Name: _____

School: _____

University of Houston

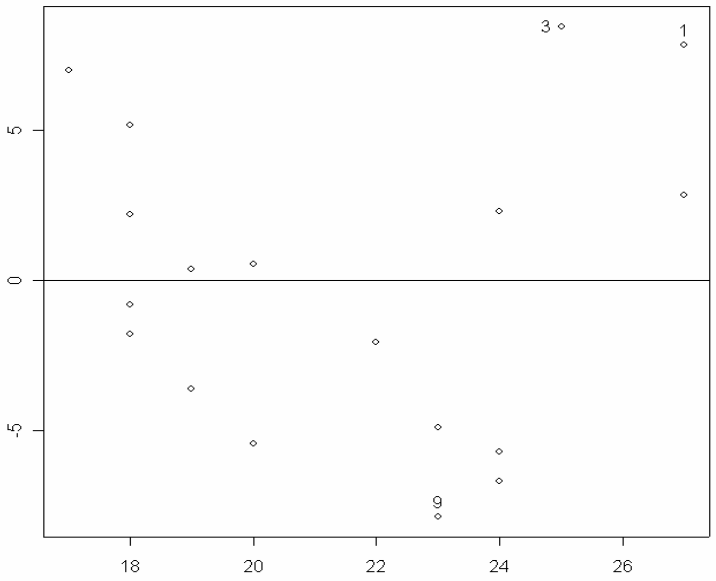
High School Contest – Spring 2006

Statistics Test

Part I – Multiple Choice. Each problem is worth 5 points.

1. Fifty cars were selected to test the effect of a gasoline additive on exhaust emissions. Emissions were measured from each car for a tank of gasoline without the additive and then for a tank with the additive. The researcher would like to test the research hypothesis that the additive decreases emissions against the null hypothesis that it has no effect. Which of the following tests would be least appropriate?
   (a) The normal test of a hypothesis about a population mean
   (b) The paired-sample Student-t test
   (c) The Wilcoxon signed-rank test
   (d) The sign test
   (e) The two-sample Student-t test for independent samples

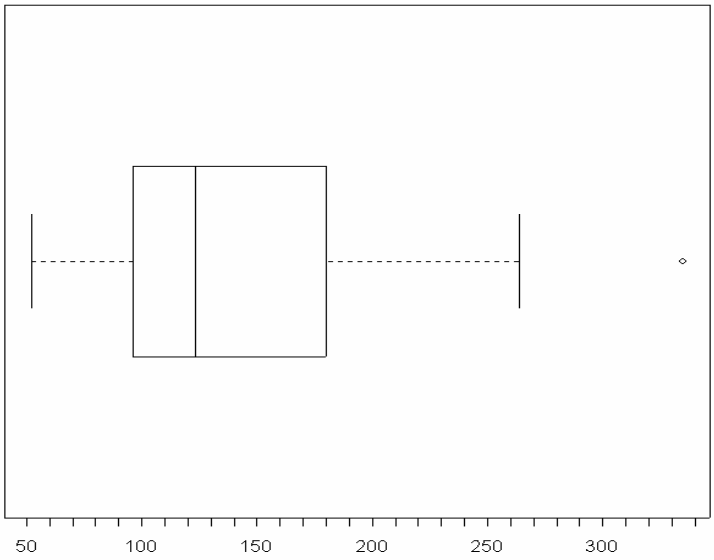
2. The diagram below shows the values of the predictor variable on the horizontal axis and the corresponding residuals from a simple linear regression. Which of the following statements is least justifiable?
   (a) The assumption of a linear relationship between the variables is unassailable.
   (b) The point labeled 1 may have a large influence on the estimated slope.
   (c) The points labeled 3 and 9 should perhaps be investigated further.
   (d) The average of the residuals is zero.



Name: _____

School: _____

3. The box and whisker diagram below indicates that the interquartile range of the data is
   (a) about 60.
   (b) about 120.
   (c) strongly skewed.
   (d) bimodal.
   (e) about 80.

