

Pledge:

10/4/2007
Dr. Lunsford

MATH 171
Test 1

Name: _____
100 Points Possible

I. Short Answer and Multiple Choice. (36 points total)

1. Circle all of the items below that are *measures of center* of a distribution: (2 points)

Mean Standard Deviation Median IQR Minimum

2. Circle all of the items below that are *measures of spread* of a distribution: (2 points)

Mean Standard Deviation Median IQR Minimum

3. A set of midterm exam scores from a large introductory statistics class has a median that is much larger than the mean. Which of the following is most consistent with this information (circle one): (2 points)

- a. A histogram of the exam scores would be symmetric.
- b. A histogram of the exam scores would be left-skewed.
- c. A histogram of the exam scores would be right-skewed.

4. Which of the following is most likely to have a normal distribution? (circle all that apply) (2 points)

- a. average household income in the USA
- b. eye color in the population of LU students
- c. the size of the military budget in 25 developed countries.
- d. SAT Math Scores from the USA in 2006
- e. none of these

5. A study is conducted on students taking a statistics class. Several variables are recorded in the survey. Identify each variable as categorical or quantitative. (2 points each, 6 points total)

_____ Color of car the student owns.

_____ Number of credit hours taken during that semester.

_____ Number of siblings.

6. The Nurses' Health Study has interviewed a sample of more than 100,000 female registered nurses every two years since 1975. The study finds that "light to moderate drinkers had a significantly lower risk of death" than either nondrinkers or heavy drinkers. The Nurses Health Study is (circle all that apply): (2 points)

- a. an experiment
- b. an observational study
- c. a stratified random sample
- d. a treatment

I. Short Answer and Multiple Choice, continued.

7. A committee on community relations in a college town plans to survey local businesses about the importance of students as customers. From telephone book listings, the committee chooses 150 business at random. Of these, 73 return the questionnaire mailed by the committee. The population for this study is: (2 points)

- (a) All businesses in the college town.
- (b) the 150 businesses chosen.
- (c) the 73 businesses that returned the questionnaire.

8. The sample in the setting of the previous exercise is: (2 points)

- (a) all business in the college town.
- (b) the 150 businesses chosen.
- (c) the 73 businesses that returned the questionnaire.

9. As the observations in a distribution become more spread out about the mean, the standard deviation gets: (2 points)

- a. larger
- b. smaller
- c. stays the same

10. Which of the following indicates a weak linear relationship? (2 points)

- a. correlations near -1
- b. correlations near 0
- c. correlations near 1

11. We want to investigate if the average age at which infants begin to crawl is influenced by the average outdoor temperature six months after birth, when they are likely to first begin crawling. The explanatory variable is: (2 points)

- a. average outdoor temperature
- b. average age at which infants begin to crawl

12. Removing an outlier will _____ the correlation, r , between two variables. (2 points)

- a. increase
- b. decrease
- c. either increase or decrease

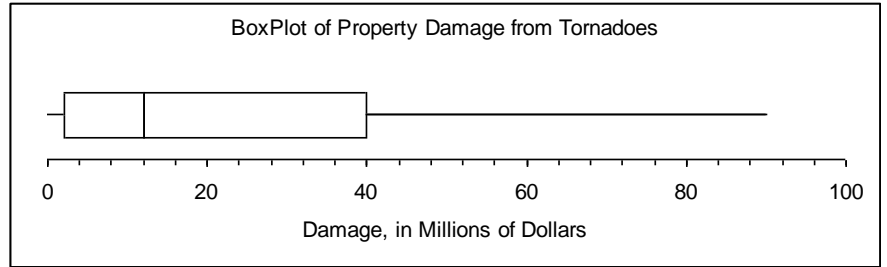
13. Indicate whether each of the following statements is true or false by writing “true” or “false” in the blank provided next to each statement. (2 points each, 8 total)

- _____ A) If the correlation between two variables is positive, then the explanatory variable *causes* the response variable.
- _____ B) The standard normal distribution has a mean equal to 1.
- _____ C) If people with larger heads tend to be more intelligent, then we would expect the correlation between head size and intelligence to be positive.
- _____ D) If we use a regression line to make predictions we are sure that the predictions are highly accurate.

Problem II. Below is a box plot of the property damage (in millions of dollars) caused by tornadoes per year over the period from 1950 to 1999 in each of the 50 states. Use it to answer the following questions. (4 points each, 12 total)

a. What is the median amount of property damage?

b. Twenty five percent of the property damage caused is greater than what amount?



c. Do you think the mean amount of property damage caused is larger or smaller than the median amount? Why?

