

AP Statistics 1st Nine Weeks Practice Exam

KEY

1. We collect these data from 50 male students. Which variable is categorical?

- A) eye color
- B) head circumference
- C) hours of homework last week
- D) number of cigarettes smoked daily
- E) number of TV sets at home

2. Which of those variables is most likely to be bimodal?

- A) eye color
- B) head circumference
- C) hours of homework last week
- D) number of cigarettes smoked daily
- E) number of TV sets at home

the more correct answer is D

smoker vs. nonsmoker

3. Which of those variables is most likely to follow a Normal model?

- A) eye color
- B) head circumference
- C) hours of homework last week
- D) number of cigarettes smoked daily
- E) number of TV sets at home



4. The mean number of hours worked for the 30 males was 6, and for the 20 females was 9. The overall mean number of hours worked ...

- A) is 6.5
- B) is 7.2
- C) is 7.5
- D) is none of these.
- E) cannot be determined.

$$\frac{30(6) + 20(9)}{50}$$

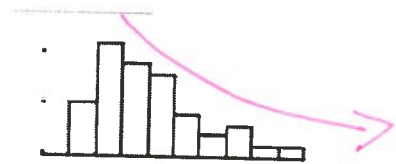
5. We might choose to display data with a stemplot rather than a boxplot because a stemplot

- I. reveals the shape of the distribution.
 - II. is better for large data sets.
 - III. displays the actual data.
- stem + leaf shows shape & all values
- box plot - only shows 5 # summary

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

6. Which is true of the data whose distribution is shown?

- I. The distribution is skewed to the right. ✓
- II. The mean is probably smaller than the median. ✗
- III. We should summarize with mean and standard deviation. ✓



mean is pulled up

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

7. Suppose that a Normal model describes fuel economy (miles per gallon) for automobiles and that a Saturn has a standardized score (z-score) of +2.2. This means that Saturns ...

- A) get 2.2 miles per gallon.
- B) get 2.2 times the gas mileage of the average car.
- C) get 2.2 mpg more than the average car.
- D) have a standard deviation of 2.2 mpg.
- E) achieve fuel economy that is 2.2 standard deviations better than the average car.

z = # of standard deviations

8. Suppose a Normal model describes the number of pages printer ink cartridges last. If we keep track of printed pages for the 47 printers at a company's office, which must be true?

- I. The page counts for those ink cartridges will be normally distributed.
- II. The histogram for those page counts will be symmetric.
- III. 95% of those page counts will be within 2 standard deviations of the mean.

- A) none B) I only C) II only D) II and III E) I, II, and III

9. Which is true?

- I. Random scatter in the residuals indicates a model with high predictive power.
- II. If two variables are very strongly associated, then the correlation between them will be near +1.0 or -1.0.
- III. The higher the correlation between two variables the more likely the association is based in cause and effect.

quantitative

we can't assume this, we need to know it is linear
Never

- A) none B) I only C) II only D) I and II only E) I, II, and III

10. It's easy to measure the circumference of a tree's trunk, but not so easy to measure its height. Foresters developed a model for ponderosa pines that they use to predict the tree's height (in feet) from the circumference of its trunk (in inches):

$$\ln \hat{h} = -1.2 + 1.4(\ln C)$$

A lumberjack finds a tree with a circumference of 60"; how tall does this model estimate the tree to be?

- A) 5' B) 11' C) 19' D) 83' E) 93'

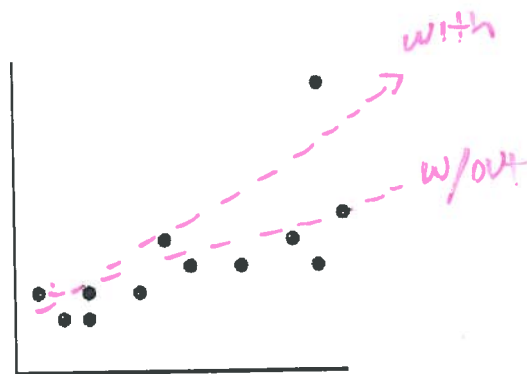
11. Two variables that are actually not related to each other may nonetheless have a very high correlation because they both result from some other, possibly hidden, factor. This is an example of

- A) leverage. B) a lurking variable. C) extrapolation. D) regression. E) an outlier

hidden may be causing relationship
predicting

12. If the point in the upper right corner of this scatterplot is removed from the data set, then what will happen to the slope of the line of best fit (b) and to the correlation (r)?

- A) both will increase.
B) both will decrease.
C) b will increase, and r will decrease.
D) b will decrease, and r will increase.
E) both will remain the same.



**estimates*

$$e = y - \hat{y}$$

positive - higher than pred.
negative - lower than pred.

13. When using midterm exam scores to predict a student's final grade in a class, the student would prefer to have a

- A) positive residual, because that means the student's final grade is higher than we would predict with the model.
- B) positive residual, because that means the student's final grade is lower than we would predict with the model. X
- C) residual equal to zero, because that means the student's final grade is exactly what we would predict with the model. *could also be reasonable*
- D) negative residual, because that means the student's final grade is lower than we would predict with the model. *wouldn't want that*
- E) negative residual, because that means the student's final grade is higher than we would predict with the model. Y

14. Which of the following is not a goal of re-expressing data?

4 GOALS FROM CH 9
POWERPOINT

- A) Make the distribution of a variable more symmetric.
- B) Make the spread of several groups more alike.
- C) Make the form of a scatterplot more nearly linear.
- D) Make the scatter in a scatterplot spread out evenly rather than following a fan shape.
- E) All of the above are goals of re-expressing data.

15. R-sq is a measure of ...

SCRIPT we follow

- A) the percentage of the accuracy of the regression equation.
- B) the proportion of the variability in the response variable that is explained by the explanatory variable.
- C) the probability that the regression line makes a correct prediction.
- D) the change in the y-variable that corresponds with the change in the x-variable.
- E) the initial predicted starting point of the response variable when x is zero.

16. The relationship between the number of hours a person practices a task and the time it takes them to complete the task is calculated to have R-sq = 56.7%. The value of the correlation coefficient is

- A) -0.753
- B) -0.238
- C) 0.238
- D) 0.753
- E) 2.38

the more practice,
the easier it is to
complete task



17. In predicting the growth of the volume of a small bay by measuring the height of the water at a dock, a

researcher is using a model of $\sqrt[3]{\text{volume}} = 2.34 + 4.56(\text{height})$ where height is measured in m and volume cubic miles. If the height rises to 3.45 m, what is the predicted volume?

- A) 2.62 m^3
- B) 18.1 m^3
- C) 5902 m^3
- D) $7 \times 10^7 \text{ m}^3$
- E) $1.2 \times 10^{18} \text{ m}^3$