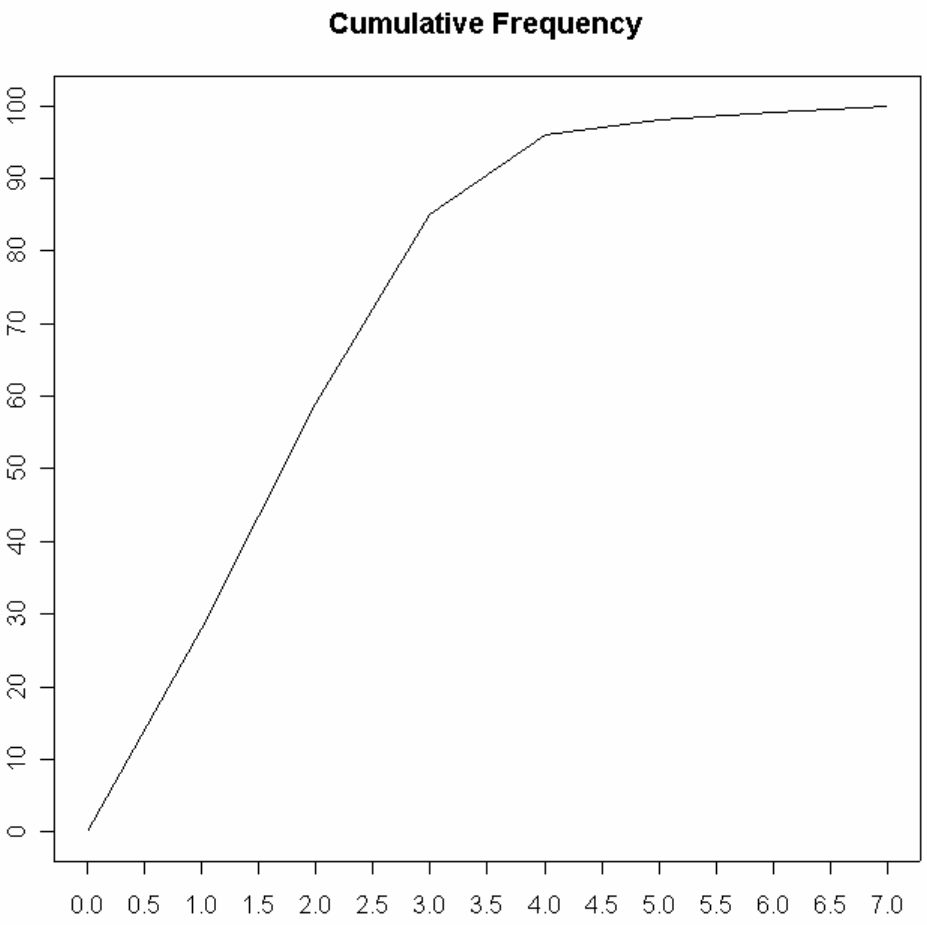


4. Which of the following is not generally true?
   (a) For large enough samples, the sample mean and sample median are close together.
   (b) For large samples, well-constructed histograms give an accurate picture of the distribution.
   (c) For a given confidence level, larger samples give confidence intervals of smaller expected width.
   (d) For large samples, procedures based on the normal distribution can often be used even when the sample data comes from a non-normal distribution.
   (e) For a normally distributed variable, the probability is about 95% that the value of the variable is within two standard deviations of its mean.

Name: _____
School: _____

Part II – Free Response. Each problem is worth 10 points. Show your work.

5. The figure below is a cumulative frequency polygon, or ogive, of 100 measurements. Estimate the median and quartiles of the data.



Name: _____

School: _____

6. Scores on a nationwide professional qualifying exam are normally distributed and have a population mean of 700 and a population standard deviation of 75. An examinee scored at the 40^{th} percentile nationally. What was the examinee's numerical score?

7. A random sample of size 36 is taken from the normal distribution with mean 5 and variance 9. Let $\overline{X}$ denote the sample average. What is the probability that $3.5 < \overline{X} < 6$?

