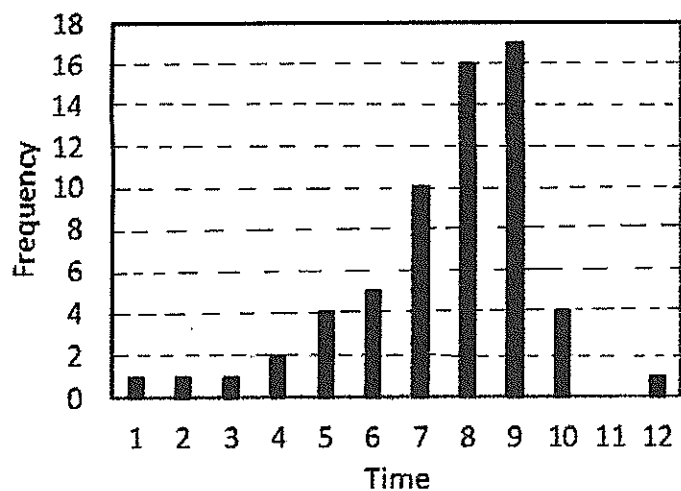


Directions: Solve each of the following problems. Decide which is the **best** of the choices given and answer in the appropriate place on the answer sheet.

1. Which of the following is the primary difference between an experiment and an observational study?

- A. Experiments are only conducted on human subjects; observational studies can be conducted on wildlife subjects.
- B. Experiments must be double-blind; observational studies do not need to be.
- C. Experiments must use randomized treatment and control groups; observational studies also use treatment and control groups.
- D. In an experiment, the researcher manipulates some variables to observe its effect on a response variable; in an observational study, he simply observes and records the observations.
- E. There is no substantive difference—they can both accomplish the same research goals.

2. Sixty-two students were asked to complete a test. The number of minutes each student needed to complete was recorded and the times are displayed in the histogram. Which of the following describes the distribution?



- A. The distribution is approximately normal.
- B. The distribution is roughly symmetric.
- C. The distribution is bimodal.
- D. The distribution is skewed left.
- E. The distribution is skewed right.

3. What is the mean of time (in minute) that these students used to complete the test?

- A. 5.35
- B. 6.25
- C. 7.15
- D. 7.53
- E. 8.53

4. A set of biostatistics test scores in a large class has the following summary:

Minimum	Q1	Median	Q3	Maximum
2	18	22.5	28.5	33

Which statement about outliers must be true?

- A. There is exactly one outlier on the lower end.
- B. There is at least one outlier on the higher end.

- C. There is exactly one outlier on the higher end.
- D. There is at least one outlier on the lower end.
- E. There are no outliers.

5. Which of the following are examples of continuous data?

- I. The speed your car goes
 - II. The number of outcomes of a binomial experiment
 - III. The average temperature in Taipei
 - IV. The wingspan of a bird
- A. I only
 - B. I and II only
 - C. I, II and III only
 - D. I, III, and IV only
 - E. I and IV only

Use the following computer output for a least-squares regression for Questions 6-8.

The regression equation is

predictor	Coef.	SE Coef.	T ratio	P
Constant	147.14	20.696	7.11	0.005
x	0.246	0.021	11.67	0.001

s=136.69 r-sq=0.55 r-sq(adj)=0.53

6. What are the assumptions about linear regression?

- I. The value of the dependent variable y must be normally distributed about the regression line.
 - II. The standard deviation of each of the dependent variables must be the same for each value of the independent variables.
 - III. The sample size must be large to produce significant results.
- A. I only
 - B. II only
 - C. III only
 - D. I and II only
 - E. I and III only

7. What is the equation of the least-squares regression line?

- A. $\hat{y} = 11.69x + 7.11$
- B. $\hat{y} = 0.021x + 20.696$
- C. $\hat{y} = 147.17x + 0.246$
- D. $\hat{y} = 0.246x + 147.14$
- E. none of the above

8. What percent of the variation in \hat{y} is explained by the regression model using the predictor?

- A. 7.6%
- B. 53.0%
- C. 55.0%
- D. 72.8%
- E. 74.2%

9. Which of the following statements is (are) true?

- I. In order to use a χ^2 procedure, the expected value for each cell of a one- or two-way table must be at least 5.
- II. In order to use χ^2 procedures, you must have at least 2...

degrees of freedom.

