

University of Houston
High School Math Contest
Spring 2011
Statistics Test

Name _____

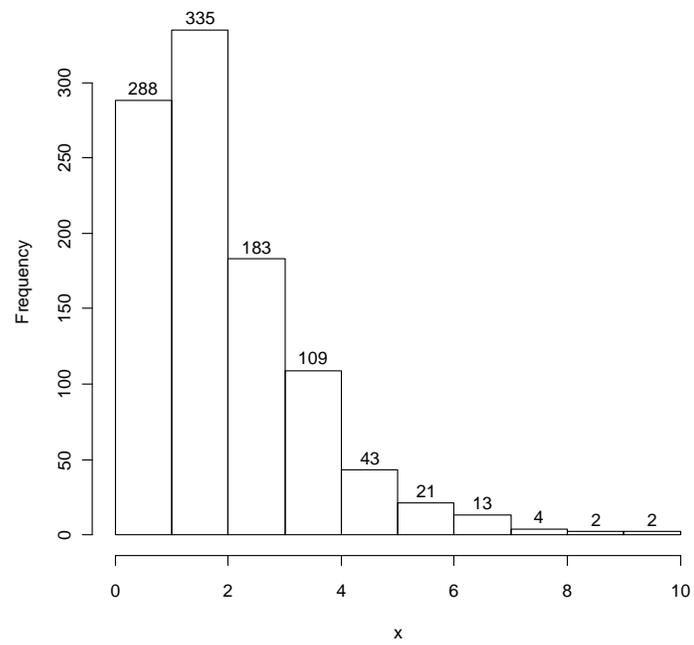
School _____

Exam Time: 1 hour

Any type of calculator is allowed. Tables are provided.

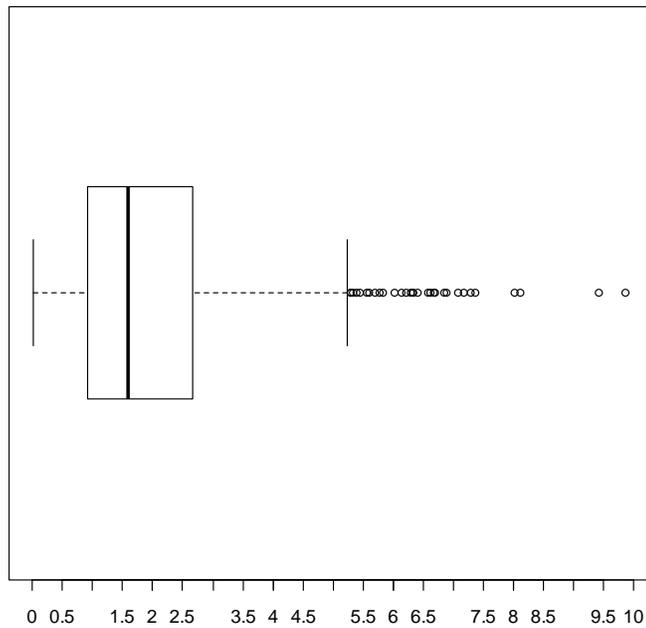
The first 15 problems are multiple choice. The last question calls for a written response. It will be used to break ties.

1. The figure below is a histogram of 1000 measurements. The mean of the data is closest to
- (a) 1.95
 - (b) 1.00
 - (c) 2.25
 - (d) 3.10.
 - (e) Since the data is not given, the mean cannot be estimated.



2. The figure below is a box and whisker diagram of some data. Which of the following is true?

- (a) The mean is less than the median.
- (b) The mean is greater than the median.
- (c) The mean is about the same as the median.
- (d) The mean is a more reliable location measure than the median.
- (e) The figure gives no information about the mean.

