to the treatment condition and 20 to the control, and he measures pain intensity on a 1-10 scale. At the end of the study the researcher finds an improvement, the standardized mean difference effect size (i.e., Cohen's d) of which is d = .25. Published studies show that the group therapy tends to produce standardized effect sizes of about d = .35. Which of the following is true about the two treatment modalities?
- Group therapy is slightly more effective

20. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. He assigns 20 subjects to the treatment condition and 20 to the control, and he measures pain intensity on a 1-10 scale. At the end of the study the researcher finds an improvement, the standardized mean difference (i.e., Cohen's d) effect size of which is d = .30. How would you characterize the impact of the new program on pain?
- The program had a small positive impact on pain levels

21. Grades on a chemistry midterm have a mean of 70 and a standard deviation of 10, and grades on a psychology midterm have a mean of 80 and a standard deviation of 5. A student obtained a score of 80 on both exams. Use z-scores to determine the test on which the student performed better.
- chemistry

22. Grades on a chemistry midterm have a mean of 70 and a standard deviation of 10. On the z-score metric, the student's score was +1.5. Which interpretation is correct regarding the student's performance?Grades on a psychology midterm have a mean of 80 and a standard deviation of 5. On the z-score metric, the student's score was -2.0. Which interpretation is correct regarding the student's performance?

- The student scored significantly better than most of his peers

23. Researchers are interested in determining whether a smoking cessation intervention reduces smoking frequency. They assign 50 subjects to the treatment condition and 50 to the control, and measure smoking frequency on a 1-5 scale. At the end of the study the researcher finds that the treatment subjects improved, with a standardized mean difference effect size of .51. Which of the following interpretations is correct?
- The mean difference between the treatment and control group is equivalent to about a half a standard deviation

8. A p
- S

8. A p
- S

Two researchers are studying a new treatment for depression using a design involving a treatment and a control group. Researcher A finds a standardized mean difference effect size (i.e., Cohen's d) of d = .20, and Researcher B finds an effect size of d = .50. Which study implemented the most effective treatment?

- Researcher B

To better understand scores, a researcher converts his sample of depression scores to z-values. Which of the following z-scores most likely corresponds to a clinical depression diagnosis? The depression scores are scaled such that higher values indicate more depressive symptoms.

- z=2.0

After converting depression scores to the z-score metric, a study participant's score takes on a positive value. Which of the following is definitely true about this person's depression score?

- The individual is above the mean

Grades on a psychology midterm have a mean of 80 and a standard deviation of 5. On the z-score metric, the student's score was -2.0. Which interpretation is correct regarding the student's performance?

- The student scored significantly worse than most of his peers

Grades on a chemistry midterm have a mean of 70 and a standard deviation of 10, and grades on a psychology midterm have a mean of 80 and a standard deviation of 5. A student obtained a score of 80 on both exams. Use z-scores to determine the test on which the student performed better.

- Chemistry

Assignment #2

1. Enter the value of the sample mean, exactly as it appears on the printout. _____

    4.00

2. Which of the following interpretations is most correct regarding the standard deviation?

    The average difference between an individual's score and the mean is 1.33 points

3. In which group are pain interference scores most similar?

    Females

4. For females, the mode = 5, median = 4.50, and mean = 4.12. Why do the three measures of spread differ so much?

    The distribution is negatively skewed, with lower scores pulling down the mean

5. To facilitate interpretation of the gender effect, use jamovi to obtain the standardized mean difference (i.e., Cohen's d effect size) effect size between males and females. There is a short video on CCLE that illustrates this process, and a jamovi screenshot also appears in the lecture slides. Enter the value of the standardized mean difference exactly as it appears on the screen. _____

    -0.138← i got this me too same same

6. Considering the male and female means and the standardized mean difference, which of the following interpretations is correct?

    Males have lower levels of pain interference, and the gender difference is less than small in magnitude

7. The researchers want to know whether the intervention condition is a variable that influences pain interference score (i.e., Do treatment and control cases experience different levels of pain interference) by how much do the centers of the two distributions differ? _____

    0.77

8. The Cohen's d (standardized mean difference) effect size value is .60. Which of the following interpretations is correct about this value?

    The treatment and control group means differ by 0.60 standard deviation units

9. Considering the raw mean value and the standardized mean difference, which of the following interpretations is correct?

    The treatment group has lower levels of pain interference, and the mean difference is moderate in magnitude

10. Given the raw mean difference and effect size value, what is your conclusion about the effectiveness of the online intervention?

    The treatment was beneficial, participants experienced a reduction in pain interference levels

6. Sampling Distributions and Error

Class Prep Quiz #6
1. Each random sample we could possibly take from a population will give a slightly different estimate of the mean. Which of the following is true about the distribution of sample means from MANY different samples?
    A histogram of the means from many different random samples will look like a bell (normal) curve
2. Which of the following best describes the sampling distribution of the mean?
    A histogram of the sample means from MANY different random samples of participants
3. The expected variability or standard deviation of the means from MANY different random samples is known as what quantity?
    Standard error
4. Suppose a research team at UCLA wants to estimate the average depression level in the general population. To do so, they recruit a random sample of 300 participants from throughout the city. What is the estimate for this scenario?
    The mean depression score for the 300 participants.
5. Which of the following statements is true concerning sample means drawn from the same population.
    Each sample we could potentially collect from a population will give us a different mean.
6. What role does a sampling distribution play in statistics?
    The shape of the sampling distribution (e.g., its skewness) helps us to understand how precise our sample mean is
    I said it would be 'draw conclusions about the mean of a larger population'; the shape of the sampling distribution would always be normal as long as there were enough samples, as large samples == more normal;
7. Psychology is very concerned with replicability of research findings (i.e., studies performed by different researchers and with different samples arrive at similar conclusions). Suppose a research team at UCLA administers the Beck Depression Inventory to a community sample gets 20 as an estimate of the mean. Researchers at UC Merced want to replicate this finding, so they also collect data from a community sample. Which of the following statements is true about the means of the two samples?
    The mean depression scores from the two samples will likely differ

8. Which of the following best describes the link between an estimate and the population.

    We collect a sample and compute a statistic that estimates some characteristic of the population (e.g., the mean)
9. A research team at UCLA wants to estimate the average depression level in the general population. The team administers the Beck Depression Inventory to a community sample of 300 participants and gets a mean of 20. This sample mean is unlikely to be exactly the same as the true mean in the full population. Which of the following terms describes the difference between the sample mean and the true mean of the population?
    Sampling variability
10. Suppose a research team at UCLA wants to estimate the average depression level in the general population. To do so, they recruit a random sample of 300 participants from throughout the city. What is the population parameter for this scenario?
    The mean depression score of all adults in the United States

Content Quiz #6

1. Across the entire Presidential election season, polling companies conduct hundreds of political opinion polls. Suppose that we obtained the percentage of Hillary Clinton voters from 100 different polls and created a histogram (sampling distribution) of the 100 estimates. Which of the following best describes the likely shape of this distribution?
    a. Normal distribution
2. A psychologist studying postpartum depression administered a questionnaire to a sample of 100 new mothers and obtained a mean depression score of 22 with a standard deviation of 4. The standard error of the mean is .40. All other things being equal, what will happen to the standard error if the standard deviation is increased to 8?
    a. The standard error will increase
3. Suppose that researchers from 10 different universities each computed the mean depression score from a sample of 100 participants. The standard deviation of the estimates across the 10 samples is what quantity?
    a. Standard error
4. A polling firm wants to know the percentage of registered Democratic voters that approve of Donald Trump's performance as President. What would the polling firm need to do to obtain an estimate?
    a. Survey a randomly selected group of 1000 registered Democrats
5. A psychologist studying postpartum depression administered a questionnaire to a sample of 100 new mothers and obtained a mean depression score of 22 with a standard deviation of 4. The standard error of the mean is .40. Which of the following interpretations is correct for the standard error?
    a. A sample mean based on $N = 100$ should differ from the population mean by .40, on average
6. Across the entire Presidential election season, polling companies conduct hundreds of political opinion polls. Suppose that we obtained the percentage of Hillary Clinton voters from 100 different polls and created a histogram (sampling distribution) of the 100 estimates. Which of the following statements is true?
    a. The true percentage of Clinton voters in the population is near the center of the histogram
7. Each random sample we could possibly take from a population will give a slightly different estimate of the 22e standard error of 0.50 mean. Which of the following is true about the distribution of sample means from MANY different samples?
    a. A histogram of the means from many different random samples will look like a bell (normal) curve
8. What role does a sampling distribution play in statistics?

a. It allows us to use a sample mean to draw conclusions about the mean of a much larger population

9. Across the entire Presidential election season, three different polling companies conducted dozens of public opinion polls that estimate the percentage of Hillary Clinton voters in the population. To keep track of their results, the polling firms graph the estimates obtained across many random samples, the results of which are shown below. These sampling distribution graphs illustrate which of the following concepts?
   a. Sampling distribution

10. Across the entire Presidential election season, three different polling companies conducted dozens of public opinion polls that estimate the percentage of Hillary Clinton voters in the population. To keep track of their results, the polling firms graph the estimates obtained across many random samples, and the resulting sampling distributions are shown below. Which of the polling firms obtained estimates with the most sampling error or sampling variability?
    a. Polling firm A's estimates exhibited the most sampling error or sampling variability

11. A social psychologist is trying to estimate the average tipping behavior in the population of restaurant patrons. She collects data from a sample of 20 participants and finds that the mean is 24% with a standard error of .50. Which of the following is true about the standard error value?
    a. The sample mean based on N = 20 should differ from the true population mean tip by .50, on average

12. A psychologist studying post-partum depression administered a questionnaire to 200 new mothers and obtained a mean depression score of 22. The 22 value represents what quantity?
    a. Estimate

13. A poll of 772 registered voters found that 42% of respondents approve of Donald Trump's performance as President. The 42% value represents what quantity?
    a. Estimate

14. A psychologist studying post-partum depression administered a questionnaire to a sample of 100 new mothers and obtained a mean depression score of 22 with a standard deviation of 4. The standard error of the mean is .40. All other things being equal, what will happen to the standard error if the sample size was increased to 200?
    a. The standard error will decrease

15. A polling firm wants to know the percentage of registered Democratic voters that approve of Donald Trump's performance as President. What would the polling firm need to do to obtain the parameter?
    a. Survey a randomly selected group of 1000 registered voters

16. A psychologist studying post-partum depression administered a questionnaire to 200 new mothers and obtained a mean depression score of 22. What would the researcher need to do to obtain the parameter?
    a. Administer the depression survey to all new mothers in the US

17. Suppose a research team at UCLA wants to estimate the average depression level in the general population. To do so, they recruit a random sample of 300 participants from throughout the city. What is the estimate for this scenario?
    a. The mean depression score the 300 participants

18. A poll of 772 registered voters found that 42% of respondents approve of Donald Trump's performance as President. The 772 respondents represent what quantity?
    a. Sample

19. A psychologist studying aggressive behavior in a sample of 30 children finds the average number of aggressive acts during a play period is 6 with a standard error of .50. Which of the following interpretations correct for the standard error?
    a. The sample mean based on 30 children should differ from the true population mean by 0.50, on average

20. Across the entire Presidential election season, three different polling companies conducted dozens of public opinion polls that estimate the percentage of Hillary Clinton voters in the population. To keep track of their results, the polling firms graphed the estimates obtained across many random samples, and the sampling

distributions are shown below. Which of the polling firms obtained estimates with the least amount of sampling error or sampling variability?

  a. Polling firm C's estimates exhibited the least sampling error or sampling variability

21. A psychologist studying post-partum depression administered a questionnaire to 200 new mothers and obtained a mean depression score of 22. Suppose that the researcher could somehow administer the questionnaire to all new mothers in the US, finding a mean of 24. The 24 value represents what quantity?

a. parameter

7. Confidence Intervals

Class Prep Quiz #7

1. A researcher studying chronic pain wants to estimate the mean pain level in a population of persons with fibromyalgia. Using a rating scale of 0 (no pain at all) to 10 (worst pain imaginable), a researcher estimates the mean at 6. Which of the following will produce the largest confidence interval (i.e., an interval with the widest range)?

    A sample size of 100

2. Two industrial organizational psychologists use a 7-point rating scale to estimate the mean job satisfaction in a particular industry. The first research's mean estimate is 4, and the 95% confidence interval ranges from 3.89 to 4.09. The second research's mean estimate is 4.1, and the 95% confidence interval ranges from 3.79 to 4.39. Which researcher's estimate is likely more accurate?

