

Multiple Choice Practice

Test 1 – Organizing Data

Name _____ Per _____

1. An outlier is an observation that

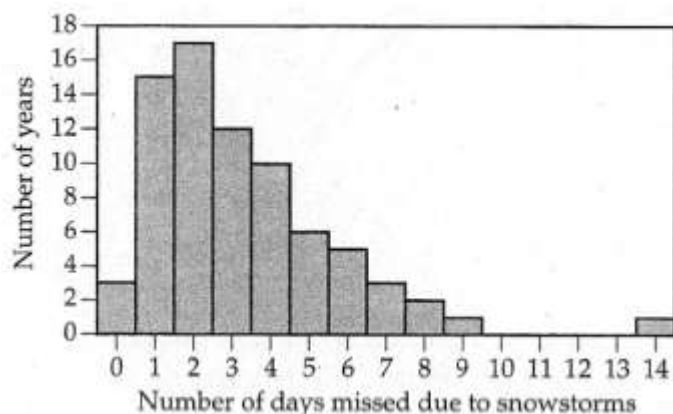
- A. is seen more frequently than the other observations in the data set.
- B. is seen less frequently than the other observations in the data set.
- C. is always smaller than the other observations in the data set.
- D. is always larger than the other observations in the data set.

➡ E. is significantly different from the other observations in the data set.

2. Suppose that a Normal model describes the acidity (pH) of rainwater, and that water tested after last week's storm had a z-score of +1.8. This means that the acidity of that rain

- ➡ A. had a pH 1.8 standard deviations higher than that of average rainwater.
- B. varied with a standard deviation of 1.8.
 - C. had a pH 1.8 higher than average rainfall.
 - D. had a pH 1.8 times that of average rainwater.
 - E. had a pH of 1.8.

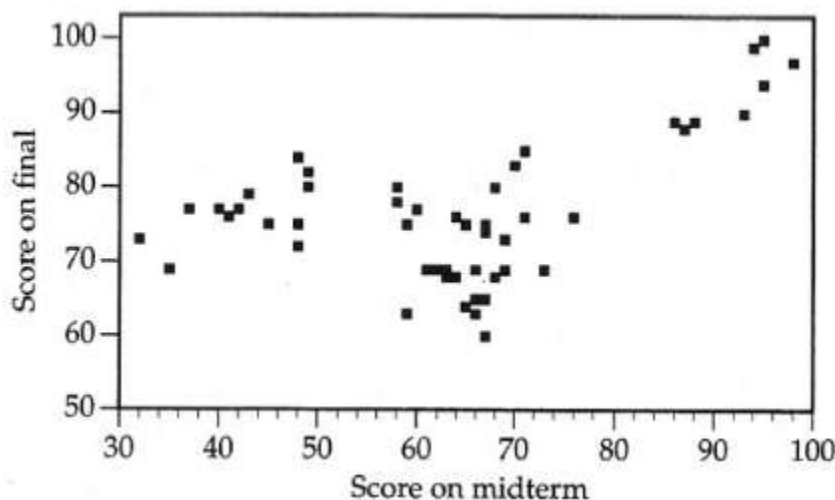
3. In the northern US, schools are sometimes closed during winter due to severe snowstorms. At the end of the school year, schools have to make up for the days missed. The following graph shows the frequency distribution of the number of days missed due to snowstorms per year, using data collected from the past 75 years.



Which of the following should be used to describe the center of this distribution?

- A. Mean, because it is an unbiased estimator.
- ➡ B. Median, because the distribution is skewed.
- C. IQR, because it excludes outliers and includes on the middle 50 percent of data.
- D. First quartile, because the distribution is left skewed.
- E. Standard deviation, because it is unaffected by the outliers.

4. Extra study sessions were offered to students after the MidTerm to help improve their understanding of statistics. Student scores on the MidTerm and the Final Exam were recorded. The following scatterplot shows the Final test scores against the MidTerm test scores.



Which of the following statements correctly interprets the scatterplot?