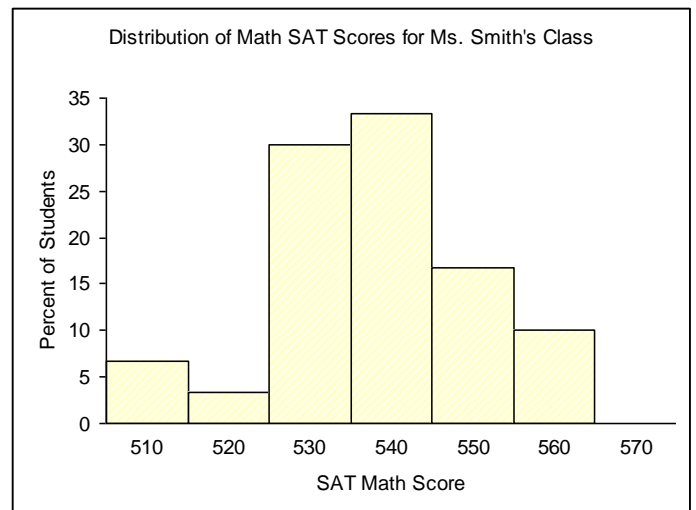
Problem III. The following 15 scores on the mathematics portion of the SAT were the scores for the students in Ms. Brown's pre-calculus class. (13 points total)

521	558	552	527	535
563	522	574	589	596
510	531	535	524	537

(a) Make a stem and leaf plot of the above data. (6 points)

(b) The mean SAT math score for Ms. Brown's students is 544.9 with a standard deviation of 26.01. To your right is a histogram of the SAT Math Scores for Ms. Smith's Class. Use the histogram to estimate the mean SAT Math score for Ms. Smith's class. You should not need to use your calculator to estimate this number! (4 points)

(c) True or False (circle one): Ms. Smith's students have more variability in their SAT Math scores than Ms. Brown's students. (3 points)



Problem IV. Many high school students take either the SAT or the ACT. However, some students take both. Data was collected from over two thousand students who took both college entrance exams. The average SAT score was 940 with a standard deviation of 160. The average ACT score was 22 with a standard deviation of 6. The correlation between the two variables equals 0.817. (13 points total)

(a) John received a 1130 on the SAT, while Richard got a 30 on the ACT. Assuming both tests measure the same thing, who did better, John or Richard? Please show the appropriate computations to justify your answer. (6 points)

(b) An admissions officer at a prestigious university used the above information to find the least squares regression line to predict an SAT score given an ACT score. The line is given by

$$(\text{Predicted SAT Score}) = 460.69 + 21.79 * (\text{ACT Score})$$

What would be the predicted SAT score for a student whose ACT score was 25? (3 points)

(c) How reliable is the prediction equation in part (b)? To answer this question, consider the percentage of explained variation in SAT scores via the linear model. If you applied to this college and had only taken the ACT, would you want them to use the equation in part (b) to predict your SAT score? Why or why not? (4 points)

Problem V. Assume women's heights are approximately normally distributed with a mean of 64 inches and a standard deviation of 2.5 inches. The height of the Rockette dancers at New York City Music Hall must be between 65.5 inches and 68.0 inches tall. What percent of women meet the height requirement to be a Rockette? Clearly show all calculator input and draw a picture to graphically represent this percent. (6 points)

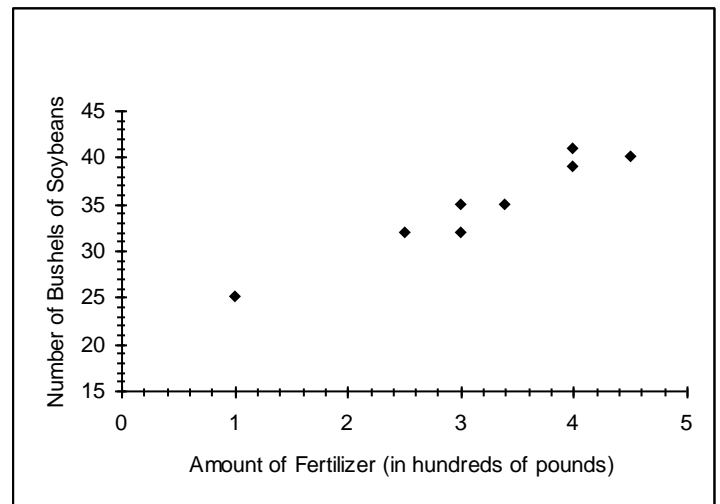
Problem VI. The scores on a university physics exam are normally distributed with a mean of 65 and a standard deviation of 11. If the bottom 10% of students will fail the exam, what is the lowest score that a student can have to be awarded a passing grade? Clearly show all calculator input. (6 points)

Problem VII. A farmer is interested in the relationship between the amount of fertilizer used (in hundreds of pounds) and the number of bushels of soybeans produced. The farmer conducted an experiment on several plots of land and obtained the following data. Below you are also given a scatterplot of the data. (14 points total)

Amount of Fertilizer (in hundreds of pounds)	1	2.5	3	3	3.4	4	4	4.5
Number of Bushels of Soybeans Produced	25	32	35	32	35	39	41	40

(a) Plot the centroid of the data on the scatterplot. Clearly indicate the point on the plot and its coordinates below. (4 points)

(b) Find and accurately plot the least squares regression line on the scatterplot to your right. Clearly indicate the equation of your line, including what the variables x and \hat{y} represent, below. (6 points)



(c) Use the regression line model given above to predict the expected number of bushels of soybeans produced when 200 pounds of fertilizer is used. Show your work below to get the answer and show this prediction graphically (by using up and over lines) on the scatterplot. (4 points)