III. In a 4×3 two-way table, the number of degrees of freedom is 6.

- A. I only
- B. II only
- C. III only
- D. all of them
- E. I and III only

10. The correlation coefficient between length in inches and body mass in pounds for a wildlife species is 0.62. If the lengths are converted from inches to centimeters, what will the correlation coefficient be? (1 in. = 2.54 cm)

- A. -0.62
- B. -0.26
- C. 0.10
- D. 0.26
- E. 0.62

11. A marine ecologist study whether a small increase in the acidity of seawater would adversely affect marine life. She randomly assigned several individuals of sea anemone to one of eight aquariums that had been prepared with environments similar to the natural habitat. Then she randomly selected four tanks, and gradually increased the acidity of the water. She monitored the health of the anemones during the study. Which of these statements is NOT true about this study?

- A. An advantage of using only one variety of sea anemones is that there should be less variability in the response to each treatment.
- B. A disadvantage of using only one variety of sea anemones is that the scope of inference is limited only to that type of anemone.
- C. An advantage of using aquariums is that it is easier to maintain control to avoid confounding factors.
- D. disadvantage of using aquariums is that the anemones might respond differently in the ocean than they do in aquariums.
- E. If the anemones in the aquariums with increased acidity are less healthy than those without increased acidity, it cannot be determined that the increased acidity *caused* the response.

12. Monitoring wildlife populations is important that appropriate management efforts can be implemented. One study found an easy way to estimate the number of nesting sites of terns by monitoring the number of calls heard per minute. More calls happen when there are more birds. In fact, it turned out that the number of calls explained 81% of the variation in the abundance of nests in the breeding sites. Which of the following statements is correct about the correlation between the number of calls and the abundance of nests?

- A. The correlation coefficient is -0.81.
- B. The correlation coefficient is 0.81.
- C. The correlation coefficient is -0.90.
- D. The correlation coefficient is 0.90.
- E. The information is insufficient to answer the question.

13. For a project, Anne recorded the heights of all 28 students in her class and calculated several statistics. She then realized she made an error when recording the height of the tallest

person in the class. She correctly had him listed as the tallest, but needed to add two centimeters to correct it. Which of these measures of spread must remain unchanged after she recalculated the statistics?

- A. Mean absolute deviation
- B. Standard deviation
- C. Variance
- D. Interquartile range
- E. Range

Use the following information for Questions 14 and 15.

In a study to determine if there is a difference in color preferences for young children based on their gender, the followings were obtained in a random sample of children.

Color	Boys	Girls
Blue	10	13
Green	9	16
Red	8	12

14. How many degrees of freedom should be used in a test of independence of gender and color preference?

- A. 1
- B. 2
- C. 3
- D. 4
- E. 5

15. What is the χ^2 statistics for the test?

- A. 0.283
- B. 1.283
- C. 2.283
- D. 3.283
- E. 4.283

16. A producer of weight control products has created a new formula for its food. To compare the effectiveness of the new formula to that of the old formula, it conducted a double-blind randomized experiment. Volunteers tried the old formula on one side of the fence and the new formula on the other in a way that which side got which formula was randomly determined. The response variable was the difference in the number of pimples (old formula - new formula). Which is the correct significance test to perform?

- A. A two-proportion z -test
- B. A two-sample t -test
- C. A matched pairs t -test
- D. A chi-square test of independence
- E. Analysis of variance

17. Which of these is the best description of a P -value?

- A. The probability of making a Type I error.
- B. The probability of making a Type II error.
- C. The probability of getting a test statistic at least as extreme as the observed test statistic, if the null hypothesis is true.
- D. The probability of getting a test statistic at least as extreme as the observed test statistic, if the null hypothesis is false.
- E. The probability of rejecting the null hypothesis if it is, in

fact, false.

Use the followings for question 18-20.

A researcher wishes to try three different techniques to lower the blood pressure of individuals diagnosed with high blood pressure. The subjects were randomly assigned to three groups: the first group takes medication, the second group exercises, and the third diets. After the experiment, he has the following analysis of variance statistics with some missing values:

Source	Sum of squares	d.f	Mean square	F
Between	160.13	?	?	?
Error	104.80	?	8.73	
Total	264.93	14		

18. What is the degree of freedom for the "Between" and "Error"?

- A. 1 and 5
- B. 2 and 12
- C. 12 and 2
- D. 3 and 12
- E. 3 and 11

19. What is the mean square value for the "Between"?

- A. 160.13
- B. 80.07
- C. 40.04
- D. 20.02
- E. 10.01

20. What is the F statistics for the analysis?

- A. 1.15
- B. 2.29
- C. 4.59
- D. 9.17
- E. 18.34