STATS-10 Midterm Practice Problems

1. Marital status of each member of a randomly selected group of adults is an example of what type of variable?
   A) Categorical variable B) Numerical variable C) Neither

2. The average number of hours spent completing statistics homework for a randomly selected group of statistics students is an example of what type of variable.
   A) Categorical variable B) Numerical variable C) Neither

3. A recent report showed there were 49 accidents involving pedestrians in City A and 62 accidents involving pedestrians in City B this year. The mayor of City A claims that his city is safer for pedestrians than City B. What information is missing that might contradict this claim?
   A) The total number of pedestrians in both City A and City B
   B) The number of crosswalks in both City A and City B
   C) The number of accidents that do not involve pedestrians in both City A and City B
   D) The number of accidents involving pedestrians from the previous year

4. A group of students is divided into two groups. One group listens to classical music while taking a math test and the other group takes the test in silence. The average test scores of the two groups are compared to see whether listening to music during a math test has an effect on scores.
   A) Observational study B) Controlled experiment

5. Consider the following statement: "Researchers conducted a large observational study and determined that children who participated in school music programs scored higher on math exams in later grades than those who did not." Suppose that upon hearing this a politician states that all children should participate in school music programs. What is wrong with the politician's statement?
   A) The politician confused correlation with causation.
   B) The controlled experiment was not double-blinded.
   C) There was a placebo effect.
   D) This study exhibits bias.

6. The distribution of marital status of a randomly selected group of adults would best be visualized by which plot?
   A) Boxplot B) Histogram C) Scatterplot D) Bar chart

7. The gender and handedness of students in a class are recorded in the table below

	Left	Right
Male	10	40
Female	5	35

What percent of the class is left-handed?

A) 5% B) 16.7% C) 20% D) 66.7%

8. Which of the following is (are) key features of a well-designed experiment? (Select all that apply)

A) The population should be large enough to observe the full range of variability
B) The subjects of the study should be assigned to the groups randomly
C) The study should use a placebo if possible
D) The researchers should know who is in which group

9. Use the following information to answer the following questions.

A group of researchers is interested in whether UCLA students would generally consider themselves happy. They waited outside of each entrance of Ackerman Union after it had been selected at random from a list of all the buildings at UCLA. They stopped every 10th person who walked into the building and did the following:

1) asked if the person was a UCLA student.
2) if the person was a UCLA student, then researchers asked whether the person was happy and recorded the student's answer and major.
3) If the person was not a UCLA student, the researchers thanked the person for stopping and no other questions were asked.

They eventually recorded the responses of 100 UCLA students and discovered that 85 said they were "happy."

i. This is called a(n)

A) Anecdotal study B) Observational study
C) Experimental study D) None of these

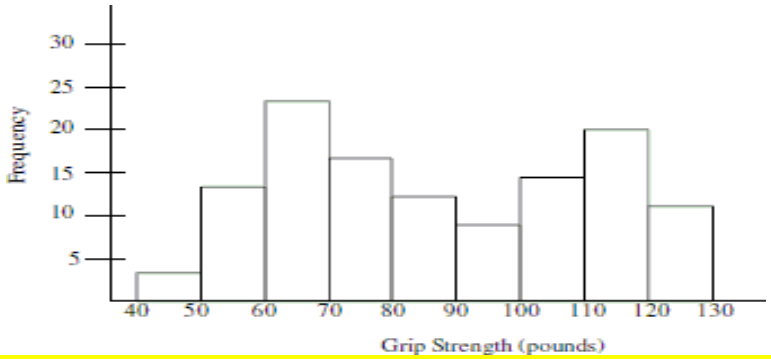
ii. The population of interest to the researchers is

A) All UCLA students
B) All UCLA students who go into Ackerman Union
C) All persons who go into Ackerman Union
D) All persons on campus

iii. The sample is

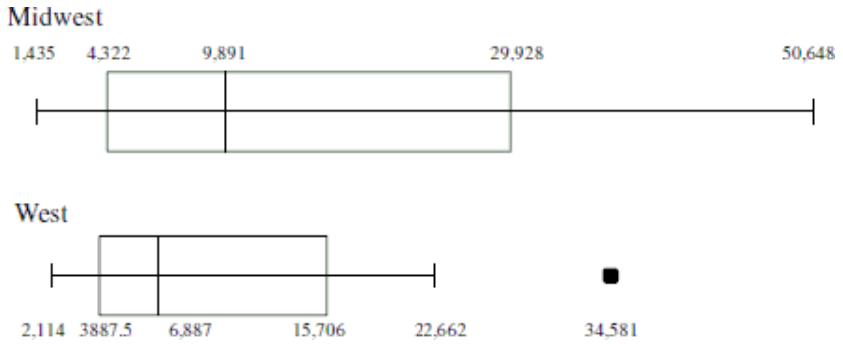
A) 100 persons who went into Ackerman Union
B) 85 happy students
C) 100 UCLA students who went into Ackerman Union
D) All persons who talked to the researchers.

