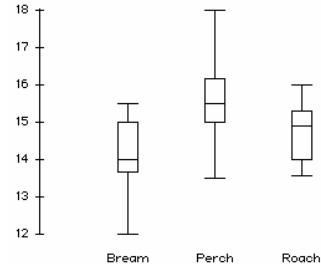


## AP Statistics Review

1. A study was conducted on the weights of three different species of fish found in a lake in Finland. These three fish (bream, perch and roach) are commercial fish. Their weights are displayed in the boxplots below. Which of the following statements comparing these boxplots is **NOT** correct?

- (a) The median weights of the three species differ.
  - (b) The spread of roach is less than the spread of the other two species.
  - (c) The distributions of weights are approximately symmetric for all three species.
  - (d) There are no outliers in weight for the three species.
- The variability in the weights for the three species exceeds the variation in the three species'



2. Foresters use regression to predict the volume of timber in a tree using easily measured quantities such as diameter. Let  $y$  be the volume of timber in cubic feet and  $x$  be the diameter in feet (measured at 3 feet above ground level). One set of data gives  $y = -30 + 60x$ . The predicted volume for a tree of 18 inches is:

- (a) 1050 cubic feet
- (b) 600 cubic feet
- (c) 105 cubic feet
- (d) 90 cubic feet
- (e) 60 cubic feet

3. Which one of the following is not a principle of experimentation?

- (a) Randomly allocating experimental units to treatments.
- (b) Stratifying the experimental units into groups of similar individuals and applying different treatments to each stratum.
- (c) Using double blindness to eliminate bias.
- (d) Replicating to measure overall experimental error and increase precision.
- (e) Using a control group to determine whether treatment really works.

4. Which of the following are true statements?

- I. Voluntary response samples often underrepresent people with strong opinions.
  - II. Convenience samples often lead to undercoverage bias.
  - III. Questionnaires with nonneutral wording are likely to have response bias.
- (a) I and II
  - (b) I and III
  - (c) II and III
  - (d) I, II, and III
  - (e) None of the above

5. The heights of married men are approximately normally distributed with a mean of 70 and a standard deviation of 3, while the heights of married women are approximately normal distributed with a mean of 65 and a standard deviation of 2.5. Determine the probability that a randomly selected married woman is taller than a randomly selected married men.

- (a) 0.05
- (b) 0.10
- (c) 0.15
- (d) 0.20
- (e) Cannot be determined from the given information.

6. Suppose the average height of policemen is 71 inches with a standard deviation of 4 inches, while the average for policewoman is 66 inches with a standard deviation of 3 inches. If a committee looks at all ways of pairing up one male with one female officer, what will be the mean and standard deviation for the difference in heights for the set of possible partners?

- (a) Mean of 5 inches with a standard deviation of 1 inch.
- (b) Mean of 5 inches with a standard deviation of 3.5 inches.
- (c) Mean of 5 inches with a standard deviation of 5 inches.
- (d) Mean of 68.5 inches with a standard deviation of 1 inch.
- (e) Mean of 68.5 inches with a standard deviation of 3.5 inches.

7. Which of the following are true statements?

- I. By the law of large numbers, the mean of a random variable will get closer and closer to a specific value.
  - II. The standard deviation of a random variable is never negative.
  - III. The standard deviation of a random variable is 0 only if the random variable takes a lone single value.
- (a) I and II
  - (b) I and III
  - (c) II and III
  - (d) II and III
  - (e) I, II, and III

8. The stemplot displays the 1988 per capita income (in hundreds of dollars) of the 50 states. Which of the following best describes the data?

- (a) Skewed distribution, mean greater than median
- (b) Skewed distribution, median greater than mean
- (c) Symmetric distribution, mean greater than median
- (d) Symmetric distribution, median greater than mean
- (e) Symmetric distribution with outliers on high end

9. In order to rate TV shows, phone surveys are sometimes used. Such a survey might record several variables, some of which are listed below. Which of these variables are categorical?

- I. The type of show being watched
- II. The number of persons watching the show
- III. The ages of persons watching the show
- IV. The name of the show being watched
- V. The number of times the show has been watched in the last month

- (a) II, III, and V
- (b) I only
- (c) I and V
- (d) I and IV
- (e) None of the above

10. Ms. Myers' precalculus class had a standard deviation of 2.4 on a trigonometry test, while Ms. Gentry's precalculus had a standard deviation of 1.2 on the same test. What can be said about these two classes?