    I put the answer with the interval 3.89 to 4.09 because they're both 95% confident, so the smaller interval range be more accurate because that means their estimate is most likely closer to the true mean

    On March 22, 2020, the Gallup Presidential opinion poll reports that 49% of registered voters approve of the job that Trump is doing as President, with a 95% confidence interval ranging from 46% to 51%. Which of the following is true?

    There is a 95% chance that the interval from 46% and 51% contains the true favorability ratings in the full populations of likely voters

3. Industrial organizational psychologists use a 7-point rating scale to estimate the mean job satisfaction in a particular industry. The mean estimate is 4, and the 95% confidence interval ranges from 3.89 to 4.09. Suppose the researcher instead wanted to compute a 99% confidence interval? Which of the following would be true?

    The width of the confidence interval would increase (i.e., cover a range greater than .20)

4. Two industrial organizational psychologists use a 7-point rating scale to estimate the mean job satisfaction in a particular industry. The first research's mean estimate is 4, and the 95% confidence interval ranges from 3.89 to 4.09. The second research's mean estimate is 4.1, and the 95% confidence interval ranges from 3.89 to 4.29. Which researcher likely had a smaller sample size?

    The researcher with an interval from 3.89 to 4.29

5. A researcher studying chronic pain wants to estimate the mean pain level in a population of persons with fibromyalgia. Using a rating scale of 0 (no pain at all) to 10 (worst pain imaginable), a researcher estimates the mean at 6. Which of the following will produce the largest confidence interval (i.e., an interval with the widest range)?
    A standard deviation of 2

6. A researcher studying chronic pain wants to estimate the mean pain level in a population of persons with fibromyalgia. Which situation would make him confident that his sample mean is close to the true mean in the population?
    A confidence interval with a very narrow (small) range of numbers" because every confidence interval will include the sample mean since you add or subtract the margin of error from the sample mean. So I think a confidence interval with a smaller range would make him more confident that his estimate is close to the true mean.<<same here

7. An industrial organizational psychologist uses a 7-point rating scale to estimate the mean job satisfaction in a particular industry. The mean estimate is 4, and the 95% confidence interval ranges from 3.89 to 4.09. What is meant by the width of the confidence interval?
    The range from 3.89 to 4.09

8. An industrial organizational psychologist uses a 7-point rating scale to estimate the mean job satisfaction in a particular industry. The mean estimate is 4, and the 95% confidence interval ranges from 3.89 to 4.09. Which of the following is true?
    There is a 95% chance that the range from 3.89 to 4.09 includes the true mean satisfaction rating in the entire population of employees

9. A researcher studying chronic pain wants to estimate the mean pain level in a population of persons with fibromyalgia. The range in which she is pretty sure the true population mean (parameter) lies is known as what?
    A confidence interval

10. A clinical researcher estimates the rate of clinical depression among new mothers at 15%, and the corresponding confidence interval ranges from 7% to 23%. Which of the following interpretations is true about the confidence interval
    - There is a 95% probability that the range from 7% to 23% includes the true percentage of mothers in the population

11. Using a 0-10 scale, a health psychologist reports that the sample mean pain rating is 7 with a 95% confidence interval ranging between 6.5 to 7.5. Which of the following is true?
    - There is a 95% probability that the interval from 6.5 and 7.5 contains the true mean pain rating in the population

12. A clinical researcher estimates the rate of clinical depression among new mothers at 15%, with a confidence interval ranging between 13% and 17%. Which of the following is true?
    - There is a 95% probability that the interval from 13% and 17% contains the true percentage of depressed mothers in the population
    -

13. A clinical researcher estimates the mean depression score among new mothers as 19, and the corresponding confidence interval ranges from 17 to 21. Which of the following interpretations is true about the confidence interval.
    - If the researcher repeated the study with a new sample, there is a 95% chance that the new sample mean would fall between 17 to 21

14. A CNN Presidential poll reports a confidence interval for the Trump approvals ranging from 38% to 44%, and a New York Times poll reports a confidence interval of 37% to 45%. The standard error (standard deviation of the estimates) is larger in which poll?

- New York Times

15. The texting study from class reports that college-age students check their phones for texts an average of 16 times with a confidence interval ranging between 12 and 20. Which of the following is a correct interpretation of the confidence interval.
    - There is a 95% probability that the range from 12 and 20 includes the true mean number of text checks in the entire population of students

100a Midterm Exam

Participants suffering from schizophrenia were randomly assigned to a placebo (fake pill) condition or a medication regimen. What level of measurement describes the DrugCondition (placebo vs. control) variable?
- Nominal

Section 1: Terminology and Concepts

1. Participants suffering from schizophrenia were randomly assigned to a placebo (fake pill) condition or a medication regimen. What level of measurement describes the DrugCondition (placebo vs. control) variable?
   Nominal

2. The outcome or dependent variable was measured on a 7-point clinician rating scale, with higher scores reflecting greater illness severity and lower scores indicating less illness severity (e.g., 1 = normal, not at all ill; 7 = among the most extremely ill). What level of measurement describes the illness severity variable?
   Ordinal

3. Participants suffering from schizophrenia were randomly assigned to a placebo (fake pill) condition or a medication regimen. What measure of central tendency is MOST appropriate for this variable?
   Mode

4. Describe the likely population of interest for the schizophrenia study.
   All people suffering from schizophrenia in the US

5. Describe the sample for the schizophrenia study.
   437 participants suffering from schizophrenia

6. The outcome or dependent variable was measured on a 7-point clinician rating scale, with higher scores reflecting greater illness severity and lower scores indicating less illness severity (e.g., 1 = normal, not at all ill; 7 = among the most extremely ill). The average illness rating is of particular interest because it represents the typical level of illness. Describe the meaning of a parameter in this context.
   Average illness rating of all people suffering schizophrenia in the US after receiving medication regimen

Section 2: Descriptive Analyses

1. What distribution shape best describes pre-test illness severity scores?
   Negatively, left-skewed

2. Referring to the frequency distribution table of pre-test illness severity scores, which value tells us how many participants have pre-test illness severity scores of 4 or lower (i.e., at or below the midpoint of the rating scale)?
   12.6%

3. Participants suffering schizophrenia were randomly assigned to a placebo (fake pill) condition or a medication regimen. What is the mode for this variable?
   Medication regimen/Drug Condition

4. The outcome or dependent variable was measured on a 7-point clinician rating scale, with higher scores reflecting greater illness severity and lower scores indicating less illness severity (e.g., 1 = normal, not at all

ill; 7 = among the most extremely ill). What pre-test illness severity score value divides the distribution in half (i.e., What is the middle score)?

6

5. The outcome or dependent variable was measured on a 7-point clinician rating scale, with higher scores reflecting greater illness severity and lower scores indicating less illness severity (e.g., 1 = normal, not at all ill; 7 = among the most extremely ill). What is the mean pre-test illness severity rating? Do not round, enter the value as you see it.

5.57

6. The outcome or dependent variable was measured on a 7-point clinician rating scale, with higher scores reflecting greater illness severity and lower scores indicating less illness severity. Why is the mean pre-test illness severity rating different than the median rating?

Because the data is negatively skewed, the mean value is pulled down lower due to extreme lower values influencing it.

7. The outcome or dependent variable was measured on a 7-point clinician rating scale, with higher scores reflecting greater illness severity and lower scores indicating less illness severity. Provide an interpretation of the pre-test illness severity standard deviation.

The average distance between a pre-test illness rating and the mean pre-test illness rating is 0.900.

8. The outcome or dependent variable was measured on a 7-point clinician rating scale, with higher scores reflecting greater illness severity and lower scores indicating less illness severity. Use jamovi to get the z-score for a participant with a pre-test illness severity illness severity rating of 4. What is the value of the z-score? Do not round, enter the value exactly as you see it.

-1.747

9. The outcome or dependent variable was measured on a 7-point clinician rating scale, with higher scores reflecting greater illness severity and lower scores indicating less illness severity. Use jamovi to get the z-score for a participant with a pre-test illness severity illness severity rating of 4. Provide an interpretation of the z-score.

The pre-test illness severity rating of 4 falls below the mean by a distance that equals 1.747 standard deviations.

Section 3: Pre-Test Comparison

1. Consider the pre-test illness severity means for the drug and placebo conditions. Which group's sample mean is probably the MOST SIMILAR to its corresponding unknown population mean (i.e., Which has the greatest precision or the least sample error?)

Drug condition

2. Use Cohen's d effect size to determine whether the drug and placebo conditions are equivalent prior to starting the studying with respect to their pre-test illness severity illness severity ratings. What is the value of Cohen's d? Do not round, enter the value exactly as it appears.

0.00294

3. Use Cohen's d effect size to determine whether the drug and placebo conditions are equivalent prior to starting the studying with respect to their pre-test illness severity illness severity ratings. Provide an interpretation of the numeric value Cohen's d effect size. Note that this question is not asking you to judge the magnitude or size of the effect.

The treatment and control group averages differ by an amount equal to 0.00294 of a standard deviation unit

4. Use Cohen's d effect size to determine whether the drug and placebo conditions are equivalent prior to starting the studying with respect to their pre-test illness severity illness severity ratings. Use established benchmarks or rules of thumb to describe the magnitude or size of the pre-test mean difference.

The magnitude of the pre-test mean difference is null or very small.

5. Consider the pre-test illness severity mean and its standard error for the drug condition. Use the "plus or minus 1.96 standard error" rule to compute the 95% confidence interval around this group's average. What

is the LOWER limit of the 95% confidence interval? Use the values exactly as jamovi prints them for your calculator, round only your final answer to 3 decimal places.
5.472

6. Consider the pre-test illness severity mean and its standard error for the drug condition. Use the "plus or minus 1.96 standard error" rule to compute the 95% confidence interval around this group's average. What is the UPPER limit of the 95% confidence interval? Use the values exactly as jamovi prints them for your calculator, round only your final answer to 3 decimal places.
5.668

7. Consider the pre-test illness severity mean for the drug condition. Another question in this section asks you to compute the 95% confidence interval around this group's average. Provide an interpretation of the confidence interval.
We are 95% confident that the true (population) pre-test severity mean falls between 5.472 to 5.668.

Section 4: Primary Analyses

1. What distribution shape best describes week 6 illness severity scores for the drug condition?
Positively, right-skewed

2. Consider the week 6 illness severity mean for the drug condition. Which value on the printout tells you the expected difference between the sample mean and the mean of the entire population of participants who could receive the medication? Do not round, enter the value exactly as you see it.
0.0809

3. Use Cohen's d effect size to determine whether the drug and placebo conditions are equivalent after participating in the study for six weeks. What is the value of Cohen's d? Do not round, enter the values exactly as it appears.
1.09 got this too

4. Use Cohen's d effect size to determine whether the drug and placebo conditions are different after participating in the study for six weeks. Provide an interpretation of the numeric value Cohen's d effect size. Note that this question is not asking you to judge the magnitude or size of the effect.
The drug and control placebo averages differ by an amount equal to 1.09 of a standard deviation unit

5. Use Cohen's d effect size to determine whether the drug and placebo conditions are different after participating in the study for six weeks. Use established benchmarks or rules of thumb to describe the magnitude or size of the mean difference at the end of the study.
The magnitude of the pre-test mean difference is very large, the difference between the two means is **larger than one** standard deviation

6. Use the group means and Cohen's d effect size to evaluate the overall results of the study. How would you characterize the effectiveness of the drug relative to a placebo (fake medication)?
Because the difference is large, according to Cohen's d, the drug is very effective relative to a placebo.

7. The standard errors of the drug and placebo group averages are quite different. What feature of the study or the data explain the difference between these two values?
The large sample size difference between the drug and the placebo groups explains the differences between their standard errors.