A. All students have shown significant improvement in the Final Exam scores as a result of the extra study sessions.

B. The extra study sessions were of no help. Each student's Final Exam score was about the same as his or her score on the MidTerm.

C. The extra study sessions further confused the students. All student scores decreased from MidTerm to Final Exam.

➡ D. Students who scored below 55 on the MidTerm showed considerable improvement on the Final Exam; those who scored between 55 and 80 on the MidTerm showed minimal improvement on the Final Exam; and those who scored above 80 on the MidTerm showed almost no improvement on the Final Exam.

E. Students who scored below 55 on the MidTerm showed minimal improvement on the Final Exam; those who scored between 55 and 80 on the MidTerm showed moderate improvement on the Final Exam; and those who scored above 80 on the MidTerm showed considerable improvement on the Final Exam.

5. All but one of the statements below contains a mistake. Which one could be true?

A. The correlation between height and weight is 0.568 inches per pound.

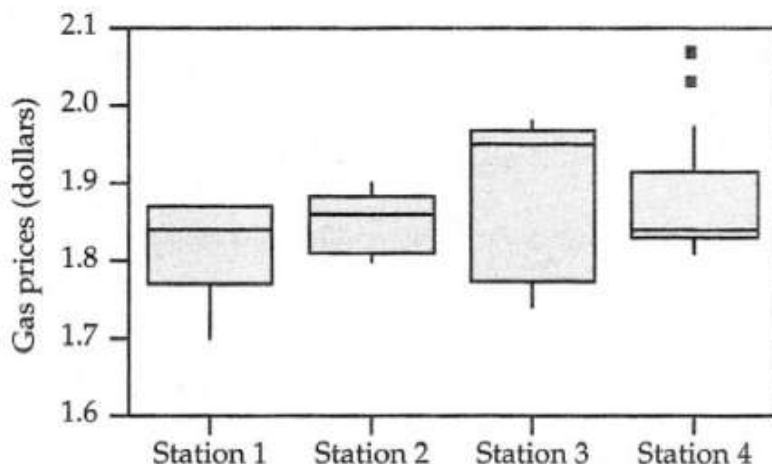
B. If the correlation between blood alcohol level and reaction time is 0.73, then the correlation between reaction time and blood alcohol level is -0.73.

C. The correlation between the breed of a dog and its weight is 0.435.

D. The correlation between gender and age is -0.171.

➡ E. The correlation between weight and length of foot is 0.488.

6. A resident of Auto Town was interested in finding the cheapest gas prices at nearby gas stations. On randomly selected days over a period of one month, he recorded the gas prices (in dollars) at four gas stations near his house. The box plots of gas prices are as follows:



Which station has more consistent gas prices?

- A. Station 1 ☒ B. Station 2 C. Station 3
D. Station 4 E. Station 4 without the outliers

7. A residuals plot is helpful because

- I. it will help us see whether our model is appropriate.
II. it might show a pattern in the data that was hard to see in the original scatterplot.
III. it will clearly identify influential points.

- A. I only B. II only ☒ C. I, II, and III
D. I and III only E. I and II only

8. The observed time (in minutes) it takes a runner to complete a marathon are approximately normally distributed. The z-score for his running time this week was -2. Which one of the following statements is a correct interpretation of his z-score?


- A. This week his time was two minutes lower than his time last week.
B. This week his time was two minutes lower than his best time ever.
C. This week his time was two minutes lower than his average time.
☒ D. This week his time was two standard deviations lower than his average time.
E. This week his time was two standard deviations lower than his time last week.

9. A regression analysis of students' AP Statistics scores and the number of hours they spent doing homework found $r^2=0.32$. Which if these is true?


- I. There's a 32% chance that you'll get the score this model predicts for you.
II. Homework accounts for 32% of your grade in AP Stats.
III. 32% of variation in student test scores can be explained with this model.