- (b) Ms. Myers' class is more homogeneous than Ms. Gentry's class.
- (c) Ms. Gentry's class is more homogeneous than Ms. Myers' class.
- (d) Ms. Gentry's class did less well on the test than Ms. Myers' class.
- (e) Ms. Myers' class performed twice as well on the test as Ms. Gentry's class.
- (f) Ms. Myers' class performed 1.2 points better on the test as Ms. Gentry's class.

11. The following is the DataDesk statistical summary of the gold medal performance in the men's long jump (measured in inches) for the modern Olympic series starting in 1900. Approximately what percent of the data lie between 298 and 334.64?

- (f) 25%
- (g) 33%
- (h) 50%
- (i) 67%
- (j) 75%

Summary of long jump	
No Selector	
Percentile	25
Count	22
Mean	314.096
Median	312.781
StdDev	20.7055
Min	281.5
Max	350.5
Lower 1th %tile	298
Upper 1th %tile	334.64

12. Data are obtained for a group of college freshman examining their SAT scores (math plus verbal) from their senior year of high school and their GPAs during their first year of college. The resulting regression equation is:

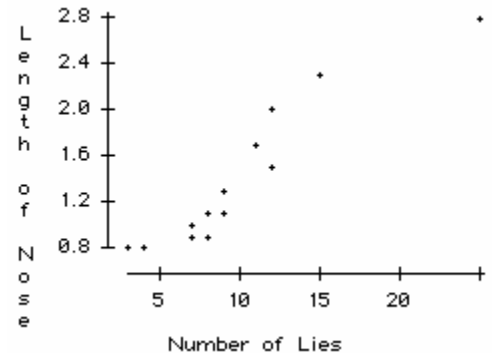
$$y = 0.00161x + 1.35 \text{ with } s_x = 120, \text{ and } s_y = .3057$$

What percentage of the variation in GPAs can be explained by looking at SAT scores?

- (a) 0.161%
- (b) 16.1%
- (c) 39.9%
- (d) 63.2%
- (b) This value cannot be computed from the information given.

13. The coefficient of determination of the data described in the scatterplot is:

- (a) 0.35
- (b) 0.65
- (c) -0.80
- (d) 0.88
- (e) 0



14. As reported in the *Journal of the American Medical Association* (June 13, 1990), for a study of ten nonagenarians, the following tabulation shows a measure of strength versus a measure of functional mobility

Strength (kg)	7.5	6	11.5	10.5	9.5	18	4	12	9	3
Walk time (s)	18	46	8	25	25	7	22	12	10	48

What does the slope of the least-squares regression line signify?

- (f) The sign is positive, signifying a direct cause-and-effect relationship between strength and mobility.
- (g) The sign is positive, signifying that the greater the strength, the greater the functional mobility.
- (h) The sign is negative, signifying that the relationship between strength and functional mobility is weak.
- (i) The sign is negative, signifying that the greater the strength, the less the functional mobility.
- (j) The slope is close to zero, signifying that the relationship between strength and functional mobility is weak.

15. Fourth grade children were asked what emotion they associated with the color red. The responses for emotion and gender of the children are summarized in the following two-way table.

	Anger	Pain	Happiness	Love
Male	35	27	12	38
Female	27	17	19	39

What proportion of the males associate the color red with love?

- (a) 0.5234
- (b) 0.3598
- (c) 0.3393
- (d) 0.1822
- (e) 0.1775

16. A strong negative association between Average State SAT scores and Percentage of students taking the SAT reflects which underlying relationship?

- (a) causation
- (b) correlation
- (c) common response
- (d) extrapolation
- (e) confounding

17. Two variables are confounded when:

- (e) The effect of one variable on the response variable is dependent upon the effect of the other variable.
- (f) The effect of one variable on the response variable cannot be separated from the other variable.
- (g) The effect of one variable on the response variable changes the impact of the other variable on the response variable.
- (h) Both variables are classified as lurking or extraneous variables.
- (i) They interact in their effects on the response variable.