Assignment #3

1. Which of the following interpretations is correct for the standard error printed by jamovi?
   a. D or A (not sure which)
   b. I got B

2. The R code displays a histogram and kernel density plot of the 10,000 sample means. How would you characterize the shape of the distribution of the means (i.e., the sampling distribution).
   a. C
3. The R code reports the standard deviation of the 10,000 sample means (i.e., a simulation- based estimate of the standard error). Enter the standard deviation of the estimates. Do not round, enter the value exactly as it appears in the output.
   a. 0.077
4. Which of the following interpretations is correct for the standard deviation of the 10,000 mean estimates?
   a. C ← I got C
   b. I got A? <- I also got A
5. How would you characterize the difference between the standard error printed by jamovi (i.e., the standard error computed via an equation) and the simulation-based standard error from the computer simulation (i.e., the standard deviation of the estimates)?
   a. B
6. The R code displays a histogram and kernel density plot of the 10,000 sample means. How would you characterize the shape of the distribution of the means (i.e., the sampling distribution).
   a. C
7. The R code reports the standard deviation of the 10,000 sample means (i.e., a simulation- based estimate of the standard error). Enter the standard deviation of the estimates. Do not round, enter the value exactly as it appears in the output.
   a. 0.294
8. Which of the following interpretations is correct for the standard deviation of the 10,000 estimates?
   a. A
   b. i got C
9. What is your conclusion about the impact of sample size on the precision of the estimates?
   a. B

8. Hypothesis Testing

Class Prep Quiz #8
1. -Fill in the blank with the statement that is true about the p-value. The p-value gets smaller as

   _____.

   The estimate from the sample data gets farther from the mean predicted by the null hypothesis (i.e., becomes more "absurd")
2. -Which of the following best describes the p-value?
   A measure of how rare the data (the sample estimate) are if the null hypothesis is true
3. An industrial organizational psychologist uses a 7-point rating scale to estimate the mean job satisfaction in a particular industry (1 = very dissatisfied ... 4 = neutral ... 7 = very satisfied). She has reason to believe that the particular industry she is studying has better satisfaction than most. To test this prediction, she uses a null hypothesis that says the true satisfaction mean is 4 (neutral). Which of the following sample means would be most likely to discredit the null hypothesis?
   7
4. A clinical psychologist is evaluating a new treatment for depression. The null hypothesis is that the mean is 18 (the mean of a national norm group with mild depression). Following two months of treatment, his sample of 50 participants has a mean of 17 (an improvement of one point), the p-value for which is .10. Which of the following conclusions is warranted?
   The new treatment is ineffective at reducing depression (the data was consistent with the null)

5. A clinical psychologist is evaluating a new treatment for depression. The null hypothesis is that the mean is 18 (the mean of a national norm group with mild depression). Following two months of treatment, his sample of 50 participants has a mean of 17 (an improvement of one point), the p-value for which is .10. Which of the following interpretations is true concerning the p-value?
   If he took samples of 50 participants over and over again, he would expect about 1 out of 10 to give an estimate that is a point or more below the hypothesized mean of 18

6. An industrial organizational psychologist uses a 7-point rating scale to estimate the mean job satisfaction in a particular industry (1 = very dissatisfied, 4 = neutral, 7 = very satisfied). Suppose the researcher wants to test the hypothesis that true mean is 4 (neutral). Which of the following sample means is most "absurd" or "rare" if the hypothesis is true? Check all that apply.
   1, 7 <- i got this too

7. A clinical psychologist is evaluating a new treatment for depression. Which of the following is consistent with the procedure for null hypothesis significance testing?
   Assume that the treatment has no effect, then refute or contradict that assumption with a sample that shows a marked reduction in depression scores

8. An industrial organizational psychologist uses a 7-point rating scale to estimate the mean job satisfaction in a particular industry (1 = very dissatisfied ... 4 = neutral ... 7 = very satisfied). She has reason to believe that the particular industry she is studying could have better or worse satisfaction than most. To test this prediction, she uses a null hypothesis that says the true satisfaction mean is 4 (neutral). The TWO-TAILED p-value for her sample mean is .04 (4%). Which of the following conclusions is true?
   The sample mean is much higher OR much lower than 4

9. An industrial organizational psychologist uses a 7-point rating scale to estimate the mean job satisfaction in a particular industry (1 = very dissatisfied ... 4 = neutral ... 7 = very satisfied). She has reason to believe that the particular industry she is studying could have better satisfaction than most. To test this prediction, she uses a null hypothesis that says the true satisfaction mean is 4 (neutral). Which of the following p-values is consistent with the finding that the estimate from her data discredits the null hypothesis?
   p = .01

10. An industrial organizational psychologist uses a 7-point rating scale to estimate the mean job satisfaction in a particular industry (1 = very dissatisfied ... 4 = neutral ... 7 = very satisfied). She has reason to believe that the particular industry she is studying has better satisfaction than most. To test this prediction, she uses a null hypothesis that says the true satisfaction mean is 4 (neutral). The ONE-TAILED p-value for her sample mean is .04 (4%). Which of the following conclusions is true?
    The sample mean is much higher than 4

Content Quiz #8

1. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group. The sample mean is 17, and the population mean for the national norm group is 20. The one-tailed probability for the sample mean is p = .04, which they declare as statistically significant. What would the conclusion be had they instead performed a two-tailed significance test?
   Results would be non-significant

2. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group. The sample mean is 17, and the population mean for the national norm group is 20. The one-tailed probability for the sample mean is p = .22. Based on the probability value, which of the following interpretations is correct?
   The new treatment had no impact on depression scores

3. A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons and recorded their average tip. Believing that the true tip average in the population is somewhere closer to 20%, the researchers conduct a one-tailed hypothesis test where the expected direction of the effect is lower than 25%. Which of the following corresponds to the alternate hypothesis for the test?

The population mean tip < 25%

4. A researcher studying chronic pain concludes that his new treatment protocol produced a "statistically significant" improvement in self-report pain levels. Which of the following is most likely true?

The sample mean is very different the mean predicted by the null hypothesis

5. A psychologist conducts a one-tailed hypothesis test to determine whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group, the value of which is $\mu = 20$. The graphic below depicts the sampling distribution that the researcher uses to evaluate statistical significance. Suppose that the researcher obtains an estimate that is deemed "non-significant". In which region(s) of the curve does the estimate fall?

B, B or C

6. A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons that produced a sample mean or 27%. Using a two-tailed hypothesis test, the researchers found no significant difference, with a probability value of .06. Which of the following interpretations is correct?

The average tip in the US is 25%

7. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group. The sample mean is 17, and the population mean for the national norm group is 20. The two-tailed probability for the sample mean is p = .08, which is non-significant. Suppose she had instead performed a one-tailed hypothesis test expecting a decrease. What would her conclusion be?

Results are statistically significant

8. A psychologist studying chronic pain reports that his new treatment protocol produced a "statistically significant" reduction in pain. What do we know about the probability value for the analysis?

The probability is less than 0.5

9. A college entrance exam has a mean of 500. A small liberal arts college has 100 applicants with a mean of 510. The college conducts a hypothesis test to determine whether its applicants are significantly better than the general population of applicants. The one-tailed probability value for the analysis is p = .08. Based on the probability value, which of the following interpretations is correct?

   The colleges applicants are no different from the general population of test takers

10. A psychologist conducts a two-tailed hypothesis test to determine whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group, the value of which is $\mu = 20$. The graphic below depicts the sampling distribution that the researcher uses to evaluate statistical significance. Suppose that the researcher obtains an estimate that is deemed "statistically significant". In which region(s) of the curve does the estimate fall?

   A or C

   A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons and recorded their average tip. Believing that the true tip average in the population is somewhere closer to 20%, the researchers conduct a one-tailed hypothesis test where the expected direction of the effect is lower. Which sample mean could potentially lead the researchers to declare their results as statistically significant?

   22%

   A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group. The sample mean is 17, and the population mean for the national norm group is 20. The one-tailed probability for the sample mean is p = .04, which they declare as statistically significant. What would the conclusion be had they instead performed a two-tailed significance test?

   Results would be non-significant

   A clinical psychologist is evaluating a new treatment for depression. The null hypothesis is that the mean is 18 (the mean of a national norm group with mild depression). Following two months of treatment, his sample of 50 participants has a mean of 17 (an improvement of one point), the p-value for which is .10. Which of the following interpretations is true concerning the p-value?

If he took samples of 50 participants over and over again, he would expect about 1 out of 10 to give an estimate that is a point or more below the hypothesized mean of 18

A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group. The sample mean is 17, and the population mean for the national norm group is 20. The one-tailed probability for the sample mean is p = .04. Which of the following interpretations is correct regarding the probability value?

If the null hypothesis is true, there is an 4% probability of obtaining a mean of 17 or lower

A psychologist studying depression reports that her new treatment produced a "non-significant" reduction in depression. What do we know about the probability value for the study?

The probability is greater than .05

Fill in the blank with the statement that is true about the p-value. The p-value gets smaller as _____.

The estimate from the sample data gets farther from the mean predicted by the null hypothesis (i.e., becomes more "absurd")

An industrial organizational psychologist uses a 7-point rating scale to estimate the mean job satisfaction in a particular industry (1 = very dissatisfied ... 4 = neutral ... 7 = very satisfied). She has reason to believe that the particular industry she is studying could have better or worse satisfaction than most. To test this prediction, she uses a null hypothesis that says the true satisfaction mean is 4 (neutral). The TWO-TAILED p-value for her sample mean is .04 (4%). Which of the following conclusions is true?

The sample mean is much higher OR much lower than 4

A college entrance exam has a mean of 500. A small liberal arts college has 100 applicants with a mean of 510. The college conducts a hypothesis test to determine whether its applicants are significantly better than the general population of applicants. The one-tailed probability for the sample mean is p = .03. Based on the probability value, which of the following interpretations is correct?

The analysis refutes the null hypothesis

A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she will compare the sample mean from her study to the mean of a national norm group, the value of which is μ = 20. Which of the following represents the null hypothesis for the study?

The population average depression score equals 20

11. A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons that produced a sample mean or 27%. Using a two-tailed hypothesis test, the researchers found a significant difference, with a probability value of .04. Which of the following interpretations is correct?

    a. The average tip in the US is significantly higher than 25%

12. A college entrance exam has a mean of 500. A small liberal arts college has 100 applicants with a mean of 510. The college conducts a hypothesis test to determine whether its applicants are significantly better than the general population of applicants. The one-tailed probability for the sample mean is $p = .08$. Based on the probability value, which of the following interpretations is correct?

    a. The analysis supports the null hypothesis

13. A researcher studying chronic pain concludes that his new treatment protocol produced a "non-significant" improvement in self-report pain levels. Which of the following is most likely true?

    a. The sample mean is very similar to the mean predicted by the null hypothesis

14. A psychologist conducts a one-tailed hypothesis test to determine whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group, the value of which is $\mu = 20$. The graphic below depicts the sampling distribution that the researcher uses to evaluate statistical significance. Suppose that the researcher obtains an estimate that refutes the null. In which region(s) of the curve does the estimate fall?

    a. A

15. A psychologist studying chronic pain reports that his new treatment protocol produced a "non-significant" reduction in pain. What do we know about the probability value for the study?

    a. The probability is greater than .05

16. A college entrance exam has a mean of 500. A small liberal arts college has two years of applications, each with 100 prospective students. The college conducts hypothesis tests to determine whether its applicants are significantly better than the general population of applicants. The first year of applicants has a mean of 510, whereas the second year of applicants has a mean of 505. Which sample will produce the smallest probability value?

    a. The sample with a mean of 510

17. A college entrance exam has a mean of 500. A small liberal arts college has 100 applicants with a mean of 510. The college conducts a hypothesis test to determine whether its applicants are significantly better than the general population of applicants. The one-tailed probability value for the analysis is $p = .08$. Based on the probability value, which of the following interpretations is correct?

    a. The college's applicants are no different from the general population of test takers

18. A college entrance exam has a mean of 500. A small liberal arts college has two years of applications, each with 100 prospective students. The college conducts hypothesis tests to determine whether its applicants are significantly better than the general population of applicants. The first year of applicants has

a mean of 510, whereas the second year of applicants has a mean of 505. Which sample will produce the largest probability value?

    a.   The sample with a mean of 505

19. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group. The sample mean is 17, and the population mean for the national norm group is 20. The one-tailed probability for the sample mean is $p = .04$. Which of the following interpretations is correct regarding the probability value?

If the null hypothesis is true, there is an 4% probability of obtaining a mean of 17 or lower

A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons and recorded their average tip. Believing that the true tip average in the population is somewhere closer to 20%, the researchers conduct a one-tailed hypothesis test where the expected direction of the effect is lower. Which sample mean could potentially lead the researchers to declare their results as statistically significant?

22%

A college entrance exam has a mean of 500. A small liberal arts college has two years of applications, each with 100 prospective students. The college conducts hypothesis tests to determine whether its applicants are significantly better than the general population of applicants. The first year of applicants has a mean of 510, whereas the second year of applicants has a

The sample with a mean of 510

A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group. The sample mean is 17, and the population mean for the national norm group is 20. The one-tailed probability for the sample mean is $p = .17$. Based on the probability value, which of the following interpretations is correct?

Results are non-significant

A psychologist studying depression reports that her new treatment produced a "non-significant" reduction in depression. What do we know about the probability value for the study?