- ☒ A. all of these B. none of these C. II only
D. III only E. I and II


10. The relationship between the selling price of a car (in \$1,000) and its age (in years) is estimated from a random sample of cars of a specific model. The relation is given by the formula $\widehat{\text{selling price}} = 15.9 - 0.983(\text{age})$. Which of the following is true from this equation?

- A. For every year the car gets older, the selling price goes down by approximately 9.83 percent.
- B. A new car costs on the average \$9,830.
- C. For every year the car gets older, the selling price drops by approximately \$1,590.
- D. A new car costs \$16,883.
-  E. For every year the car gets older, the selling price drops by approximately \$983.

11. A small kiosk at the Atlanta airport carries souvenirs in the price range of \$3.99 to \$29.99, with a mean price of \$14.75. The airport authorities decide to increase the rent charged for a kiosk by five percent. To make up for the increased rent, the kiosk owner decides to increase the prices of all items by 50 cents. As a result, which of the following will happen?


- A. The mean price and the range of prices will increase by 50¢.
- B. The mean price will remain the same, but the range of prices will increase by 50¢.
- C. The mean price and the standard deviation of prices will increase by 50¢.
-  D. The mean price will increase by 50¢., but the standard deviation of prices will remain the same.
- E. The mean price and the standard deviation will remain the same.

12. On a boxplot, an observation is classified as an outlier if


- A. it is smaller than $\bar{x} - 3s_x$ or larger than $\bar{x} + 3s_x$.
- B. it is larger than $\bar{x} - 3s_x$ and smaller than $\bar{x} + 3s_x$.
- C. it is smaller than Q_1 or larger than Q_3 .
-  D. it is smaller than $Q_1 - (1.5)(IQR)$ or larger than $Q_3 + (1.5)(IQR)$.
- E. it is larger than $Q_1 - (1.5)(IQR)$ and smaller than $Q_3 + (1.5)(IQR)$.

13. The relation between studying time (in hours) and grade on a semester exam (0-100) in a random sample of students in math class was found to be $\widehat{\text{grade}} = 50.25 + 10.2(\text{hours})$.

