

Name \_\_\_\_\_

Date \_\_\_\_\_

Unit I Review

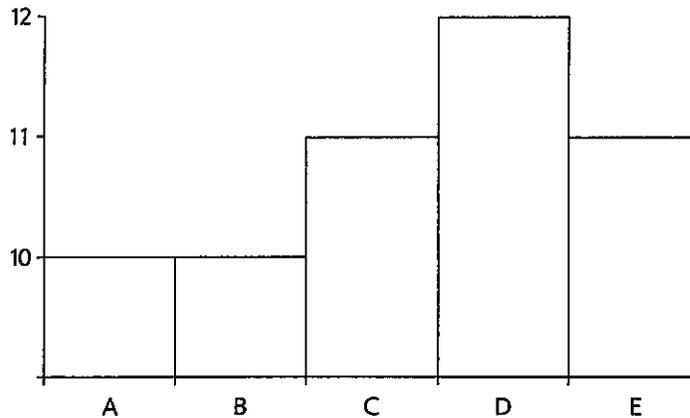
AP Statistics

Part I

1) Which of the following is/are suitable to determine whether a distribution is skewed?

- I. Stemplot
  - II. Histogram
  - III. Cumulative frequency charts
- A. III only
  - B. I and II
  - C. I and III
  - D. II and III
  - E. I, II, and III

2) Which statement is best represented by this histogram?



- A. The histogram is skewed right.
- B. Division D sold more than A and B combined.
- C. Division D made  $1/3$  of all sales.
- D. Division C sold more than division B.
- E. Division C sold twice as much as division B.

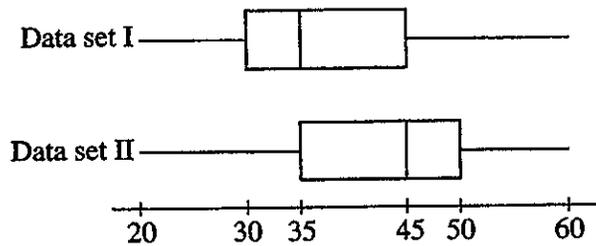
3) Which of the following statements is false?

- A. A symmetric distribution can have outliers.
- B. A skewed left distribution cannot have outliers on the right.
- C. A skewed right distribution may have outliers on the right.
- D. A symmetric distribution may have more outliers than a skewed distribution.
- E. Stemplots are suitable for showing outliers.

4) Which of the following would be greatly affected by outliers?

- I. IQR
  - II. range
  - III. median
- A. I only
  - B. II only
  - C. I and II
  - D. I and III
  - E. II and III

5)



The boxplots shown above summarize two data sets, I and II. Based on the boxplots, which of the following statements about these two data sets CANNOT be justified?

- (A) The range of data set I is equal to the range of data set II.
- (B) The interquartile range of data set I is equal to the interquartile range of data set II.
- (C) The median of data set I is less than the median of data set II.
- (D) Data set I and data set II have the same number of data points.
- (E) About 75% of the values in data set II are greater than or equal to about 50% of the values in data set I.

6)

The following list is ordered from smallest to largest: 25, 26, 26, 30,  $y$ ,  $y$ ,  $y$ , 33, 150. Which of the following statements is (are) true?

- I. The mean is greater than the median
  - II. The mode is 26
  - III. There are no outliers in the data
- a. I only
  - b. I and II only
  - c. III only
  - d. I and III only
  - e. II and III only

7)

Which of the following is not true about histograms?

- A. The bars must touch each other.
- B. The width of each class must be the same.
- C. A histogram for a given set of data may be symmetric or skewed depending on the selection of class width and boundaries.
- D. Histograms may have gaps and clusters.
- E. Histograms may represent categorical data.

8) The mean of a bell-shaped, symmetric distribution is 0, and the standard deviation of the distribution is 5. If 5 is added to each value in the distribution and then the resulting values are each divided by 5, what would be the new mean and standard deviation?

- A. mean = 0, standard deviation = 1
- B. mean = 1, standard deviation = 1
- C. mean = 0, standard deviation = 2
- D. mean = 1, standard deviation = 2
- E. mean = 5, standard deviation = 1

9) The salary schedule of a school district has six columns. The number of employees in each column is shown in the chart that follows. The salary increases as you move from column 1 to column 2 to column 3, and so on.

Column	1	2	3	4	5	6
# of Employees	32	58	110	230	650	90

Which of the following statements is accurate?

- A. The distribution is skewed right with a mean salary that is greater than the median salary.
- B. The distribution is skewed right with a median salary that is greater than the mean salary.
- C. The distribution is skewed left with a mean salary that is greater than the median salary.
- D. The distribution is skewed left with a median salary that is greater than the mean salary.
- E. Without knowing the salary for each column, the relationship between mean and median salaries cannot be determined.