The probability is greater than .05

A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group. The sample mean is 17, and the population mean for the national norm group is 20. The two-tailed probability for the sample mean is $p = .08$. Suppose she had instead performed a one-tailed hypothesis test expecting a decrease. What would the one-tailed probability value be (roughly)?

.04

A college entrance exam has a mean of 500. A small liberal arts college has 100 applicants with a mean of 510. The college conducts a hypothesis test to determine whether its applicants are significantly different than the general population of applicants. The two-tailed probability value for the analysis is $p = .08$. Which of the following interpretations is correct regarding the probability value?

If the null hypothesis is true, there is an 8% probability of obtaining a mean of 510 or larger or 490 or smaller

A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons that produced a sample mean or 22%. Using a two-tailed hypothesis test, the researchers found a significant difference, with a probability value of .04. Which of the following interpretations is correct?

The average tip in the US is significantly lower than 25%

A psychologist studying depression reports that her new treatment produced a "statistically significant" reduction in depression. What do we know about the probability value for the study?
The probability is less than .05

A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons that produced a sample mean or 22%. Using a two-tailed hypothesis test, the researchers found no significant difference, with a probability value of .06. Suppose that they had instead conducted a one-tailed test where the expected direction of the effect was lower than 25%. What conclusion would result from the one-tailed test?
The average tip in the US is significantly lower than 25%

A college entrance exam has a mean of 500. A small liberal arts college has 100 applicants with a mean of 510. The college conducts a hypothesis test to determine whether its applicants are significantly better than the general population of applicants. The one-tailed probability for the sample mean is $p = .08$. Based on the probability value, which of the following interpretations is correct?
The analysis supports the null hypothesis

A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group. The sample mean is 17, and the population mean for the national norm group is 20. The one-tailed probability for the sample mean is $p = .04$. Suppose that she instead performs a two-tailed hypothesis test. What would the two-tailed probability value be (roughly)?

.08

A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons and recorded their average tip. Believing that the true tip average in the population is somewhere closer to 20%, the researchers conduct a one-tailed hypothesis test where the expected direction of the effect is lower than 25%. Which of the following corresponds to the null hypothesis for the test?
The population mean tip is 25%

A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group. The sample mean is 17, and the population mean for the national norm group is 20. The one-tailed probability for the sample A psychologist conducts a two-tailed hypothesis test to determine whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group, the value of which is $\mu = 20$. The graphic below depicts the sampling distribution that the researcher uses to evaluate statistical significance. Suppose that the researcher obtains an estimate that is deemed "statistically significant". In which region(s) of the curve does the estimate fall?mean is p = .04. Based on the probability value, which of the following interpretations is correct?
The new treatment was effective at reducing depression

A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons and recorded their average tip. Assuming that researchers perform a two-tailed hypothesis test, which of the following means is most likely to produce a non-significant finding?

25%
A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons and recorded their average tip. Assuming that researchers perform a two-tailed hypothesis test, which of the following means is most likely to produce a statistically significant finding?
21%

A psychologist conducts a one-tailed hypothesis test to determine whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group, the value of which is $\mu = 20$. The graphic below depicts the sampling distribution that the researcher uses to evaluate statistical significance. Suppose that the researcher obtains an estimate that is deemed "statistically significant". In which region(s) of the curve does the estimate fall?
A

 A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group. The sample mean is 17, and the population mean for the national norm group is 20. The one-tailed probability for the sample mean is p = .04. Based on the probability value, which of the following interpretations is correct?
Results are statistically significant

A psychologist conducts a two-tailed hypothesis test to determine whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she compares the sample mean from her study to the mean of a national norm group, the value of which is $\mu = 20$. The graphic below depicts the sampling distribution that the

researcher uses to evaluate statistical significance. Suppose that the researcher obtains an estimate that is deemed "statistically significant". In which region(s) of the curve does the estimate fall?

A or C

A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons that produced a sample mean or 22%. Using a two-tailed hypothesis test, the researchers found a significant difference, with a probability value of .04. Which of the following interpretations is correct?

The average tip in the US is significantly lower than 25%

A college entrance exam has a mean of 500 and a standard deviation of 100 in the population. A small liberal arts college has 100 applicants with a mean of 510. The college wants to determine whether its applicants have significantly higher scores than the general population of applicants. Which of the following null hypotheses is most appropriate for this scenario?

The population average exam score equals 500

A USA Today newspaper article suggested that a 25% tip might be the new national norm for restaurant patrons (i.e., the average tip % in the US). To test this proposition, suppose that researchers collected a sample of 25 restaurant patrons and recorded their average tip. Believing that the true tip average in the population is somewhere closer to 20%, the researchers conduct a one-tailed hypothesis test where the expected direction of the effect is lower than 25%. Which of the following means is most likely to produce a significant finding?

21%

A college entrance exam has a mean of 500. A small liberal arts college has 100 applicants with a mean of 510. The college conducts a hypothesis test to determine whether its applicants are significantly better than the general population of applicants. The one-tailed probability value for the analysis is p = .03. Based on the probability value, which of the following interpretations is correct?

The college's applicants score higher than the general population of test takers

Class Prep Quiz #9
1. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she used a test statistic to compare the sample mean from her study to the mean of a national norm group (the population mean). She obtains a t test statistic of 1.64 and a one-tailed probability of .06. What is the correct conclusion about results of the study?

   The treatment did not impact depression
2. A university runs a new class designed to improve the Freshman experience. The hypothesis is that participation in the new class will change attitudes toward school. To evaluate the hypothesis, the research team uses a test statistic to compare a sample mean to the mean of the general population of students who don't take the class. What size and magnitude of the test statistic would lead the team to refute the null hypothesis (i.e., prove it "guilty", demonstrate it to be "absurd"), concluding that the class changes attitudes?

   A large positive test statistic
3. A college entrance exam has a mean of 500 and a standard deviation of 100 in the population. A small liberal arts college has 100 applicants with a mean of 510. The college wants to determine whether its applicants have significantly higher scores than the general population of applicants. A researcher performs uses a test statistic and obtains a one-tailed probability of .40. What test statistic most likely produced this p-value?

   .2

4. A college entrance exam has a mean of 500 and a standard deviation of 100 in the population. A small liberal arts college has 100 applicants with a mean of 510. The college wants to determine whether its applicants have significantly higher scores than the general population of applicants. A researcher uses a test statistic and obtains a probability value of .02. What test statistic most likely produced this p-value?
2.5
5. A psychologist is interested in determining whether a new treatment for depression decreases (improves) depression scores. To evaluate the treatment, she uses a one-tailed t test to compare the sample mean from her study to the mean of a national norm group. What size and magnitude of a test statistic would lead her to refute the null hypothesis (i.e., prove it "guilty", demonstrate it to be "absurd"), concluding that the treatment is effective?
A very large negative test statistic
6. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she will use a test statistic. The null hypothesis involves a prediction about what parameter?
The true population mean
7. A college entrance exam has a mean of 500 and a standard deviation of 100 in the population. A small liberal arts college has 100 applicants with a mean of 510. The college wants to determine whether its applicants have significantly higher scores than the general population of applicants. A researcher obtains a t test statistic of 1.96 and a probability of .049. What is your conclusion about results of the study?
The college's applicants score significantly higher than the general population
8. A university runs a new class designed to improve the Freshman experience. The hypothesis is that participation in the new class will improve attitudes toward school. To evaluate the hypothesis, the research team uses a test statistic to compare the sample mean from a sample to the mean of a national norm group. The assumed population mean (the expected mean if the null is true) is 20, the mean from a sample of 25 students is 22, and the standard error is 4. What is the value of the test statistic from this analysis?
0.5

Content Quiz #9
1. A university runs a new class designed to improve the Freshman experience. The researchers want to evaluate whether participation in the new class will improve attitudes toward school. To evaluate the hypothesis, the research team uses a t test statistic to compare the sample mean from a sample to the mean of a national norm group. The population mean is 20, the mean from a sample of 25 students is 23, and the standard error is 2. What is the null hypothesis for the analysis?
The population mean = 20
2. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she used a test statistic to compare the sample mean from her study to the mean of a national norm group (the population mean). The sample mean is identical to the national norm group mean. Which test statistic is most consistent with this result?
0
3. A university runs a new class designed to improve the Freshman experience. The hypothesis is that participation in the new class will improve (increase) attitudes toward school. To evaluate the hypothesis, the research team uses a test statistic to compare a same that the class improves attitudes?
A large positive test statistic
4. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she used a test statistic to compare the sample mean from her study to the mean of a national norm group (the population mean). She obtains a t test statistic of -2.64 and a one-tailed probability of .02. What is the correct conclusion about results of the study?

The treatment improved (decreased) depression

5. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she used a z-statistic to compare the sample mean from her study to the mean of a national norm group (the population). She then reanalyzed the data with a t-statistic and found that the two tests produced identical probability values. Which situation explains the similarity of the two tests?
The sample size is very large

6. Which of the following represents the numeric value that you get from a test statistic like a z-test or t-test?
The degree of difference between the sample mean and expected mean if the null is true on a standardized metric

7. Suppose a researcher obtained a test statistic value of 2. Which of the following interpretations is correct?
The difference between the sample mean and the mean predicted by the null hypothesis is twice as large as the difference we would expect from sampling error ← not sure about this one ← I just took it and this is the right answer

8. A college entrance exam has a mean of 500 and a standard deviation of 100 in the population. A small liberal arts college has 100 applicants with a mean of 510. The college wants to determine whether its applicants have significantly higher scores than the general population of applicants. The data analyst uses a test statistic to evaluate the research question. The null hypothesis involves a prediction about what parameter?
The true population mean

9. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the hypothesis, the researcher uses a test statistic to compare the sample mean from a sample to the mean of a national norm group. The population mean is 10, the mean from a sample of 9, and the standard error is 2. What is the value of the test statistic from this analysis?
0.5

10. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she used a test statistic to compare the sample mean from her study to the mean of a national norm group (the population mean). She obtains a test statistic of -1.64 and a one-tailed probability of .06. What would the probability value be if she instead performed a two-tailed test?
.12

11. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she used a test statistic to compare the sample mean from her study to the mean of a national norm group (the population mean). The sample mean is significantly lower than the national norm group mean, and she concludes that the new treatment is effective. Which test statistic is most consistent with this result?
-2

12. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she used a test statistic to compare the sample mean from her study to the mean of a national norm group (the population mean). She finds a "non-significant" decrease in depression. Which test statistic is most consistent with this result?
-.20

13. A college entrance exam has a mean of 500 and a standard deviation of 100 in the population. A small liberal arts college has 100 applicants with a mean of 510. The college wants to determine whether its applicants have significantly higher scores than the general population of applicants. The data analyst obtains a t statistic of 1.70 and a two-tailed probability of .04. What would the probability be if the analyst performed a one-tailed test?
.02

14. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she used a test statistic to compare the sample mean from her study to the mean of a national norm group (the population mean). She obtains a t statistic of -1.64 and a two-tailed probability of .06. What would the probability value be if she instead performed a one-tailed test?

.03

15. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. The researcher uses a t test statistic to compare the sample mean from a sample to the mean of a national norm group. The population mean is 10, the mean from a sample of 9 clients is 8, and the standard error is 2. What is the null hypothesis for the analysis?

The population mean= 10

16. A college entrance exam has a mean of 500 and a standard deviation of 100 in the population. A small liberal arts college has 100 applicants with a mean of 510. The college wants to determine whether its applicants have significantly higher scores than the general population of applicants. The data analyst obtains a t statistic of 1.70 and a one-tailed probability of .04. What would the probability be if the analyst performed a two-tailed test?

.08

17. A college entrance exam has a mean of 500 and a standard deviation of 100 in the population. A small liberal arts college has 100 applicants with a mean of 510. The college wants to determine whether its applicants have significantly higher scores than the general population of applicants. A researcher obtains a t test statistic of 1.70 and a probability of .051. What is your conclusion about the results of the study?

The college's applicants score the same as the general population

18. A college entrance exam has a mean of 500 and a standard deviation of 100 in the population. A small liberal arts college has 100 applicants with a mean of 510. The college wants to determine whether its applicants have significantly higher scores than the general population of applicants. The data analyst obtains a t statistic of 1.70 and a one-tailed probability of .04. What would the probability be if the analyst performed a two-tailed test?

0.08

19. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she used a test statistic to compare the sample mean from her study to the mean of a national norm group (the population mean). She obtains a t test statistic of 1.64 and a one-tailed probability of .06. What is the correct conclusion about results of the study?