18. A simple random sample of size  $n$  is selected in such a way that

- (g) Each member of the population has an equal chance of being selected.
- (h) Each member of the population is given an opportunity to respond to the survey.
- (i) All samples of size  $n$  have the same chance of being selected.
- (j) The probability of selecting any sample is known to be  $\frac{1}{n} \rightarrow rand$ .
- (k) The sample is guaranteed to represent the entire population.

19. In sample surveys, bias can be controlled by all of the following except

- (k) Using a probability or chance sampling procedure.
- (l) Wording questions so they are not confusing or misleading.
- (m) Carefully training and supervising interviewers.
- (n) Prompting respondents so that they give correct responses.
- (o) Reducing non-response and undercoverage.

20. A graduate student a study to determine whether a new activity-based method is better than the traditional lecture of teaching statistics. He found two teachers to help him in his study for one semester. Mr. Dull volunteered to continue teaching with traditional lectures and Ms. Perky agreed to try the new activity-based method. Each teacher planned to teach two sections of approximately forty students each for adequate replication. At the end of the semester, all sections would take the same final exam and their scores would be compared. What is the treatment variable in this study?

- (a) Teacher
- (b) Section of the Course
- (c) Teaching Method
- (d) Final Exam Score
- (e) Student

21. In a study on the effect of reinforcement on learning from programmed text, two experimental treatments are planned: reinforcement given after every frame of programmed text or reinforcement given after every three frames. Which one of the following control groups would serve best in this study?

- (b) A group which does not read the programmed text material.
- (c) A group that reads the programmed material in prose formats.

- (d) A group which reads the programmed material but does not receive reinforcement.
- (e) A group that reads the programmed text material and reinforcement is given at random.
- (f) A group which watches the video of the programmed material.

22. Which of the following are true statements?

- I. Voluntary response samples often underrepresent people with strong opinions.
- II. Convenience samples often lead to undercoverage bias.
- III. Questionnaires with nonneutral wording are likely to have response bias.

- (a) I and II      (b) I and III      (c) II and III      (d) I, II, and III      (e) None of the above

23. To survey the opinions of bleacher fans at Wrigley Field, a surveyor plans to select every one-hundredth fan entering the bleachers one afternoon. Will this result in a random sample?

- (f) Yes, because each bleacher fan has the same chance of being selected.
- (g) Yes, but only if there is a single entrance to the bleachers.
- (h) Yes, because the 99 out of 100 bleacher fans that are not selected will form a control group.
- (i) Yes, because this is an example of systematic sampling, which is a special case of random sampling.
- (j) No, because each fan does not have the same chance of being selected.

24. What fault do all these sampling designs have in common?

- I. The *Wall Street Journal* plans to make a prediction for a presidential election based on a survey of its readers.
- II. A radio talk show asks people to phone in their views on whether the United States should pay off its huge debt to the United Nations.
- III. A police detective interested in determining a sample of high school students and interviews each one about any illegal drug use by the student during the past year.

- (a) All the designs make improper use of stratification.
- (b) All the designs have errors that can lead to strong bias.
- (c) All the designs confuse association with cause and effect.
- (d) None of the designs satisfactorily controls for sampling error.
- (e) None of the designs makes use of chance in selecting a sample.

25. Event A has probability 0.4. Event B has probability 0.5. If A and B are disjoint then the probability that both events occur is:

- (a) 0.0.      (b) 0.1.      (c) 0.2.      (d) 0.9.      (e) None of the above.

26. Event A has probability 0.4. Event B has probability 0.5. If A and B are independent then the probability that both events occur is:

- (a) 0.0.      (b) 0.1.      (c) 0.2.      (d) 0.9.      (e) None of the above.

27. The amount of pollutants a factory dumps into a river is approximately normally distributed with a mean of 2.43 and a standard deviation of 0.88 tons. What is the probability that it dumps more than 3 tons?

- (a) 0.2578      (b) 0.2843      (c) 0.6500      (c) 0.7157      (e) 0.7422

28. Which of the following is **not** true concerning discrete probability distribution?

- (k) The probability of any specific value is between 0 and 1, inclusive.
- (l) The mean of the distribution is between the smallest and largest value in the distribution.
- (m) The sum of all probabilities is 1.
- (n) The standard deviation of the distribution is between -1 and 1.
- (o) The distribution may be displayed using a probability histogram.