10) The following is the 5-number summary of a set of animal weights:  $a$ , 12,  $b$ , 16, and  $c$ . If a boxplot is drawn to show outliers, what is maximum length of either boxplot whisker?

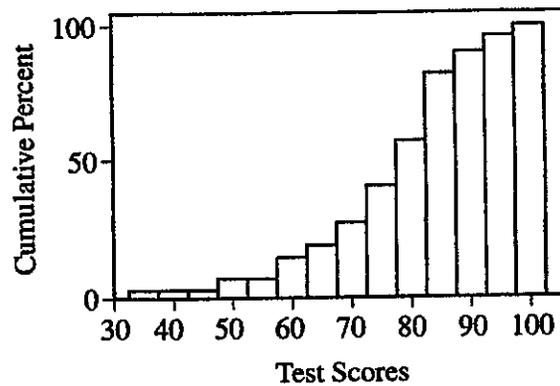
- A. the larger of  $(12 - a)$  and  $(c - 16)$
- B. the larger of  $(b - a)$  and  $(c - b)$
- C. 6
- D. 8
- E. cannot be determined from the given data

11) If two distributions are symmetric and have the same range and the same mean, then which of the following statements are true?

- I. They have the same standard deviation.
  - II. They have the same median.
  - III. They have the same IQR.
- A. I only
  - B. II only
  - C. III only
  - D. II and III
  - E. I, II, and III

12)

**AP STATISTICS  
TEST SCORES**



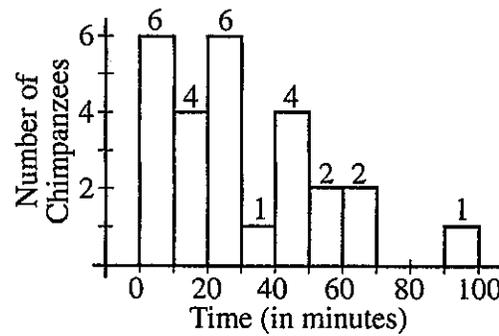
The figure above shows a cumulative relative frequency histogram of 40 scores on a test given in an AP Statistics class. Which of the following conclusions can be made from the graph?

- (A) There is greater variability in the lower 20 test scores than in the higher 20 test scores.
- (B) The median test score is less than 50.
- (C) Sixty percent of the students had test scores above 80.
- (D) If the passing score is 70, most students did not pass the test.
- (E) The horizontal nature of the graph for test scores of 60 and below indicates that those scores occurred most frequently.

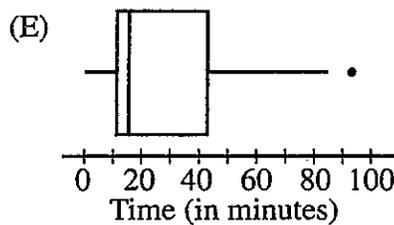
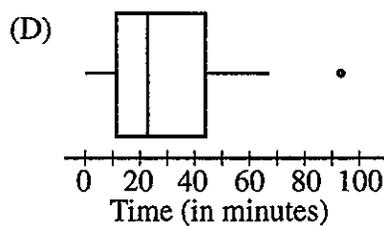
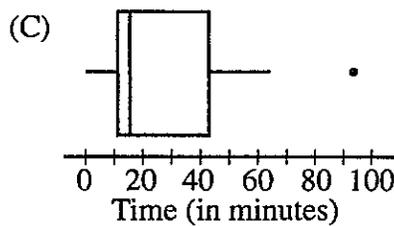
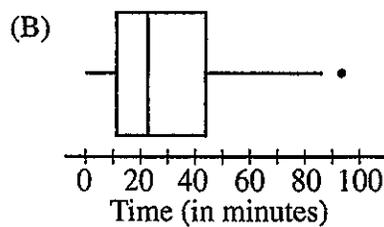
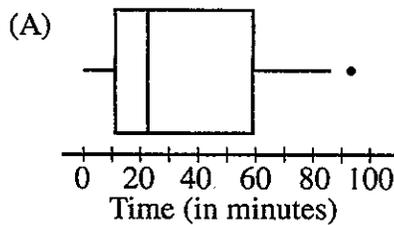
13) The five-number summary for a one-variable dataset is {5, 18, 20, 40, 75}. If you wanted to construct a modified boxplot for the dataset (that is, one that would show outliers if there are any), what would be the maximum possible length of the right side "whisker"?

- a. 35
- b. 33
- c. 5
- d. 55
- e. 53

- 14) The histogram below displays the times, in minutes, needed for each chimpanzee in a sample of 26 to complete a simple navigational task.



It was determined that the largest observation, 93, is an outlier since  $Q_3 + 1.5(Q_3 - Q_1) = 87.125$ . Which of the following boxplots could represent the information in the histogram?