Name: _____

School: _____

8. It is known that about 10% of all first-grade children have a certain learning disability. Educators would like to know that percentage more precisely, to two significant digits. They intend to conduct a study of a sample of first-graders. How large a sample should they take to achieve the desired accuracy with 95% confidence?

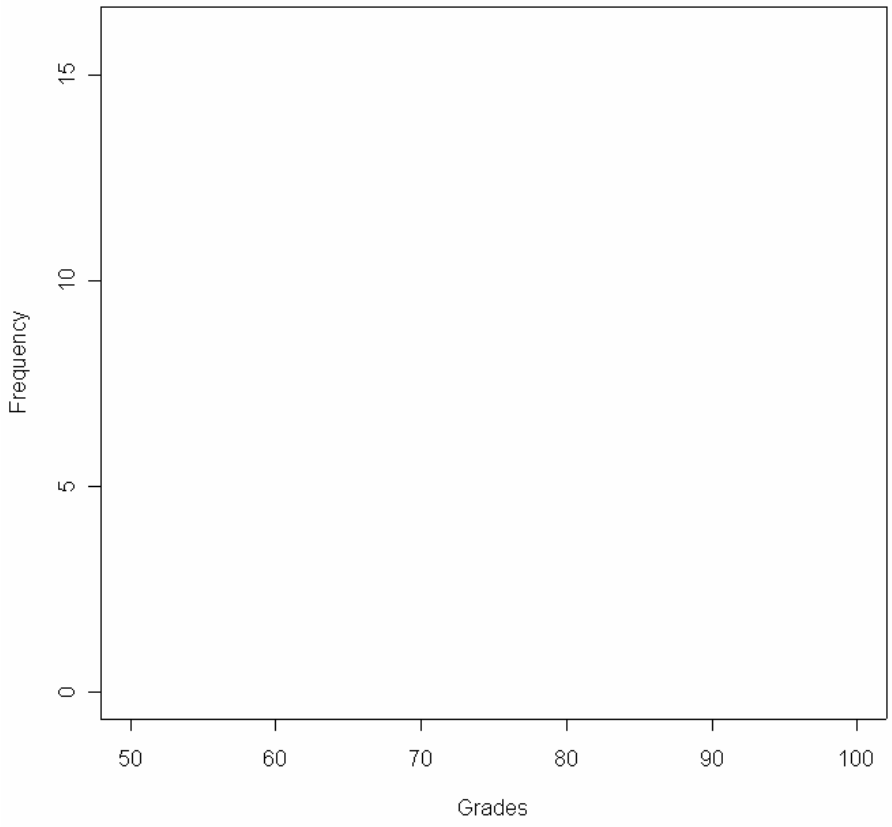
9. A sample of 25 observations comes from a distribution that is known to be normal with a standard deviation of 3 and an unknown mean μ. A researcher is interested in testing the research hypothesis H_1: $\mu > 0$ against the null hypothesis H_0: $\mu = 0$ using a normal test based on the sample mean and a significance level of 0.05. What is the power of the test if the true mean is 2?

Name: _____
School: _____

10. The stem and leaf diagram below shows exam grades of 69 statistics students. The stem is the first digit of a student's score and the leaf is the second digit. Beneath the stem plot, construct a histogram of the data.

    Stem|Leaves

    5 | 778
    6 | 1122334
    6 | 555666777799999
    7 | 001123333344
    7 | 5555666777779
    8 | 00222244
    8 | 5567788999
    9 | 1



Name: _____

School: _____

11. Students in a random sample of size 40 from university H were classified as to gender and preference for morning or afternoon classes. At a significance level of 10%, is there evidence of a difference in the preferences of men and women at university H?

	Men	Women
Morning	7	14
Afternoon	10	9

Name: _____

School: _____

12. For June through September of 1999, the rainfall and yield per acre of cotton were measured for 8 farms in Texas. The results are tabulated below. Find a 95% confidence interval for the expected increase in yield for each additional inch of rain.

Rain	Yield
8.2	17.2
6.4	17.6
12.8	19.6
7.6	19.8
13.8	20.0
16.1	23.5
12.4	20.2
12.8	20.2