Part I: Multiple Choice
This portion of the test will determine 60% of your overall test grade. Each question is worth 3 pts.

1. The back-to-back stemplot below gives the average bowling scores of male and female participants in the finals of a national tournament.

Key: 28:6 = 286 participants

MALES

FEMALES

28
27
26
25
378
0267
028
89
997632 88421 80 3 1

Which of the following is not true?
(A) The same number of male and female bowlers participated.
(B) The males’ scores are skewed while the female scores are roughly symmetric.
(C) The median and mean female scores are roughly equal.
(D) The mean male score is greater than the median male score.
(E) The range of the male scores equals the range of the female scores.

2. A 1995 poll by the Program for International Policy asked respondents what percentage of the U.S. budget they thought went to foreign aid. The mean response was 18% and the median was 12%. What do these responses indicate about the shape of the distribution of all responses?
(A) The distribution is skewed to the left.
(B) The distribution is skewed to the right.
(C) The distribution is symmetric.
(D) The distribution is uniform.
(E) The shape of the distribution cannot be approximated with just the mean and the median. More information is necessary to answer this question.

3. A reporter from a consumer magazine notes that the mean price of 30 selected grocery items at a particular supermarket is $0.75 with a standard deviation of $0.20. Suppose the following week the store raises all prices by 5 cents, and then the next week they lower all prices by 5%. What are the new mean and standard deviation for the 30 selected items?

(A) $0.76, $0.19  
(B) $0.76, $0.20  
(C) $0.76, $0.24  
(D) $0.80, $0.20  
(E) $0.80, $0.24

4. When there are multiple gaps and clusters, which of the following is the best choice to give an overall picture of a distribution?

(A) Mean and standard deviation  
(B) Median and interquartile range  
(C) Boxplot with its five number summary  
(D) Stemplot or histogram  
(E) None of the above are really helpful in showing gaps and clusters.

5. A data set has the following five-number summary: {31, 50, 58, 62, 87}. Which of the following pairs of values in this data set would be considered outliers? Please show your work.

(A) 31, 81  
(B) 32, 81  
(C) 31, 80  
(D) 32, 80
6. The 70 highest dams in the world have an average height of 200 meters with a standard deviation of 35 meters. The Russian dams, Nurek and Charvak have heights of 240 and 168 meters respectively. The Hoover and Grand Coulee dams have heights with z-scores of +2.69 and -1.13 respectively. List the dams in order by ascending size.
(A) Charvak, Grand Coulee, Hoover, and Nurek
(B) Charvak, Grand Coulee, Nurek, and Hoover
(C) Grand Coulee, Charvak, Hoover, and Nurek
(D) **Grand Coulee, Charvak, Nurek, and Hoover**
(E) None of the above orders is correct.

7. A trucking firm determines that its fleet of trucks averages a mean of 12.4 miles with a standard deviation of 1.2 miles per gallon on cross-country hauls. If the distribution of miles per gallon follows a normal distribution, what is the probability that one of the trucks averages fewer than 10 miles per gallon?
(A) 0.0082
(B) **0.0228**
(C) 0.4772
(D) 0.5228
(E) None of the above. The correct answer is ________________.

8. Company I manufactures bomb fuses that burn an average of 50 minutes with a standard deviation of 10 minutes, while Company II advertises fuses that burn on average 55 minutes with a standard deviation of 5 minutes. Which company’s fuse is more likely to last at least 1 hour? Assume normal distributions of fuse times.
(A) Company I’s, because of its greater standard deviation.
(B) Company II’s, because of its greater mean.
(C) For both companies, the probability that a fuse will last at least 1 hour is 15.9%.
(D) For both companies, the probability that a fuse will last at least 1 hour is 84.1%.
(E) The problem cannot be solved from the information given.

_____9. Cucumbers grown on a certain farm have weights that follow a normal distribution with a standard deviation of 2 ounces. What is the mean weight if 85% of the cucumbers weigh less than 16 ounces?
(A) 13.92
(B) 14.30
(C) 14.40
(D) 14.88
(E) None of the above

_____10. Suppose that the distribution of math SAT scores from your state this year is normally distributed with mean 480 and standard deviation 100 for males, and mean 440 and standard deviation 120 for females. If someone who scores 780 or higher on math SAT can be considered a genius, what is the proportion of geniuses among the male SAT takers?
(A) 28%
(B) 14%
(C) 3%
(D) 1.4%
(E) None of the above, the correct answer is ________________________.