The treatment did not impact depression

10. Paired Samples t-Test

Class Prep Quiz #10

1. A psychologist is interested in determining whether a new treatment for depression impacts depression scores. She recruits 30 subjects and measures their depression at baseline then again following the intervention. At the conclusion of the study she uses a paired-samples t-test to evaluate whether scores change. The paired-samples t-statistic is 1.80 with a two-tailed probability of .20. Which of the following conclusions is true regarding the treatment?

The treatment changed scores in a manner that is consistent with random chance

2. An educational psychologist is interested in whether test scores improve between 7th and 8th grade. She recruits a sample of 100 students and finds test score averages of 50 and 52 in 7th and 8th grade, respectively. The standard error for the paired-samples t-statistic is 2.25. Which of the following conclusions is true regarding the change in scores?
The change in scores from 7th to 8th grade is about what you would expect due to sampling error (random chance) alone ← i took the content quiz and this was correct

3. An educational psychologist is interested in whether test scores improve between 7th and 8th grade. She recruits a sample of 100 students and finds test score averages of 50 and 53 in 7th and 8th grade, respectively. The paired-samples t-statistic is 1.85 and has a one-tailed probability value of .08. Which of the following conclusions is true regarding the change in scores?
Scores did not change from 7th to 8th grade

4. An educational psychologist is interested in whether math test scores improve between 7th and 8th grade. She recruits a sample of 100 students and uses a paired-samples t-test to analyze the data. Which of the following is a valid null hypothesis for the study?
The difference between the average 7th and 8th grade math scores is 0 in the population

5. Which of the following is NOT an example of a within-subjects design that would yield data appropriate for a paired-samples t-test?
Males and females are compared to see whether they differ in their depression levels

6. A clinical psychologist is studying the progression of alcohol use in a teenage population. She collects a measure of alcohol use from the same group of teens at ages 16 and 18. Because the two measurements are obtained from the same people, she uses a paired-samples t-test to analyze the data. The null hypothesis predicts no change in alcohol use. The sample data are as follows: age 16 mean = 6, age 18 mean = 6.25, p = .01. Based on the probability value, which is the conclusion about the null hypothesis?
The data refute (i.e., are inconsistent with) the null

7. A clinical researcher measures depression at the beginning of a study, then she measures it a second time after participants have completed three months of therapy. Which of the following describes a difference or change scores in this context?
The amount by which each person's score changes from the beginning to the end of the study ← i took the content quiz and this was correct

8. A clinical psychologist is studying the progression of alcohol use in a teenage population. She collects a measure of alcohol use from the same group of teens at ages 16 and 18. Because the two measurements are obtained from the same people, she uses a paired-samples t-test to analyze the data. The null hypothesis predicts no change in alcohol use. The sample means are as follows: age 17 mean = 6, age 18 mean = 6.25, p = .15. Based on the probability value, which is the conclusion about the null hypothesis?
The data support (i.e., are consistent with) the null

9. A psychologist is studying marital satisfaction, and she wants to determine whether husbands OR wives report higher satisfaction. She recruits a sample of 100 husband-wife pairs and measures marital satisfaction on a 1-10 scale. If she uses a paired-samples t-test to analyze the data, which of the following sets of means provides the MOST evidence against the null hypothesis?
Husband mean = 5, wife mean = 7

10. Which of the following is an example of a within-subjects design that would yield data that could be analyzed with a paired-samples t-test?
Pretest and posttest scores are obtained from a group of people prior to and after they participate in a smoking cessation intervention

Content Quiz #10

1. An exercise physiologist is studying energy exertion from playing video games. Participants first play Wii bowling for 15 minutes, during which researchers measure calories burned. After a period of rest, participants then play Will tennis. Because the two measurements are obtained from the same people, the researchers use a paired-samples t-test to analyze the data. The null hypothesis predicts no change in energy expenditure. The sample means are as follows: bowling mean = 190, tennis mean = 200, p = .02. Based on the probability value, which is the conclusion about the null hypothesis?
   The data refute (i.e., are inconsistent with) the null

2. An educational psychologist is interested in whether test scores improve between 7th and 8th grade. She recruits a sample of 100 students and finds test score averages of 50 and 53 in 7th and 8th grade, respectively. The standard error for the paired-samples t-statistic is 2. Which of the following interpretations is true regarding the standard error?
   If the null hypothesis is true in the population, sampling error (random chance) would cause a difference of 2 points, on average

3. A clinical psychologist is studying the progression of alcohol use in a teenage population. She collects a measure of alcohol use from the same group of teens at ages 16 and 18. The average increase in alcohol use scores was .30 points, and a paired-samples t test gave a probability value of .15 (15%). Which of the following interpretations is correct regarding the probability value?
   If there is truly no change in the population, an increase of .30 or more would occur in 15% of all random samples from such a population

4. A psychologist is studying marital satisfaction, and she wants to determine whether husbands OR wives report higher satisfaction. She recruits a sample of 100 husband-wife pairs and measures marital satisfaction on a 1-10 scale. If she uses a paired-samples t-test to analyze the data, which of the following sets of means provides the MOST evidence against the null hypothesis?
   Husband mean = 5, wife mean = 7

5. An educational psychologist is interested in whether test scores improve between 7th and 8th grade. She recruits a sample of 100 students and finds test score averages of 50 and 52 in 7th and 8th grade, respectively. The standard error for the paired-samples t-statistic is 2.25. Which of the following conclusions is true regarding the change in scores?
   The change in scores from 7th to 8th grade is about what you would expect due to sampling error (random chance) alone

6. A psychologist is interested in whether participation in a smoking cessation intervention reduces smoking. She recruits a sample of 50 participants and has them rate their smoking frequency on a 1-10 scale. The pre-test and post-test means are 7 and 5, respectively. The standard error for the paired-samples t-statistic is .50. Which of the following conclusions is true regarding the change in smoking behavior?
   The change in smoking frequency is larger than what is expected due to sampling error (random chance)

7. An exercise physiologist is studying energy exertion from playing video games. Participants first play Xbox for 15 minutes, during which researchers measure calories burned. After a period of rest, participants then play Will tennis. Because the two measurements are obtained from the same people, the researchers use a paired-samples t-test to analyze the data. The null hypothesis predicts no change in energy expenditure. The sample means are as follows: Xbox = 125, tennis = 200, p = .006. Based on the probability value, which is the conclusion about the null hypothesis?
   The data refute (i.e., are inconsistent with) the null hypothesis

8. A researcher wants to study the effectiveness of a smoking cessation intervention. Which of the following research designs would require a paired-samples t-test as the analysis?
   Recruit a sample of 50 smokers, measure their smoking frequency, conduct the intervention, then measure their smoking frequency. Compare the average smoking frequency before and after the intervention.

9. A psychologist is interested in whether participation in a smoking cessation intervention reduces smoking. She recruits a sample of 50 participants and has them rate their smoking frequency on a 1-10 scale. The pre-test and post-test means are 7 and 5, respectively, for a mean change of -2. The 95% confidence interval for the mean change is -3.5 to -.5. Which of the following is true?

    There is a 95% probability that the interval from -3.5 to -.5 contains the true mean change in the population

10. Which of the following procedures is consistent with a within-subjects design that would yield data that could be analyzed with a paired-samples t-test?

    Take a sample, give all subjects a pretest before treatment, expose all to the treatment, then test all subjects a second time after treatment

11. Which of the following is NOT an example of a within-subjects design that would yield data appropriate for a paired-samples t-test?

    Males and females are compared to see whether they differ in their depression levels

12. An exercise physiologist is studying energy exertion from playing video games. Participants first play Wii bowling for 15 minutes, during which researchers measure calories burned. After a period of rest, participants then play Will tennis. The average change in energy expenditure scores was 77 points, and a paired-samples t test gave a probability value of .005 (.5% or one half of one percent). Which of the following interpretations is correct regarding these results?

    If there is truly no difference in the population, a change of 77 points or more would occur in .5% of all random samples from such a population

13. An exercise physiologist is studying energy exertion from playing video games. Participants first play Wii bowling for 15 minutes, during which researchers measure calories burned. After a period of rest, participants then play Will tennis. The average change in energy expenditure scores was 77 points, and a paired-samples t test gave a probability value of .005 (.5% or one half of one percent). Which of the following interpretations is correct regarding these results?

    There is a reliable difference in energy expenditure between the two active video games

14. A psychologist is studying marital satisfaction, and she wants to determine whether husbands or wives report higher satisfaction. She recruits a sample of 100 husband-wife pairs and measures marital satisfaction on a 1-10 scale. If she uses a paired-samples t-test to analyze the data, which of the following sets of means provides the MOST evidence in support of the null hypothesis?

    Husband mean = 7, wife mean = 7

15. A clinical psychologist is studying the progression of alcohol use in a teenage population. She collects a measure of alcohol use from the same group of teens at ages 16 and 18. The average increase in alcohol use scores was .30 points, and a paired-samples t test gave a probability value of .15 (15%). Which of the following interpretations is correct regarding these results?

    Alcohol use is steady and does not change between the ages of 16 and 18

16. Which of the following research scenarios would require a paired-samples t-test?

    Identical twins are compared to see whether they differ in their temperament

17. An educational psychologist is interested in whether test scores improve between 7th and 8th grade. She recruits a sample of 100 students and finds test score averages of 50 and 53 in 7th and 8th grade, respectively, for an average change of 3. The 95% confidence interval for the mean change is 1 to 6. Which of the following is true?

    There is a 95% probability that the interval from 1 to 6 contains the true mean change in the population

18. A clinical psychologist is studying the progression of alcohol use in a teenage population. She collects a measure of alcohol use from the same group of teens at ages 16 and 18. One a 10-point frequency scale, he finds a alcohol mean of 4.5 at age 16 and a mean of 7 at age 18, for a mean difference (change) of 2.5. The 95% confidence interval for the mean change ranges from 1.5 to 3.5. Which of the following is true?

    There is a 95% probability that the interval from 1.5 to 3.5 contains the true mean change in the population

19. Which of the following is NOT an example of a within-subjects design?

    To examine test bias, an educational researcher compares test scores between Hispanics and Caucasians

20. A psychologist is interested in whether participation in a smoking cessation intervention reduces smoking. She recruits a sample of 50 participants and has them rate their smoking frequency on a 1-10 scale. The pre-test and post-test means are 7 and 5, respectively. The standard error for the paired-samples t-statistic is .50. Which of the following conclusions is true regarding the standard error?
    a. Sampling error (random chanceWhat is the null hypothesis for the analysis?) would produce a difference of .50, even if the intervention is ineffective

11. Independent-Samples t-Test

Class Prep Quiz #11
1. -A psychologist studying chronic pain reports that his new treatment protocol produced a "non significant" difference between the treatment and control group. Which of the following conclusions is most likely true?
    The treatment and control means were very similar
2. -A psychologist is interested in determining whether a new treatment for depression impacts depression scores. She randomly assigns 30 subjects to the treatment and 30 subjects to the control. At the conclusion of the study she uses a t-test to evaluate whether the groups differ. The independent t-statistic is 2.30, and its two-tailed probability is .0251. Which of the following conclusions is true regarding the treatment?
    The treatment improves scores beyond what is expected due to random chance (estimation error)
3. -Which of the following research scenarios would require an INDEPENDENT t-test?
    To examine test bias, an educational researcher compares test scores between males and females
4. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. At the end of the study the researcher administered a widely used pain level scale to the intervention participants and a control group that received generic reading materials. Which of the following sets of means provides the strongest evidence against the null hypothesis?
    Intervention mean = 7, control mean = 8.50
5. -A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its parttwo-tailed pricipants. A group of pain patients are randomly assigned to participate in either the online program or a control group. At the end of the study the researcher administered a widely used pain level scale ranging from 1-10. The two-tailed probability value for the independent t-statistic is .02. Which of the following conclusions is definitely true?
    The treatment and control group means were different
6. -A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. She randomly assigns 30 subjects to the treatment and 30 subjects to the control. At the conclusion of the study she uses an independent t-test to evaluate whether the groups differ on a widely used depression scale that ranges between 0 and 30. The treatment group mean is 18 and the control group mean is 22. What is the null hypothesis for the analysis?
    The difference between the population means is zero
7. -A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she randomly assigns 30 subjects to the treatment and 30 subjects to the control, and she then performs an independent t-test. What population parameter is the target/focus of the hypothesis?
    The difference between the treatment and control group means
8. -A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. A group of pain patients are randomly assigned to participate in either the online program or a control group. At the end of the study the researcher administered a widely used pain level scale ranging from 1-10. The two-tailed probability value for the independent t-statistic is .67. Which of the following conclusions is definitely true?