2011 AP<sup>®</sup> STATISTICS FREE-RESPONSE QUESTIONS (Form B)

STATISTICS

SECTION II

Part A

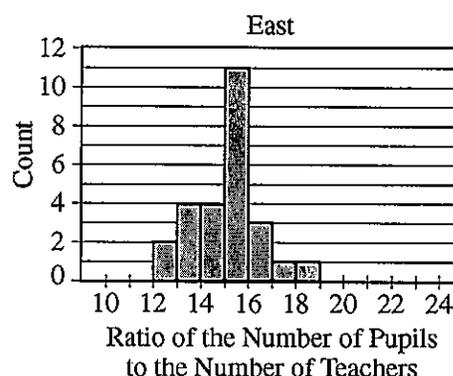
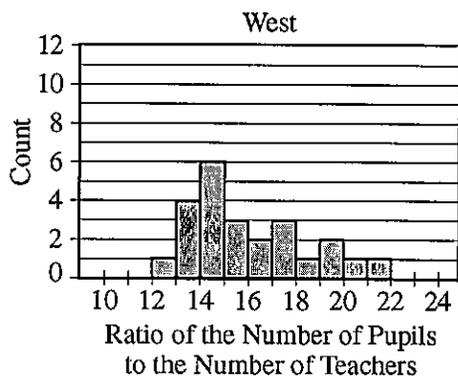
Questions 1-5

Spend about 65 minutes on this part of the exam.

Percent of Section II score—75

**Directions:** Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

1. Records are kept by each state in the United States on the number of pupils enrolled in public schools and the number of teachers employed by public schools for each school year. From these records, the ratio of the number of pupils to the number of teachers (P-T ratio) can be calculated for each state. The histograms below show the P-T ratio for every state during the 2001–2002 school year. The histogram on the left displays the ratios for the 24 states that are west of the Mississippi River, and the histogram on the right displays the ratios for the 26 states that are east of the Mississippi River.



- (a) Describe how you would use the histograms to estimate the median P-T ratio for each group (west and east) of states. Then use this procedure to estimate the median of the west group and the median of the east group.
- (b) Write a few sentences comparing the distributions of P-T ratios for states in the two groups (west and east) during the 2001–2002 school year.
- (c) Using your answers in parts (a) and (b), explain how you think the mean P-T ratio during the 2001–2002 school year will compare for the two groups (west and east).

# 2004 AP<sup>®</sup> STATISTICS FREE-RESPONSE QUESTIONS

## STATISTICS

### SECTION II

#### Part A

#### Questions 1-5

Spend about 65 minutes on this part of the exam.

Percent of Section II grade—75

**Directions:** Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy of your results and explanation.

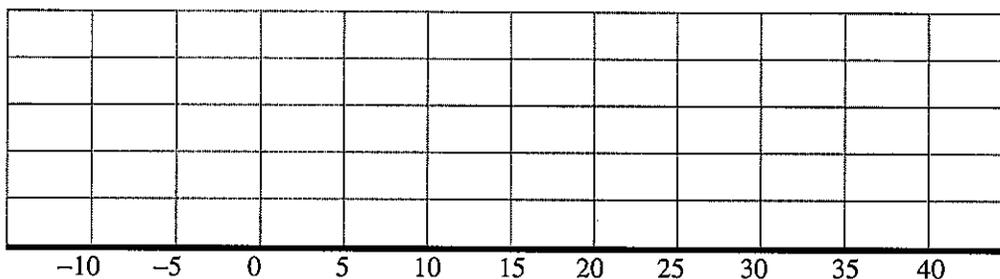
1. A consumer advocate conducted a test of two popular gasoline additives, A and B. There are claims that the use of either of these additives will increase gasoline mileage in cars. A random sample of 30 cars was selected. Each car was filled with gasoline and the cars were run under the same driving conditions until the gas tanks were empty. The distance traveled was recorded for each car.

Additive A was randomly assigned to 15 of the cars and additive B was randomly assigned to the other 15 cars. The gas tank of each car was filled with gasoline and the assigned additive. The cars were again run under the same driving conditions until the tanks were empty. The distance traveled was recorded and the difference in the distance with the additive minus the distance without the additive for each car was calculated.

The following table summarizes the calculated differences. Note that negative values indicate less distance was traveled with the additive than without the additive.

Additive	Values Below $Q_1$	$Q_1$	Median	$Q_3$	Values Above $Q_3$
A	-10, -8, -2	1	3	4	5, 7, 9
B	-5, -3, -3	-2	1	25	35, 37, 40

- (a) On the grid below, display parallel boxplots (showing outliers, if any) of the differences of the two additives.



- (b) Two ways that the effectiveness of a gasoline additive can be evaluated are by looking at either
- the proportion of cars that have increased gas mileage when the additive is used in those cars
  - or
  - the mean increase in gas mileage when the additive is used in those cars.
- i. Which additive, A or B, would you recommend if the goal is to increase gas mileage in the highest proportion of cars? Explain your choice.
- ii. Which additive, A or B, would you recommend if the goal is to have the highest mean increase in gas mileage? Explain your choice.