10. Data was collected on hand grip strength of adults. The histogram below summarizes the data. Which statement is true about the distribution of the data shown in the graph?



A) The graph shows evidence that two different groups may have been combined into one data set.
B) The graph is useless because it is bimodal.
C) The best estimate of typical grip strength is 80-90 pounds because it is in the center of the distribution.
D) There must have been a mistake made in data set because the distribution should be bell-shaped.

11. Use the side-by-side boxplots below to answer the following two questions.
The boxplots summarize the number of sentenced prisoners by state in the Midwest and West.



i. Pick the statement that best describes the shape of the distribution for the states in the West.
A) The data appears to be left-skewed with large variability.
B) The data appears to be right-skewed with a possible outlier.
C) The data appears to be roughly symmetrical with a possible outlier.

ii. Using the boxplot for the Midwest, determine which of the following statements about the distribution cannot be justified.
A) About 75% of the Midwest states had 4,322 or more prisoners.
B) There are much more states with 6887 to 15,706 prisoners than states with 3887.5 to 6887 prisoners.
C) The interquartile range of West states is about 11,819.
D) The range of the boxplot for the West states is 32,467.

12. Which of the following can be calculated from a dotplot?
    A) Range
    B) IQR
    C) Mode
    D) Mean
    E) All of the above

13. A large university conducted a survey among their students and received 300 responses. The survey asked the students for the following information: Age, Year in school (Freshman, Sophomore, Junior, Senior, Graduate student), GPA, Gender.
    What type of graph would you use to describe the variables Gender and Year in school?
    A) A side-by-side histogram should be used since these are two numerical variables
    B) A side-by-side bar chart should be used since these are two categorical variables
    C) A boxplot should be used since these are two categorical variables
    D) A scatterplot should be used since these are two numerical variables

14. The mean price of a pound of ground beef in 75 cities in the Midwest is $2.11 and the standard deviation is $0.56. Suppose the histogram of the data shows that the distribution is unimodal and symmetrical. A local grocer is selling a pound of ground beef for $3.50. What is this price in standard units? Assuming the Empirical Rule applies, would this price be unusual or not? Round to the nearest hundredth.
    A) z = 2.48; This is unusually expensive ground beef.
    B) z = -2.48; This price would not be unusual.
    C) z = 2.48; This price would not be unusual.
    D) z = -2.48; This is unusually inexpensive ground beef.

15. Use the five-number summaries below to answer the following four questions.
    A group of elementary school students is given a reading test and the scores are reported by reading grade level. The five-number
    Five number summaries

	Min -- Q1—Q2 -- Q3 -- Max
Boys:	2.0 -- 3.9 -- 4.3 -- 4.9 -- 6.0
Girls:	2.8 -- 3.8 -- 4.5 -- 5.2 -- 5.9

i. Which group had the highest score?
    A) Boys B) Girls C) They are equal

ii. Which group had the greater range?
    A) Boys B) Girls C) They are equal

iii. Which group had the greater interquartile range?
    A) Boys B) Girls C) They are equal

iv. Which group generally performed better on this test?
    A) Boys B) Girls (C)The groups performed equally well

16. The number of students enrolled in a college algebra class for the last seven semesters are listed here:

Find the median: 60 61 55 57 64 58 58.

A) 57 B) 58 C) 59 D) 60

17. Which one of the following best describes the relationship between the correlation and the slope of the regression line modeling the relationship between X and Y?

A) The correlation between X and Y equals the slope of the regression line modeling the relationship between X and Y.

B) The sign of the correlation between X and Y is the same as the sign of the slope of the regression line modeling the relationship between X and Y.

C) When the correlation between X and Y is zero, the slope of the regression line modeling the relationship between X and Y is negative.

D) The correlation between X and Y is not related to the slope of the regression line modeling the relationship between X and Y.

18. The following linear regression model can be used to predict ticket sales at a popular water park.

Predicted ticket sales per hour = -631.25 + 11.25 * (current temperature in °F)

i. What is the predicted number of tickets sold per hour if the temperature is 86°F? Round to the nearest whole ticket.

A) About 336 tickets B) About 252 tickets C) About 276 tickets D) About 301 tickets

ii. Choose the statement that best states the meaning of the slope in this context.

A) The slope tells us that if ticket sales are decreasing there must have been a drop in temperature.

B) The slope tells us that a one degree increase in temperature is associated with an average increase in ticket sales of 11.25 tickets.

C) The slope tells us that high temperatures are causing more people to buy tickets to the water park.