The treatment and control group means were essentially identical

9. -Which of the following is NOT an example of a between-subjects design?
Pretest and posttest scores are obtained from a group of people prior to and after they participate in an intervention

10. -A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. A group of pain patients are randomly assigned to participate in either the online program or a control group. At the end of the study the researcher administered a widely used pain level scale ranging from 1-10. The standard error for the analysis is .50. Which of the following interpretations is correct regarding the standard error?
Due to estimation error (random chance) alone, we would expect a .50 mean difference on a 10-point scale

Content Quiz # 11

1. Social psychologists conduct a study that attempts to determine whether drawing a smiley face (the treatment) improves tip percentages relative to a control with a blank check. The mean tip percentages for the treatment and control groups are 35 and 29. The mean difference of 6 has a 95% confidence interval ranging from 1 to 11. Which of the following is true?
Boys and girls are the same in their aggression behavior

2. A clinical psychologist is studying the association between gender and aggression in young children. Using an established observational measure, she records the number of aggressive acts committed by boys and girls during a free play period at school. Because the two measurements are obtained from different samples or people, she uses an independent-samples t test to analyze the data. The difference in the number of aggressive behaviors was 1.3 points higher for males, and the t test gave a probability value of .13 (13%). Which of the following interpretations is correct regarding these results?
There is a 95% probability that the interval from 1 to 11 contains the true mean difference in the population

3. Which of the following procedures is consistent with a between-subjects design?
Take a sample, assign half of the subjects to the treatment, half to the control

4. A cognitive psychologist is comparing ADHD children to a sample of non-ADHD controls using an established measure of reaction time that captures attention span. The mean difference is 150 (the ADHD group mean is higher by 150 points), with a confidence interval for the mean difference ranging between 25 and 350. Which of the following population parameters are not supported by the data? That is, if the following values were used as the null hypothesis, which value(s) would the data refute (i.e., make appear absurd)? CHECK ALL THAT APPLY.

500, -23, 0

5. A clinical psychologist is studying the association between education level and smoking. She collects a measure of smoking frequency from a group of participants with a high school or lower education, and she collects the same measure from a group of college graduates. Because the two measurements are obtained from different samples or people, she uses an independent-samples t test to analyze the data. The difference in the number of cigarettes smoked was 3 points, and the t test gave a probability value of .01 (1%). Which of the following interpretations is correct regarding the probability value?

If there is truly no difference in the population, a difference of 3 cigarettes or more would occur in 1% of all random samples from such a population

6. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. A group of pain patients are randomly assigned to participate in

either the online program or a control group. At the end of the study the researcher administered a widely used pain level scale ranging from 1-10. The standard error for the analysis is .50. Which of the following interpretations is correct regarding the standard error?

Due to estimation error (random chance) alone, we would expect a .50 mean difference on a 10-point scale

7. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. A group of pain patients are randomly assigned to participate in either the online program or a control group. At the end of the study the researcher administered a widely used pain level scale ranging from 1-10. The two-tailed probability value for the independent t-statistic is .02. Which of the following conclusions is definitely true?

The treatment and control group means were different A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she used a test statistic to compare the sample mean from her study to the mean of a national norm group (the population mean). The sample mean is identical to the national norm group mean. Which test statistic is most consistent with this result?

8. A psychologist studying chronic pain reports that his new treatment protocol produced a "statistically significant" difference between the treatment and control group. Which of the following conclusions is most likely true?

The treatment mean was noticeably lower (better) than the control group mean

9. A psychologist is interested in determining whether a new treatment for depression impacts depression scores. She randomly assigns 30 subjects to the treatment and 30 subjects to the control. At the conclusion of the study she uses a t-test to evaluate whether the groups differ. The independent t-statistic is 2.30 with a two-tailed probability of .0251. Which of the following conclusions is true regarding the treatment?

The treatment improves scores beyond what is expected due to random chance (estimation error)

10. A psychologist studying chronic pain reports that his new treatment protocol produced a "non significant" difference between the treatment and control group. Which of the following conclusions is most likely true?

The treatment and control means were very similar

11. A clinical psychologist is studying the association between gender and aggression in young children. Using an established observational measure, she records the number of aggressive acts committed by boys and girls during a free play period at school. Because the two measurements are obtained from different samples or people, she uses an independent-samples t test to analyze the data. The analysis results are as follows: t = -1.5638, df = 20.303, p-value = 0.1333. Based on the probability value, which is the conclusion about the null hypothesis?

The data support (i.e., are consistent with) the null

12. A university runs a new class designed to improve the Freshman experience. The hypothesis is that participation in the new class will improve attitudes toward school. To evaluate the hypothesis, the research team assigns a sample of incoming Freshman to the intervention class or a standard Freshman class (control). Which of the following sets of means provides the strongest support for the researcher's prediction?

Intervention mean = 20, control mean = 14

13. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. He assigns 20 subjects to the treatment condition and 20 to the control. At the end of the study the researcher finds a standardized mean difference effect size of .30. Based on widely accepted rules of thumb, how would you characterize the magnitude of this treatment effect.

Small

14. Which of the following is NOT an example of a between-subjects design?

Pretest and posttest scores are obtained from a group of people prior to and after they participate in an intervention

15. A clinical psychologist is studying whether a call-in help line or individual therapy leads to better outcomes for individuals who are trying to quit smoking. Using an established diary measure, participants in both groups record the number of cigarettes smoked during a specified period of time during the intervention. Because the two measurements are obtained from different samples or people, the researcher uses an independent-samples t test to analyze the data. The analysis results are as follows: t = -1.2792, df = 7.4462, p-value = 0.2392. Based on the probability value, which of the following interpretations is likely true?

The quit line and therapy produce identical outcomes

16. A clinical psychologist is studying whether males or females have better outcomes when trying to quit smoking. Using an established diary measure, participants in both groups record the number of cigarettes smoked during a specified period of time during the intervention. Because the two measurements are obtained from different samples or people, the researcher uses an independent-samples t test to analyze the data. The analysis results are as follows: t = -0.893, df = 12.963, p-value = 0.3881. Based on the probability value and the estimates, which of the following interpretations is correct?
Select one:

 Males and females smoke the same amount

17. A psychologist is interested in determining whether a new treatment for depression impacts depression scores. She randomly assigns 30 subjects to the treatment and 30 subjects to the control. At the conclusion of the study she uses an independent t-test to evaluate whether the groups differ on a widely used depression scale that ranges between 0 and 30. The standard error for the analysis is 3. Which of the following interpretations is correct regarding the standard error?

Even if there is no treatment effect, you would expect the groups to differ by 3 points due to estimation error (random chance) alone

18. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. She randomly assigns 30 subjects to the treatment and 30 subjects to the control. At the conclusion of the study she uses a t-test to evaluate whether the groups differ on a widely used depression scale that ranges between 0 and 30. The treatment group (group 1) mean is 18, the control group (group 2) mean is 22, and the standard error is 2. What is the value for the t-statistic?
    a.   -2

19. A university runs a new class designed to improve the Freshman experience. The hypothesis is that participation in the new class will improve attitudes toward school. To evaluate the hypothesis, the research team assigns a sample of incoming Freshman to the intervention class or a standard Freshman course load (control). Which of the following sets of means provides the strongest support for the null hypothesis?

    a.Intervention mean = 20, control mean = 21

20. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. A group of pain patients are randomly assigned to participate in either the online program or a control group. At the end of the study the researcher administered a widely used pain level scale ranging from 1-10. What is the null hypothesis for the study?

    A. If the experiment were conducted on the entire population, the difference in the treatment and control group means would be zero

    B. If the experiment were conducted on the entire population, the treatment group mean would be lower than the control group mean

    C. If the experiment were conducted on the entire population, the treatment and control group means would both equal 5 (the midpoint on the scale)

    D. The percentage of treatment and control group participants with significant pain is the same in the population

21. A clinical psychologist is studying whether males or females have better outcomes when trying to quit smoking. Using an established diary measure, participants in both groups record the number of cigarettes smoked during a specified period of time during the intervention. Because the two measurements are obtained from different samples or people, the researcher uses an independent-samples t test to analyze the data. The analysis results are as follows: t = -0.893, df = 12.963, p-value = 0.3881. Based on the probability value, which is the conclusion about the null hypothesis?
    a.   The data support (i.e., are consistent with) the null
22. Researchers are interested in determining whether a smoking cessation intervention reduces smoking frequency. They assign 50 subjects to the treatment condition and 50 to the control, and measure smoking frequency on a 1-5 scale. At the end of the study the researcher finds that the treatment subjects improved, with a standardized mean difference effect size of .51. Which of the following interpretations is correct?
    a.   The mean difference between the treatment and control group is equivalent to about a half a standard deviation
23. A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. She randomly assigns 30 subjects to the treatment and 30 subjects to the control. At the conclusion of the study she uses a t-test to evaluate whether the groups differ on a widely used depression scale that ranges between 0 and 30. The treatment group (group 1) mean is 18, the control group (group 2) mean is 22, and the standard error is 2. What is the value for the t-statistic?
    a.   -2
24. A cognitive psychologist is comparing ADHD children to a sample of non-ADHD controls using an established measure of reaction time that captures attention span. The mean difference is 150 (the ADHD

group mean is higher by 150 points), with a confidence interval for the mean difference ranging between 25 and 350. Which of the following population parameters are supported by the data? That is, if the following values were used as the null hypothesis, which value(s) would be supported by the data? CHECK ALL THAT APPLY.

    a.   50

25. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. At the end of the study the researcher administered a widely used pain level scale to the intervention participants and a control group that received generic reading materials. Which of the following sets of means provides the strongest evidence against the null hypothesis?

    a.   Intervention mean = 7, control mean = 8.50

26. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. A group of pain patients are randomly assigned to participate in either the online program or a control group. At the end of the study the researcher administered a widely used pain level scale ranging from 1-10. The two-tailed probability value for the independent t-statistic is .67. Which of the following conclusions is definitely true?

    a.   The treatment and control group means were essentially identical

27. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. A group of pain patients are randomly assigned to participate in either the online program or a control group. At the end of the study the researcher administered a widely used pain level scale ranging from 1-10. The treatment group (group 1) mean is 8, the control group (group 2) mean is 7, and the standard error for the analysis is .50. What is the value for the t-statistic?

    a.   2

28. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. A group of pain patients are randomly assigned to participate in either the online program or a control group. At the end of the study the researcher administered a widely used pain level scale ranging from 1-10. The standard error for the analysis is .50. Which of the following interpretations is correct regarding the standard error?

    a.   Due to estimation error (random chance) alone, we would expect a .50 mean difference on a 10-point scale

29. A clinical psychologist is studying whether a call-in help line or individual therapy leads to better outcomes for individuals who are trying to quit smoking. Using an established diary measure, participants in both groups record the number of cigarettes smoked during a specified period of time during the intervention. Because the two measurements are obtained from different samples or people, the researcher uses an independent-samples t test to analyze the data. The analysis results are as follows: t = -1.2792, df = 7.4462, p-value = 0.2392. Based on the probability value, which of the following interpretations is likely true?

    a.   The quit line and therapy produce identical outcomes

30. A clinical psychologist is studying whether individual therapy leads to better outcomes for individuals who are trying to quit smoking relative to a control group. Using an established diary measure, participants in both groups record the number of cigarettes smoked during a specified period of time during the intervention. Because the two measurements are obtained from different samples or people, the researcher uses an independent-samples t test to analyze the data. The analysis results are as follows: t = 2.3534, df = 7.9529, p-value = 0.04662. Based on the probability value, which is the conclusion about the null hypothesis?

    a.   The data refute (i.e., are inconsistent with) the null

31. A clinical psychologist is studying whether a call-in help line or individual therapy leads to better outcomes for individuals who are trying to quit smoking. Using an established diary measure, participants in both groups record the number of cigarettes smoked during a specified period of time during the intervention. Because the two measurements are obtained from different samples or people, the researcher uses an independent-samples t test to analyze the data. The analysis results are as follows: t = -1.2792, df = 7.4462, p-value = 0.2392. Based on the probability value, which is the conclusion about the null hypothesis?

    a.   The data support (i.e., are consistent with) the null

32. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. He assigns 20 subjects to the treatment condition and 20 to the control, and he measures pain intensity on a 1-10 scale. At the end of the study the researcher finds an improvement, the standardized mean difference effect size of which is .30. Which of the following describes the meaning of this effect size value?