_____11. The Wechsler Adult Intelligence Scale results in a normal distribution with a mean of 110 and a standard deviation of 25. If someone tests at the 80th percentile, what score did the individual have?
(A) 89
(B) 113
(C) 118
(D) 130
(E) 131
12. Four pairs of data are used in determining a regression line \( y = 3x + 4 \). If the four values of the explanatory variable are 32, 24, 29, and 27, respectively, what is the mean of the four values of the response variable?

(A) 68  
(B) 84  
(C) **88**  
(D) 100  
(E) The mean cannot be determined from the given information.

13. Which of the following statements about the correlation coefficient are true?

I. The correlation coefficient and the slope of the regression line may have opposite signs.  
II. A correlation of 1 indicates a perfect cause-and-effect relationship between the variables.  
III. Correlations of .87 and -.87 indicate the same degree of clustering around the regression line.

(A) I only  
(B) **II only**  
(C) III only  
(D) I and II  
(E) I, II, and III

An argument can certainly be made here that III is a reasonable statement.

14. Jay Bennett (Chance, Winter 1995, page 38) calculated the regression line for average 1991 SAT scores (total math plus verbal) versus number of dollars spent per student in 1991 for New Jersey school districts and obtained a slope of 0.0227 and a \( y \)-intercept of 707. What average SAT result does this regression line predict for students in a district that spends $10,000 per student?

(A) 467  
(B) 480  
(C) 730  
(D) **934**  
(E) More information is needed to make this calculation.

15. Which of the following statements about residuals are true?

I. The mean of the residuals is always zero.  
II. A residual is the difference between the predicted and actual value of the explanatory variable.
III. A curved pattern in the residual plot is an indication that a linear model is a good fit.

(A) I and II
(B) II and III
(C) I and III
(D) I, II and III
(E) None of the above gives the complete set of true responses.

16. A survey is to be taken to ascertain student opinions about the quality of teaching at a high school. Consider the following survey methods of picking 100 students out of the 2000 students registered at the school.

I. Randomly pick one day of the week. As the students arrive at school that day, they are asked to write their names on slips of paper and place these slips in a box. After all 2000 students have entered their names, the principal mixes the papers up in the box and reaches into the box and pulls out 100 names.

II. Using an official school roster of the 2000 students, pick every 20th name.

III. Using a random number table, 25 names are chosen from each of separate lists of freshmen, sophomores, juniors, and seniors.

IV. Randomly pick 20 classrooms. Then randomly pick 5 students from each of the 20 classrooms.

How many of the above survey methods will result in a simple random sample of 100 students?

(A) None
(B) One
(C) Two I think that I and II are sound strategies
(D) Three
(E) Four

17. A set of data relates the amount of annual salary raise and the performance rating. The least squares regression equation is \( y = 1,400 + 2,000x \) where \( y \) is the estimated raise and \( x \) is the performance rating. Which of the following statements is not correct?

(A) For each increase of one point in performance rating, the salary raise will increase on average by $2,000.
(B) This equation approximates predicted salary raises with an average error of 0.
(C) A performance rating of 0 will yield a predicted raise of $1,400.
(D) The correlation for the data is positive.
(E) All of the above are true.
18.) A reading specialist in a large public school system believes that the more time students spend reading, the better they will do in school. She plans a middle school experiment in which an SRS of 30 eighth graders will be assigned four extra hours of reading per week, and an SRS of 30 seventh graders will be assigned two extra hours of reading per week, and an SRS of 30 sixth graders with no extra assigned reading will be a control group. After one year, the mean GPAs from each group will be compared. Is this a good experimental design?

(A) Yes
(B) No, because while this design may point out an association between reading and GPR, it cannot establish a cause-and-effect relationship.
(C) No, because we do not have blinding and thus we cannot draw any relevant conclusions from this experiment.
(D) No, because any conclusion would be flawed because of blocking bias.
(E) No, because grade level is a lurking variable which may well be confounded with the variables under consideration.

19. Which of the following is not important in the design of experiments?

(A) Control of confounding variables
(B) Randomization in assigning subjects to different treatments
(C) Use of a lurking variable to control the placebo effect
(D) Replication of the experiment using sufficient numbers of subjects
(E) All of the above are important in the design of experiments.

20. A nutritionist wants to study the effect of storage time (6, 12, and 18 months) on the amount of vitamin C present in freeze dried fruit when stored for these lengths of time. Vitamin C is measured in milligrams per 100 milligrams of fruit. Six fruit packs were randomly assigned
to each of these specific storage times. The treatment, experimental unit, and response are respectively:

a. A specific storage time, amount of vitamin C, a fruit pack
b. A fruit pack, amount of vitamin C, a specific storage time
c. Random assignment, a fruit pack, amount of vitamin C
d. **A specific storage time, a fruit pack, amount of vitamin C**
e. A specific storage time, the nutritionist, amount of vitamin C