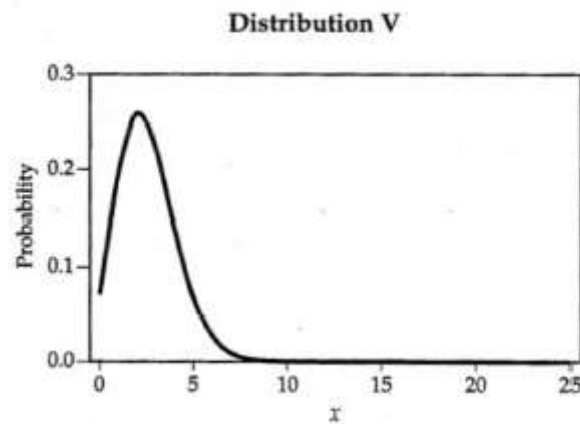
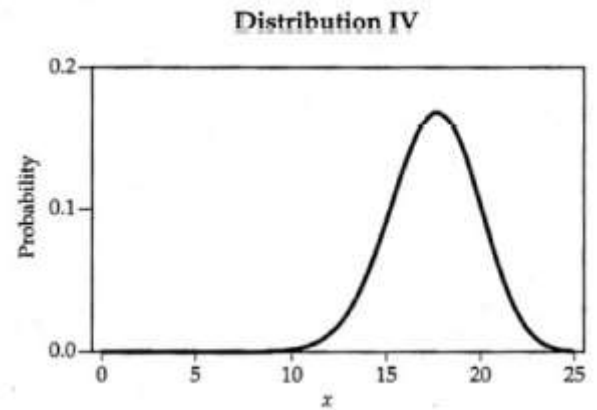
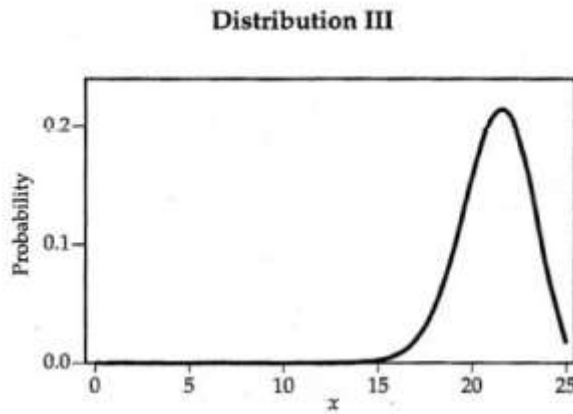
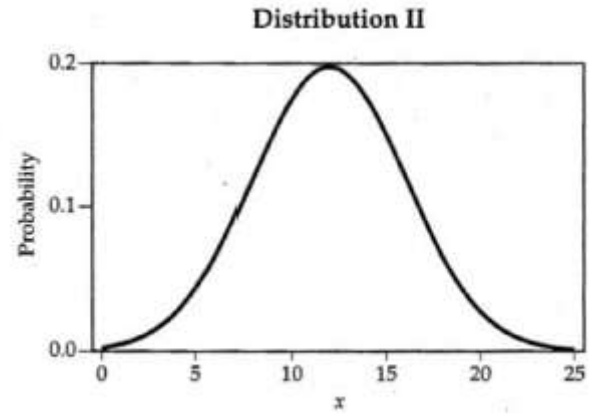
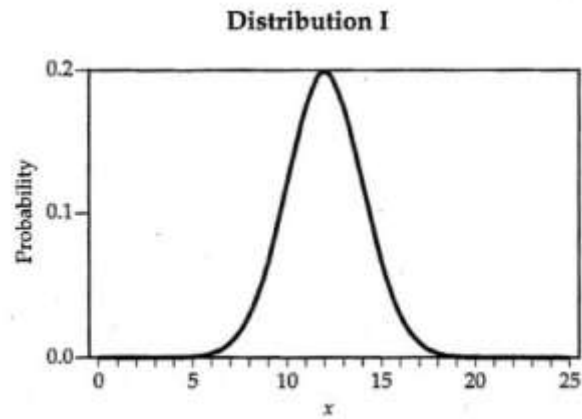
How will a student's grade be effected is she studies for two hours?

- A. Her grade will go up by 40.05 points on average.
-  B. Her grade will go up by 20.4 points on average.
- C. Her grade will remain unchanged.
- D. Her grade will go down by 10.2 points on average.
- E. It cannot be determined from the information given.

14. A numerical measure of the direction and strength of a linear association is its

- A. association.
- B. slope.
- C. predicted value.
- D. intercept.
-  E. correlation.

15. Which one of the following distributions could have a mean of approximately 12 and a standard deviation of approximately 2?

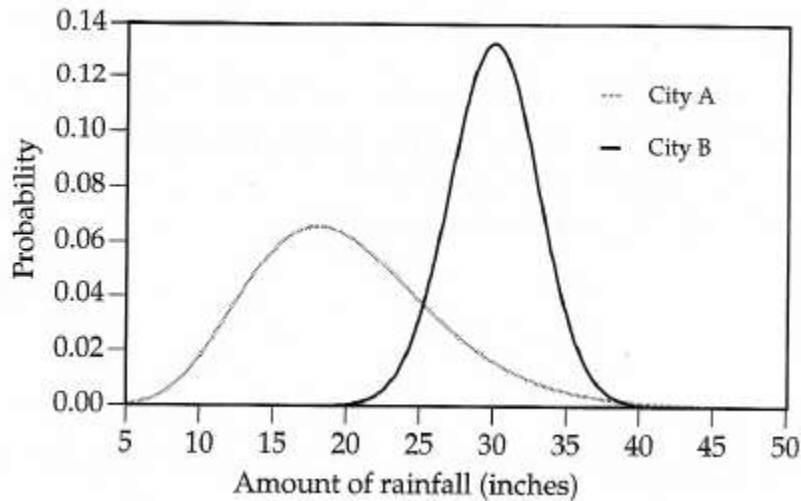


A. Distribution I
D. Distribution IV

B. Distribution II
E. Distribution V

C. Distribution III

16. The following graph summarizes data collected on annual rainfall in two cities for the past 150 years. Which of the following conclusions can be made from this graph?



