Stat3011: Sample Midterm Exam

Spring, 2003, Tiefeng Jiang

Problem 1 (20 points). Mark "T" in the parenthesis if the statement is true, and mark "F" otherwise.

1. ( ) The shape of a histogram, based on the sample data, will always be similar to the shape of the population histogram.

2. ( ) Outliers are observations which fall far from the bulk of the data.

3. ( ) The sample mean, $\bar{x}$, is sensitive to outliers, however, the sample median is not.

4. ( ) For a data set having a positively skewed unimodal histogram, the median is larger than the mean.

5. ( ) A good residual plot will have no obvious patterns in the scatter of points.

6. ( ) If the correlation between x and y is 0, then there is no relationship between x and y.

7. ( ) When regressing y on x, y is called response variable.

8. ( ) If three different outcomes, A, B and C are independent, of each other, then $P(A \text{ and } B \text{ and } C) = P(A) \cdot P(B) \cdot P(C)$.

9. ( ) Estimates of probabilities should be based on large samples in order to approximate the long run proportions.

Problem 2 (16 points) In the four statements for each question, only one of them is correct. Please mark the correct letter in the parenthesis.

1. ( ) Which of the following variables could yield continuous data?
   a. color of hair
   b. brand of stereo
   c. the number of typing errors in a letter
   d. the time between two consecutive phone calls

2. ( ) A histogram that is positively skewed
   a. is multimodal
   b. is symmetric
   c. has a longer upper tail than lower tail
   d. has a longer lower tail than upper tail

3. ( ) Which of the following is <u>not</u> a measure of the center of a data set?

   a. third quartile

   b. trimmed mean

   c. median

   d. mean

4. ( ) The empirical Rule states that the proportion of observations that are within 3 standard deviations of the mean is at least

   a. 1/3

   b. 2/3

   c. 1/9

   d. 8/9

5. ( ) A value $r = -.40$ indicates that there is a

   a. strong positive relationship between x and y

   b. strong negative relationship between x and y

   c. weak positive relationship between x and y

   d. weak negative relationship between x and y

6. ( ) A good fitting regression line has

   a. small r^2

   b. residual plot with a pattern

   c. big intercept

   d. large r^2

7. ( ) In the linear regression line, $y = a + bx$,

   a. a is the intercept and b is the slope

   b. a is the slope and b is the intercept

   c. y is the independent variable

8. ( ) If a fair coin is flipped twice with the outcome of each flip independent of each other, then the probability that at least one of the two flips results in a head is

   a. 1/4

   b. 2/4

c. 3/4

d. 2/5

Problem 3 (10 points). A highway patrol officer gave 18 speeding tickets last week on a stretch of highway having a 55 mph speed limit. The speeds recorded for the 15 tickets are listed below. Construct a stem-and-leaf display of the data and interpret it:
62, 68, 79, 59, 61, 88, 64, 71, 78, 92, 66, 68, 82, 77, 69, 70, 76, 64.

Problem 4 (12 points). A biologist studying the age of clams has collected a sample of 10 razor clams and aged them. The data is listed below:
1.2, 2.7, 1.1, 0.8, 1.9, 2.3, 1.5, 2.4, 2.9, 5.7.

a. Compute the sample mean and median.

b. Which is a better measure of center in this case?

Problem 5 (17 points). The following summary statistics resulted from a study of the relationship between the cost of a barrel of crude oil (x) and a gallon of regular unleaded gasoline (y).

$$n = 12, \quad \sum x = 24.1 \quad \sum x^2 = 4,932.8$$
$$\sum y = 14.34 \quad \sum y^2 = 17.288 \quad \sum xy = 291.55$$

a. Compute the equation of the least squares line.

b. Roughly what change in the price of a gallon of regular unleaded gasoline results from a 1 dollar increase in the cost of a barrel of crude oil?

c. Predict the price of a gallon of regular unleaded gasoline oil when the cost of a barrel of crude oil is 40.

Problem 6 (15 points). A college math professor has surveyed his records and found the following frequency distribution of grades he has assigned to the 1187 students who have taken his calculus classes.

Grade	A	B	C	D	F

Frequency 125 352 461 187 62

a. What is the probability that a randomly selected student received an A in calculus from this professor?

b. What is the probability that a randomly selected student received an A or a B in calculus from this professor?

c. What is the probability that two independently selected students both failed calculus?

Problem 7 (20 points). A random sample is to be selected from a population that has a proportion of success $p = .7$.

a) What is the smallest value of n for which it is reasonable to assume that the sampling distribution of p is approximately normal?

b) For a sample of $n = 25$, determine the approximate probability that the sample proportion will be within .25 of p.

Problem 8 (20 points). There are no daytime speed limits for automobiles on sections of the state highways and interstate freeways in the state Montana. Let $x =$ the speed of an automobile on a Montana highway. Suppose the mean speed of an automobile on a Montana highway is 72 miles per hour and the standard deviation is 4 miles per hour. Let $x_1,\ x_2,\ \cdots,\ x_{100}$ denote the highway speed of 100 randomly selected automobiles in Montana.

a) What is the approximate probability the the sample mean exceeds 73 mph?

b) What is the approximate probability the the sample mean is between 71 mph and 74 mph?

c) Suppose we don't know that the real mean speed is 72 mph, but know that the sample mean from the 100 numbers is 72 mph. Calculate the 95 percent confidence interval for the true speed.