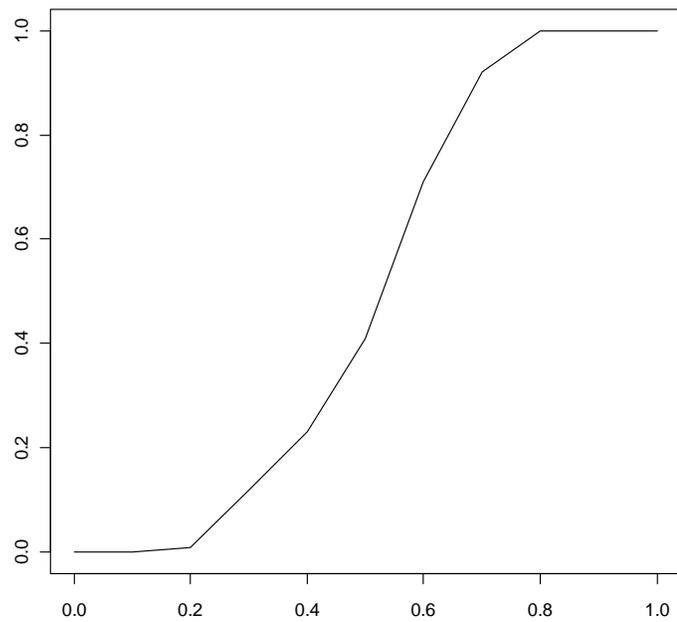


3. Refer to the same figure above. The interquartile range of the data is closest to

- (a) 10.00
- (b) 2.75
- (c) 1.75
- (d) 0.50
- (e) 5.00

4. Scores on a national achievement test are normally distributed with a mean of 700 and a standard deviation of 75. The 60th percentile of test scores is closest to
- (a) 800
 - (b) 775
 - (c) 850
 - (d) 625
 - (e) 720
5. DOE officials are planning a survey to estimate the percentage of high school students who have experimented with peppermint candy. They want the error to be no larger than 2.5% with at least 95% confidence and they do not want to survey more students than is necessary. How large a sample of students should they take? You may assume that the sample is taken with replacement.
- (a) 960
 - (b) 2350
 - (c) 1540
 - (d) 9020
 - (e) 370
6. Which of the following random variables does not have a binomial distribution?
- (a) The number of shirts that have defects in a sample of 20 coming from a production line.
 - (b) The number of boxes of Cracker Jacks that must be purchased to get all 6 of the prizes offered.
 - (c) The number of defective items in a sample of 10 randomly chosen without replacement from a large population of such items.
 - (d) The number of times a total of 7 spots appears in 100 rolls of a pair of dice.
 - (e) The number of base hits by Albert Pujols in the first three games next season in which he has 4 official at bats.

7. A random sample of size 25 is taken from the normal distribution with mean 20 and variance 9. Let \bar{X} denote the sample average. What is the probability that \bar{X} is larger than 21?
- (a) between 0.0475 and 0.0485
 - (b) between 0.1020 and 0.1038
 - (c) between 0.2877 and 0.2912
 - (d) between 0.0192 and 0.0197
 - (e) between 0.0222 and 0.0228
8. An opinion polling firm conducts a survey by the following method. A random sample of 5 precincts is selected and then in each of these precincts a sample of 200 registered voters is selected. The selected voters are polled about their approval or disapproval on a question before the city council. This is an example of
- (a) a randomized complete block design
 - (b) stratified sampling with proportional representation
 - (c) an observational study
 - (d) interval data
 - (e) cluster sampling
9. The figure on the next page is a relative frequency polygon of 100 measurements. Which of the following is not true.
- (a) There are more measurements between 0.7 and 0.9 than between 0.4 and 0.6.
 - (b) The third quartile of the data is about 0.6.
 - (c) The median of the data is greater than 0.4.
 - (d) There is no data between 0.8 and 1.0.
 - (e) More than 80 of the data values are less than or equal to 0.7.