a. Pain scores improved by about a third of a standard deviation
33. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. At the end of the study the researcher administered a widely used pain level scale to the intervention participants and a control group that received generic reading materials. Which of the following sets of means provides the strongest support for the null hypothesis?
   a. Intervention mean = 7, control mean = 7.25
34. A clinical psychologist is studying whether individual therapy leads to better outcomes for individuals who are trying to quit smoking relative to a control group. Using an established diary measure, participants in both groups record the number of cigarettes smoked during a specified period of time during the intervention. Because the two measurements are obtained from different samples or people, the researcher uses an independent-samples t test to analyze the data. The analysis results are as follows: t = 2.3534, df = 7.9529, p-value = 0.04662. Based on the probability value and the estimates, which of the following interpretations is most likely correct?
   a. The treatment condition produced a significant reduction in the number of cigarettes smoked
35. A clinical psychologist is studying the association between gender and aggression in young children. Using an established observational measure, she records the number of aggressive acts committed by boys and girls during a free play period at school. Because the two measurements are obtained from different samples or people, she uses an independent-samples t test to analyze the data. The difference in the number of aggressive behaviors was 1.3 points higher for males, and the t test gave a probability value of .13 (13%). Which of the following interpretations is correct regarding the probability value?
   a. If there is truly no difference in the population, a difference of 1.3 or more on the aggression scale would occur in 13% of all random samples from such a population
36. A clinical psychologist is studying the association between gender and aggression in young children. Using an established observational measure, she records the number of aggressive acts committed by boys and girls during a free play period at school. Because the two measurements are obtained from different samples or people, she uses an independent-samples t test to analyze the data. The difference in the number of aggressive behaviors was 1.3 points higher for males, and the t test gave a probability value of .13 (13%). Which of the following interpretations is correct regarding these results?
   a. Boys and girls are the same in their aggression behavior
37. A researcher wants to study the effectiveness of a depression intervention. Which of the following research designs would require an INDEPENDENT t-test as the analysis?
   a. Recruit a sample of 50 depressed individuals, assign half to receive the intervention, and assign the other half to receive a placebo. Compare the average depression score between the two groups.
38. A clinical psychologist is studying the association between playing violent video games and aggression in young children. Using an established observational measure, she records the number of aggressive acts committed by children after playing non-violent games, and does the same for children who have played violent video games. Because the two measurements are obtained from different samples or people, she uses an independent-samples t test to analyze the data. The analysis results are as follows: t = -2.7116, df = 15.391, p-value = 0.0158. Based on the probability value, which is the conclusion about the null hypothesis?
   a. The data refute (i.e., are inconsistent with) the null
39. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. She randomly assigns 50 subjects to the online intervention and 50 subjects to a waitlist control. At the conclusion of the study she uses a t-test to evaluate whether the groups differ. The independent t-statistic is 2.50, and its two-tailed probability value is .0141. Which of the following conclusions is true regarding the treatment?
   a. The treatment improves scores beyond what is expected due to random chance (estimation error)
40. A psychologist studying chronic pain reports that his new treatment protocol produced a "statistically significant" difference between the treatment and control group. Which of the following conclusions is most likely true?
   a. The treatment mean was noticeably lower (better) than the control group mean
41. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. At the end of the study the researcher administered a widely used pain level scale to the intervention participants and a control group that received generic reading materials. Which of the following sets of means provides the strongest support for the null hypothesis?

a. Intervention mean = 7, control mean = 8.50

12. Anova

1. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. A group of pain patients are randomly assigned to participate in either the online program or a control group. At the end of the study the researcher administered a widely used pain level scale ranging from 1-10. What is the null hypothesis for the study?
   a. The correct answer is: If the experiment were conducted on the entire population, the difference in the treatment and control group means would be zero
2. **A psychologist is interested in determining whether a new treatment for depression improves (decreases) depression scores. To evaluate the treatment, she randomly assigns 30 subjects to the treatment and 30 subjects to the control, and she then performs an independent t-test. What population parameter is the target/focus of the hypothesis?**
   a. The difference between the treatment and control group means
3. A psychologist is interested in determining whether a new online pain management program improves (reduces) pain scores for its participants. He assigns 20 subjects to the treatment condition and 20 to the control. At the end of the study the researcher finds a standardized mean difference effect size of .30. Based on widely accepted rules of thumb, how would you characterize the magnitude of this treatment effect.
   a. Small
4. Suppose an educational researcher conducts an ANOVA analysis that compares standardized test scores for four different ethnic groups: African American, Hispanic, Caucasian, and Asian. What is the null hypothesis for the ANOVA analysis?
   a. All four groups have identical mean achievement levels
5. Suppose an educational researcher conducts an ANOVA analysis that compares standardized test scores for four different ethnic groups: African American, Hispanic, Caucasian, and Asian. The probability value for the F statistic is $p = .12$. Which of the following conclusions is correct?
   a. All four groups have identical mean achievement levels, within sampling error
6. The following table of descriptive statistics displays the means and standard deviations from a study that compares three different intervention conditions designed to change attitudes toward vaccinations. Which values from the table contribute to or define between-group variability in an ANOVA analysis (i.e., variation explained by the independent variable)?
   a. Means only
7. Consider the following sets of graphs that show the score distributions for two groups being compared in an ANOVA analysis. Which graph/graphs correspond to the ANOVA with the SMALLEST between-group variability (i.e., mean difference, variation due to the model or independent variable)?
   a. Graphs A and B
8. Consider the following sets of graphs that show the score distributions for two groups being compared in an ANOVA analysis. Which graph/graphs correspond to the ANOVA with the LARGEST residual variability (i.e., error, or variation NOT explained by the independent variable)?

9. Suppose a researcher conducts an ANOVA analysis that compares three different treatments for depression: medication, cognitive behavioral therapy, and a control condition. The probability value for the F statistic is p = .01. Which of the following steps should the researcher take next?
    a.   Conduct t-tests that compare whether every possible pair of groups differs in their mean depression level

Suppose an educational researcher conducts an ANOVA analysis that compares standardized test scores for four different ethnic groups: African American, Hispanic, Caucasian, and Asian. Assuming that the F statistic is significant, how many different post hoc (pairwise) comparisons are possible for this analysis?

   a. 6

Suppose a researcher conducts an ANOVA analysis that compares three different treatments for depression: medication, cognitive behavioral therapy, and a control condition. The ANOVA analysis indicates that the between-group variability (i.e., variation due to the model or the independent variable) is greater than zero. Which of the following must be true in this case?

   a. The three groups have different mean depression scores

Suppose a researcher conducts an ANOVA analysis that compares three different treatments for depression: medication, cognitive behavioral therapy, and a control condition. Assuming that the F statistic is significant, how many different post hoc (pairwise) comparisons are possible for this analysis?

   a. 3

The following table of descriptive statistics displays the means and standard deviations from a study that compares three different intervention conditions designed to change attitudes toward vaccinations. Which values from the table contribute to or define residual variability in an ANOVA analysis (i.e., error, variation not explained by the model or the independent variable)?

   a. The standard deviations: 1.00, 1.40, and .90

To investigate test bias, an educational researcher conducts an ANOVA analysis that compares standardized test scores for four different ethnic groups: African American, Hispanic, Caucasian, and Asian. What type of evidence would be needed to refute the null (i.e., make the null look "absurd")?

   a. At least one of the four groups has different mean achievement levels

Consider the following sets of graphs that show the score distributions for two groups being compared in an ANOVA analysis. Which graph/graphs correspond to the ANOVA with the SMALLEST residual variability (i.e., error, or variation not explained by the model or independent variable)?

a. Graphs A and C

The following table of descriptive statistics displays the means and standard deviations from a study that compares three different intervention conditions designed to change depression levels. Which values from the table contribute to or define between-group variability in an ANOVA analysis (i.e., variation explained by the independent variable)?

a. The means: 11.00, 12.00, 15.00

Suppose an educational researcher conducts an ANOVA analysis that compares standardized test scores for four different ethnic groups: African American, Hispanic, Caucasian, and Asian. The probability value for the F statistic is $p = .12$. Which of the following steps should the researcher take next?

a. Stop, no further analysis is required

Suppose an educational researcher conducts an ANOVA analysis that compares standardized test scores for four different ethnic groups: African American, Hispanic, Caucasian, and Asian. The probability value for the F statistic is $p = .02$. Which of the following conclusions is correct?

a. At least one of the four groups has different mean achievement levels

Suppose an educational researcher conducts an ANOVA analysis that compares standardized test scores for four different ethnic groups: African American, Hispanic, Caucasian, and Asian. The probability value for the F statistic is $p = .02$. Which of the following steps should the researcher take next?

a. Conduct t-tests that compare whether every possible pair of groups differs in their mean achievement level

Suppose a researcher conducts an ANOVA analysis that compares three different treatments for depression: medication, cognitive behavioral therapy, and a control condition. The probability value for the F statistic is $p = .25$. Which of the following steps should the researcher take next?

a. Stop, no further analysis is required

Suppose a researcher conducts an ANOVA analysis that compares three different treatments for depression: medication, cognitive behavioral therapy, and a control condition. The ANOVA analysis indicates that the between-group variability (i.e., variation due to the model, or the independent variable) is exactly zero. Which of the following must be true in this case?

    a. The three groups have identical mean depression scores

The following table of descriptive statistics displays the means and standard deviations from a study that compares three different intervention conditions designed to change depression levels. Which values from the table contribute to or define residual variability in an ANOVA analysis (i.e., error, variation not explained by the model or the independent variable)?
Descriptive Statistics

Condition	Group Mean	Group Std. Dev.	Group N
Medication	11	2.5	50
Cognitive Therapy	12	3.0	50
Control Group	15	3.5	50

a.The standard deviations: 2.50, 3.00, and 3.50

Class Prep Quiz #12

1.   -Suppose a researcher conducts an ANOVA analysis that compares three different treatments for depression: medication, cognitive behavioral therapy, and a control condition. The probability value for the F statistic is $p = .25$, which is non-significant. Which of the following steps should the researcher take next?
Stop, no further analysis is required

2.   -Which of the following best characterizes the difference between an independent t-test and an ANOVA analysis?
ANOVA can accommodate 2 or more independent variable groups, the t-test can compare only 2 groups

3.   -Suppose a researcher conducts an ANOVA analysis that compares three different treatments for depression: medication, cognitive behavioral therapy, and a control condition. The p-value for the F statistic is .04. Which of the following statements is true?
At least one pair of treatment conditions has different means

4.   -Which of the following best describes the construction of the F statistic?
A fraction reflecting variation explained by the independent variable relative to unexplained (error) variation

5.   Suppose a researcher conducts an ANOVA analysis that compares three different treatments for depression: medication, cognitive behavioral therapy, and a control condition. The probability value for the F statistic is $p = .01$, which is statistically significant. Which of the following steps should the researcher take next?
Conduct t-tests that compare whether every possible pair of groups differs in their mean depression level

6.   In an ANOVA analyses, the independents variables are _____ and the dependent variable is _____.
Categorical, continuous

7.   -Suppose a researcher conducts an ANOVA analysis that compares three different treatments for depression: medication, cognitive behavioral therapy, and a control condition. What evidence is needed from the data in order to refute the null hypothesis (i.e., make the null look "absurd").
At least one of the three groups has a different mean depression level than the others

8. -Consider the following sets of graphs that show the score distributions for two groups being compared in an ANOVA analysis. Which graph/graphs correspond to the ANOVA with the LARGEST between-group variability (i.e., mean difference, variation explained by the model or independent variable)?
   Graph C (the graph with the peaks furthest apart)

9. -Suppose a researcher conducts an ANOVA analysis that compares three different treatments for depression: medication, cognitive, behavioral therapy, and a control condition. What is the null hypothesis for the ANOVA analysis?
   All three groups have identical mean depression levels

10. -Suppose a researcher conducts an ANOVA analysis that compares three different treatments for depression: medication, cognitive behavioral therapy, and a control condition. The p-value for the F statistic is .25. Which of the following statements is true?
    All three groups have identical means, within sampling error

Assignment 4

1. What is the null hypothesis for the analysis?
A

2. Enter the value from the printout that tells you how much scores changed from pretest to posttest on the 0 to 20 metric of data._____
1.377
I said -1.377 because it's mostly it went down, posttest-pretest

3. The t statistic value is 4.637 (or -4.637 depending on how you entered the variables). Which of the following interpretations is correct for the t statistic?
C --- I also put C ← i did too, because there was a problem similar to this on content quiz 9 but I don't know why D wouldn't be right too?
I put D ← same, I did as well
Which did majority put??^^^

4. Based on the probability value, which of the following is true regarding the null hypothesis?
B

5. Based on your response to the previous question, which of the following interpretations is correct?
A

6. Considering the value of the standardized mean difference, which of the following interpretations is correct.
B

1. What is the null hypothesis for the analysis?

A

2. By how much did the treatment group means differ on the pain interference scale? _____

1.184

3. Which value of the printout tells you the expected size of the mean difference between two groups of participants drawn from a population with equal groups (i.e., the mean difference due to sampling error or "random chance").