D) None of the above

19. A horticulturist conducted an experiment on 110 thirty-six-inch plant boxes to see if the amount of plant food given to the plant boxes was associated with the number of tomatoes harvested from the plants. The mean amount of plant food given was 27.8 milliliters with a standard deviation of 2.1 milliliters. The mean number of tomatoes harvested was 7.5 with a standard deviation of 1.5. The correlation coefficient was 0.7691. Use the information given to calculate the slope of the linear model that predicts the number of tomatoes harvested from the amount of plant food given. Round to the nearest hundredth.

A) -7.50

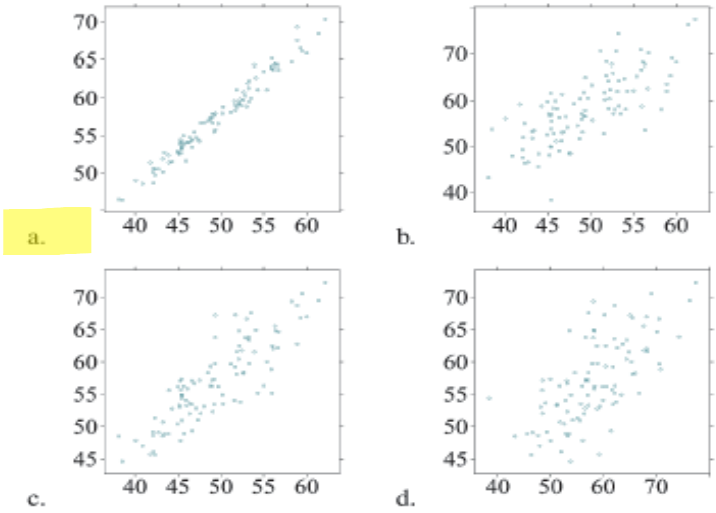
B) 0.55

C) 1.08

D) The slope cannot be determined without the actual data.

20. Which of the following statements regarding the correlation coefficient is true?
   A) The correlation coefficient is a non-negative value
   B) A high correlation tells us the data is linear
   C) A correlation coefficient of 0 means that the two variables have no relationship at all
   D) A correlation coefficient of -1 means that as one variable increases, the other decreases

21. Which of the following scatterplots shows data with the highest correlation between the explanatory and response variables?



22. In the NBA, the correlation between "steals per game" and "blocked shots per game" is found to be 0.8045. Choose the statement that is true about the coefficient of determination.
   A) The coefficient of determination, r2, is equal to approximately 0.6472.
   B) The coefficient of determination states that about 64.72% of the variation in the blocked shots per game is explained by steals per game.
   C) When given as a percent, the coefficient of determination is always between 0 and 100%.
   D) All of the above are true statements.

23. If two numerical variables X and Y have a correlation coefficient of 0.90, what percentage of the variation in one variable can be accounted for by the other variable?
   A) 45%
   B) 81%
   C) 90%
   D) The answer will depend on which one you designate as the response variable

24. Is the following an example of theoretical probability or empirical probability?
   A homeowner notes that five out of seven days the newspaper arrives before 5pm. He concludes that the probability that the newspaper will arrive before 5pm tomorrow is about 71%.

A) Theoretical B) Empirical

25. A study found that highly experienced teachers may be associated with higher student achievement. Suppose fourth-grade students at an elementary school are randomly assigned to one of eight teachers. Teachers Williams, Crouse, Perry, and Martinez are considered highly experienced. Teachers Tran, Lee, Cochran, and King are considered less experienced.
    i. What is the probability that a fourth-grade student at this school is assigned to a highly experienced teacher?
        A) 40%
        B) 20%
        C) 80%
        D) 50%

    ii. What event is the complement of the event described in part a? What is the probability of this event? Select the correct choice below and fill in the answer box to complete your choice
        A) The complement is the probability that a fourth-grade student at this school is assigned to a less experienced teacher.
        B) The complement is the probability that a fourth-grade student at this school is assigned to a highly experienced teacher.
        C) The complement is the probability that a fourth-grade student at this school is assigned to any teacher.

26. A multiple-choice test has six possible answers, only two of which is correct, for each question. (Type an integer or a simplified fraction)
    i. The probability of guessing correctly on one question is 1/3.

    ii. The probability of guessing incorrectly on one question is 2/3.

27. The sample space given here shows all possible sequences for tossing a fair coin 4 times. The sequences have been organized by the number of tails in the sequence.

0 Tails	1 Tail	2 Tails	3 Tails	4 Tails
HHHH	THHH	TTHH	HTTT	TTTT
	HTHH	THTH	THTT	
	HHTH	THHT	TTHT	
	HHHT	HTTH	TTTH	
		HTHT		
		HHTT		

i. How many outcomes are in the sample space?
16

Assuming all of the outcomes in the sample space are equally likely, find the following probabilities. (Type an integer or a simplified fraction)

ii. Probability of 0 tails in 4 tosses 1/16 .

iii. Probability of only 3 heads in 4 tosses: 1/4 .

iv. Probability of at most 3 heads in 4 tosses 15/16 .