10. A six-sided die is tossed 180 times with the following results.

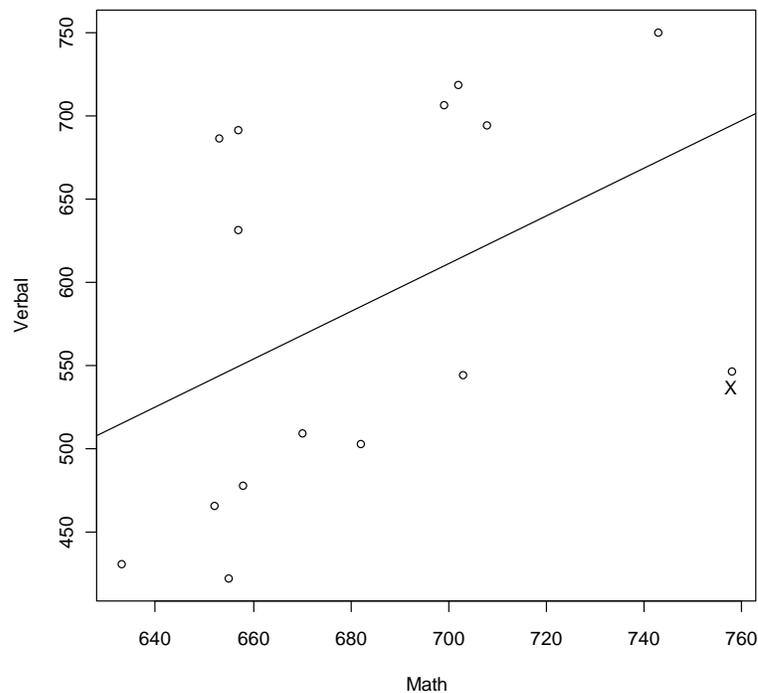
Face	1	2	3	4	5	6
Frequency	39	51	20	22	26	22

We would like to test the null hypothesis that the die is fair against the alternative hypothesis that it is not fair by using a chi-square test. Which of the following significance levels (maximum acceptable probability of type 1 error) is the smallest that would result in rejecting the null hypothesis?

- (a) 20%
- (b) 10%
- (c) 5%
- (d) 2%
- (e) 1%

11. The figure below shows math and verbal scores for 15 students. Seven of them are native speakers of the language and eight are not. Which of the following is not true.

- (a) The correlation between math and verbal scores is positive and moderately strong.
- (b) Correlations calculated separately for each group are even stronger than the overall correlation.
- (c) If this is a sample from the general population of students, an increase of 20 in the math score corresponds to an expected increase of about 30 in the verbal score.
- (d) The least squares line passes through the point whose coordinates are the average math and average verbal scores.
- (e) The point labeled with an X is an outlier but it has little influence on the fitted line.



12. A random sample of 400 measurements of numeric variable were recorded. The sample average was 12.56 and the sample standard deviation was 2.44. A 95% confidence interval for the population mean is
- (a) from 12.321 to 12.799
 - (b) from 7.778 to 17.342
 - (c) from 10.778 to 14.342
 - (d) from 11.321 to 13.799
 - (e) Not enough information is given to answer the question.
13. Let X denote the number of spots on a roll of a single fair, six-sided die. The standard deviation of the random variable X is
- (a) 2.917
 - (b) 2.012
 - (c) 0.822
 - (d) 1.708
 - (e) 1.385
14. Twenty adult men were selected to test the effect of a drug on systolic blood pressure. Each subject had his blood pressure measured before administration of the drug and again 30 minutes after administration. The researcher would like to test the null hypothesis that the drug has no effect against the alternative research hypothesis that the drug reduces blood pressure. Which of the following procedures should the researcher use?
- (a) A least squares fit of the “after” measurements as a linear function of the “before” measurements should be calculated.
 - (b) The “before” measurements and the “after” measurements should be treated as independent samples from two populations.
 - (c) The mean and standard deviation of the combined measurements should be calculated. If the absolute value of the sample mean is greater than 1.96 times the sample standard deviation, the researcher can conclude that the drug is effective.
 - (d) Women should be included in the study. Otherwise, no valid conclusions can be drawn.
 - (e) A paired sample student-t test should be applied to the differences between the “before” and “after” measurements of the subjects.

15. Scores on a national achievement test are normally distributed with a mean of 700 and a standard deviation of 75. A student scores 830. What is her percentile score?
- (a) 85%
 - (b) 75%
 - (c) 91%
 - (d) 96%
 - (e) 99%
16. A random variable Z has a standard normal distribution. Estimate the probability that $Z^4 + Z \leq 1$. Explain your approach.