_____

0.1856

4. The t statistic value is 6.38. Which of the following interpretations is correct for the t statistic?

C

5. The two-tailed probability value for the t-test prints is < .001 (less than 1 in 1000, or less than one-tenth of one percent, .10%). Which of the following interpretations is correct?

C

6. Based on the probability value, which of the following is true regarding the null hypothesis?

B

7. Based on your response to the previous question, which of the following interpretations is correct?

A

8. In addition to determining whether the treatment has a "statistically significant" impact on pain interference, researchers also want to determine the strength of the association. Which of the following interpretations is correct for the Cohen's d statistic (i.e., the standardized mean difference)?

B

9. Considering the value of the standardized mean difference, which of the following interpretations is correct (do not round the effect size).

C

Final Exam

Section 1

    1.   What is the null hypothesis for the analysis?

There is no change in illness severity between the pre-test and six-week follow up scores in the population.

2. What type of research design best describes this part of the study, where the goal is to examine change in illness severity scores between the two assessments?

Within subjects design
Edit: Wouldn't it be between subjects design since different people are getting subjected to different interventions and not going through multiple?
-Participants are assigned to one of two treatments, and their illness severity scores are measured at different stages of the treatment.
I think it's still within since we're comparing the scores of the same people before and after six weeks; participants are separated into two groups (week 0 and week 6) instead of by the treatment condition ← same, especially since the next one is definitely a between-subjects design

3. Enter the value from the printout that tells you how much scores changed from pretest to posttest on the 7-point illness severity rating scale.

2.021 is it 2.021 or -2.021

4. Consider the t-statistic from the analysis. What is the appropriate interpretation for this value? Note that the value could be positive or negative depending on how you selected the variables for the analysis.

T-statistic: -24.4 (our observed mean change is -24.4 standard error units higher than the null hypothesis change of 0--- OR could also say the difference is -24.4 times as large as what we would expect from sampling error alone)

5. Consider the standard error of the mean difference. What is the appropriate interpretation for this value?

SE difference: 0.0829

On average a sample size of 100 would give an estimate that differs from the null hypothesis by about 0.0829

- Why 100? Shouldn't it be 437? yeah i put 437
6. What can you conclude about statistical significance?

The observed p-value is smaller than .001 therefore our p-value is less than 0.05, meaning that our results are significant. We can reject the null hypothesis.

7. Consider the two-tailed probability value from the analysis. What is the appropriate interpretation for this value?

??? If the two-tailed probability value is 0.001, then our results are statistically significant (different from the null) because there is less than a 1% probability that the null is correct(?) Can someone else confirm? I got the same coolio!

This is what I put since I thought they were asking for an interpretation of what the p-value means: If there truly is no difference between the two groups (null hypothesis is true), then the probability of producing a t-statistic as or more extreme than 2.02059 is less than 0.001. ← same ← same!

8. The two-tailed probability value for the analysis is .001. What is the probability value for a one-tailed test that predicts a decrease in illness severity scores (the outcome observed in this analysis)?

??? Wouldn't this just be half of the two-tailed probability, so 0.0005? Can someone else confirm? yeah i got 0.0005 dope! Ye same!

9. Considering the entirety of the analysis results, what practical conclusion can you draw about the change in symptom severity across the duration of the study?

The symptom severity significantly decreased across the duration of the study

10. Considering the value of the standardized mean difference. How would you characterize the magnitude of the change between the pre-test and six-week follow-up?

Cohen's d equates this mean change to (-) 1.17 standard deviation units. The magnitude of this change is large.

Section 2

11. What is the null hypothesis for the t-test analysis?

Males and females do not differ in their depression scale post scores

12. What type of research design best describes this part of the study, where the goal is to examine differences between two distinct groups of persons?

Between subjects design

13. By how much did the group means differ on the 7-point illness severity scale?

-0.0628

14. Which value from the printout expresses the magnitude of the observed gender difference in the data relative to the expected mean difference due to sampling error alone (i.e., the difference that should result from drawing two samples from the population)?

SE difference: 0.155 ← i got the same ← same

-0.404 (t-statistic; it's asking for the magnitude of the observed difference RELATIVE to the SE, which is the definition of a test statistic)<-- I agree ← same!

15. What can you conclude about statistical significance?

P = 0.686 therefore, because this value is not below 0.05; not statistically significant

16. Considering your entire analysis results, what practical conclusion can you draw about the gender difference in illness severity scores at the end of the study?

The gender difference in illness severity scores at the end of the study are not significant. There is not a significant difference between male and female end severity scores.

17. Which value on the printout tells you the expected size of the mean difference that would result from comparing two random samples of 230 and 207 participants?

??? Would it be just the mean difference: -0.0628

SE difference: 0.155 ← same

18. Consider the two-tailed probability value from the analysis. What is the appropriate interpretation for this value?

P = 0.686. If there is truly no difference between the posttest means of males and females, results like these have a 68.6% chance of occurring.

19. In addition to determining whether there is a "statistically significant" gender difference, researchers also want to determine the strength of the association. What is the interpretation of the Cohen's *d* statistic (i.e., the standardized mean difference) from the analysis? Note that this question is not asking how large the effect is, but rather what its numeric value means.

??? Cohen's d: -0.0387 The means are 0.0387 standard deviations apart. ← I agree; "The male and female group averages differ by an amount equal to 0.0387 standard deviation units"

20. Considering the value of the standardized mean difference, which of the following interpretations is correct (do not round the effect size).

very small, nil effect size ← I got small-medium size ← I got negligible/nil effect too

Section 3

21. What is the null hypothesis for the ANOVA analysis?

There is no difference between the means of each group.

22. What type of evidence from the data (i.e., pattern of mean differences) are needed to reject the ANOVA null hypothesis and conclude that the F statistic is statistically significant?

At least one of the three groups has a different mean reduction in illness severity at the end of the study than the others. The F statistic is significantly larger than 1.

23. Which value from the ANOVA printout captures the magnitude of the mean differences among the three drug conditions on the variance (i.e., average squared distance) metric?

105.79

24. Which value from the ANOVA printout captures just the magnitude of the naturally occurring score differences on the variance (i.e., average squared distance) metric?

2.15

25. Consider the value of the residual mean square, the value of which is about 2.15. Which values from the descriptive statistics table influence the magnitude of this value?

Standard deviation

-Isn't it degrees of freedom (df) since to get the mean square value, you divide the sum of the squares by the degrees of freedom?

I put sum of squares and df, since residual mean square = {sum of squares}/{df}

Although these values aren't on the table labelled "descriptives." It's possible that the answer is standard deviation and N, since sum of squares of residuals is stdev squared and sample size determines df.

26. Consider the means of the three groups: placebo = 4.76, 200 mg = 3.24, 400 mg = 3.08. Suppose the mean of the placebo group changed from 4.76 to 5.00 (i.e., this group became more different from the others). How would the F statistic change, if at all?

The F statistic will increase.

27. The F statistic is 49.24, and its corresponding probability value is .001 (i.e., 1 out of 1000). What is the interpretation of this the probability value?

If the population means for the three treatment conditions are truly the same, the probability of having three sample means that produce an F statistic of 49.24 is less than .001.

28. What number on the printout represents the ratio (fraction) of variation in illness severity scores explained by treatment group membership (i.e., the group means) relative to naturally occurring scores differences among individuals?

49.24

29. The ANOVA F statistic is statistically significant. What is the next step that you should take (i.e., what analysis would you perform next, if any) after finding a significant F statistic?

Conduct t-tests that compare whether every possible pair of groups differs in their mean reduction in illness severity

30. Considering the results of the post hoc analyses and the group means, what are the practical conclusions regarding which groups produced a change in illness severity?

The 200 mg group and placebo group as well as the 400 mg group and placebo group produced a significant change in illness severity scores.

Also no difference in effect between a 200 mg and 400 mg dose treatment

31. Based on the $R2$ statistic (labeled as η! in jamovi), how would you characterize the magnitude of the mean differences among the three groups?

The magnitude of the mean differences among the three groups is large.

32. What is the interpretation of the $R2$ statistic from the printout (labeled as η! in jamovi)? Note that this question is not asking how large the effect is, but rather what its numeric value means.

18.49% of the variance was caused by the drug condition.

^^not sure of section 3 answers

13. Correlation

1.   With a positive correlation, as scores on one variable increase, scores on the other variable _____.
**Increase**

2.   A researcher investigating the relationship between job performance and IQ finds a correlation of .55. Which of the following interpretations is correct?
Large

3. A researcher investigating the relationship between smoking frequency and self-efficacy (confidence) to quit finds a correlation of -.25. Which of the following best describes the magnitude of this correlation?
small

4. Consider the following scatterplot. Which of the following values best represents the correlation between the two variables?
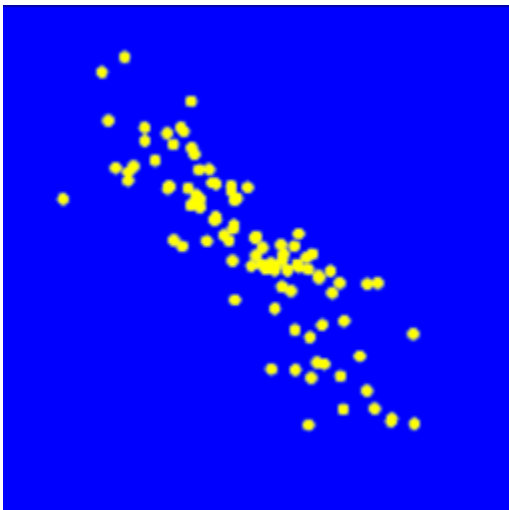
-.79

5. Which of the following pairs of variables would yield a negative correlation? Check all that apply.
**b. Chronic pain level and and life happiness**
c. Marital satisfaction and frequency of arguments with spouse

6. A researcher finds a correlation of .55 between one pair of variables, and she finds a correlation of -.55 for another pair of variables. Which correlation represents the strongest association between the two variables?

The strength of the relationship is the same for both correlations

6. Consider the following scatterplot. Which of the following values best represents the correlation between the two variables?
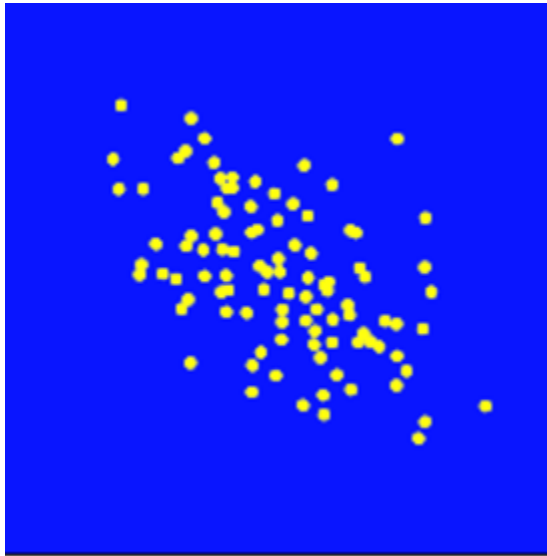
-.51

8. Which of the following pairs of variables would yield a negative correlation? Check all that apply.
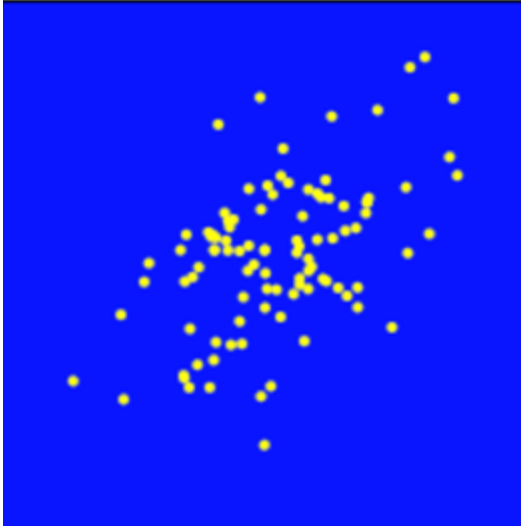**b. Hours per week playing video games and GPA**
d. Weight of a car and miles per gallon

9. A researcher investigating the relationship between job performance and IQ finds a correlation of .55. Which of the following interpretations is correct?
People with low performance ratings tend to have lower levels of IQ

10. Consider the following scatterplot. Which of the following values best represents the correlation between the two variables?

0.51