A. The cities have different mean annual rainfalls, but the range of their annual rainfalls is approximately the same.

➡ B. On average, City B gets more rain than City A, but it has a smaller range of annual rainfall.

C. On average, City B gets more rain than City A, and it has a larger range of annual rainfall.

D. On average, City A gets more rain than City B, but it has a smaller range of annual rainfall.

E. On average, City A gets more rain than City B, and it has a larger range of average rainfall.

17. Data were collected on two variables, X and Y , and a least squares regression line was fitted to the data. The estimated equation for this data is

$$\hat{y} = -2.29 + 1.70x.$$

One point has $x = 5$, $y = 6$.

What is the residual for this point?

A. 7.91

B. 6.21

C. 0.21

➡ D. -0.21

E. -2.91

18. A company has 400 employees. Their mean income is \$20,500, and the standard deviation of their incomes is \$3,750. The distribution of incomes is normally distributed. How many of the 400 employees do you expect to have an income of between \$13,000 and \$28,000?

A. 50

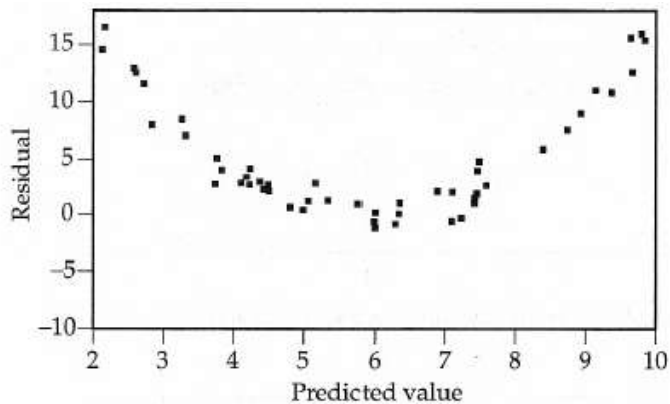
B. 100

C. 200

D. 390

➡ E. 400

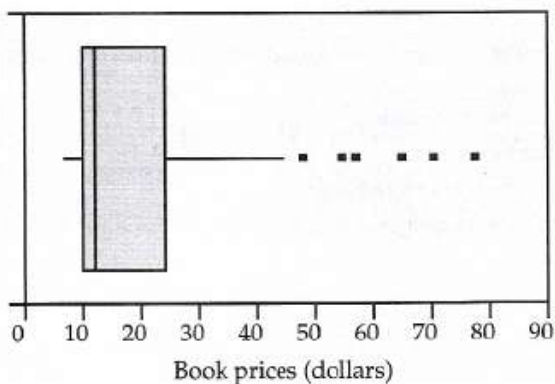
19. Sixty pairs of measurements were taken at random to estimate the relation between variables X and Y . A least squares line was fitted to the collected data. The resulting residual plot is as follows:



Which of the following conclusions is appropriate?

- ☐ A. A line is an appropriate model to describe the relation between X and Y .
- ☒ B. A line is not an appropriate model to describe the relation between X and Y .
- ☐ C. The assumption of The Law of Averages has been violated.
- ☐ D. The variables X and Y are not related at all.
- ☐ E. Without knowing what variables X and Y are, there is no way to form an appropriate conclusion about a possible linear connection between them.

20. The following boxplot summarizes the prices of books at a bookstore.



Which of the following is true about the prices of the books at this store?

- ☐ A. This store carries more books priced above the mean than below the mean.
- ☒ B. This store carries more books priced below the mean than above the mean.
- ☐ C. This store carries about the same number of books at different prices in the entire price range.
- ☐ D. The mean price of books at this store is the same as the median price of books.
- ☐ E. The mean price of books at this store is lower than the median price of books.