2.6 Review

Name_____

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Select the most appropriate answer.

1) The y-intercept is the 1) _____
   A) smallest value for the residual sum of squares.
   B) point where the regression line crosses the x-axis.
   C) predicted value of y when x = 0.
   D) predicted value of y.
   E) change in the predicted value of y per unit increase in x.

2) The slope is the 2) _____
   A) predicted value of y.
   B) smallest value for the residual sum of squares.
   C) point where the regression line crosses the y-axis.
   D) predicted value of y when x = 0.
   E) change in the predicted value of y per unit increase in x.

3) The prediction error for an observation, which is the difference between the actual value and the predicted value of the response variable, is called _____. 3) _____
   A) a correlation
   B) an intercept
   C) an outlier
   D) an extrapolation
   E) a residual

Solve the problem.

4) The data below are the final exam scores of 10 randomly selected statistics students and the number of hours they studied for the exam. Find the equation of the regression line for the given data. 4) _____

Hours, x	3	5	2	8	2	4	4	5	6	3
Scores, y	65	80	60	88	66	78	85	90	90	71

   A) $\hat{y} = 5.044x + 56.11$
   B) $\hat{y} = -56.11x - 5.044$
   C) $\hat{y} = 56.11x - 5.044$
   D) $\hat{y} = -5.044x + 56.11$

5) The data below are the final exam scores of 10 randomly selected statistics students and the number of hours they studied for the exam. What is the best predicted value for y given x = 3? 5) _____

Hours, x	3	5	2	8	2	4	4	5	6	3
Scores, y	65	80	60	88	66	78	85	90	90	71

   A) 69.78
   B) 70.15
   C) 72.23
   D) 71.24

6) The data below are the final exam scores of 10 randomly selected statistics students and the number of hours they studied for the exam. Determine the residual of a data point for which x = 3 and y = 71 6) _____

Hours, x	3	5	2	8	2	4	4	5	6	3
Scores, y	65	80	60	88	66	78	85	90	90	71

A) 71.24 B) 0.24 C) -0.24 D) 71

7) A county real estate appraiser wants to develop a statistical model to predict the appraised value of houses in a section of the county called East Meadow. One of the many variables thought to be an important predictor of appraised value is the number of rooms in the house. Consequently, the appraiser decided to fit the simple linear regression model: $\hat{y} = 74.80 + 22.86x$ 7) _____
where y = appraised value of the house (in $thousands) and x = number of rooms.
Range of the x-values: 5 - 11 Range of the y-values: 160 - 300
Give a practical interpretation of the estimate of the slope of the least squares line.
A) For a house with 0 rooms, we estimate the appraised value to be $74,800.
B) For each additional room in the house, we estimate the appraised value to increase $74,800.
C) For each additional dollar of appraised value, we estimate the number of rooms in the house to increase by 22.86 rooms.
D) For each additional room in the house, we estimate the appraised value to increase $22,860.

8) A county real estate appraiser wants to develop a statistical model to predict the appraised value of houses in a section of the county called East Meadow. One of the many variables thought to be an important predictor of appraised value is the number of rooms in the house. Consequently, the appraiser decided to fit the simple linear regression model: $\hat{y} = 74.80 + 22.86x$ 8) _____
where y = appraised value of the house (in $thousands) and x = number of rooms.
Range of the x-values: 5 - 11 Range of the y-values: 160 - 300
give a practical interpretation of the estimate of the y-intercept of the least squares line.
A) There is no practical interpretation, since a house with 0 rooms is nonsensical.
B) For each additional room in the house, we estimate the appraised value to increase $74,800.
C) We estimate the base appraised value for any house to be $74,800.
D) For each additional room in the house, we estimate the appraised value to increase $19,720.

Provide an appropriate response.

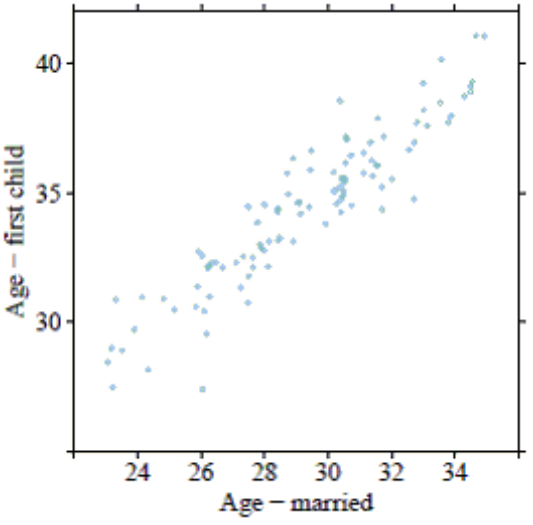
9) A regression line for predicting the selling prices of homes in Chicago is $\hat{y} = 168 + 102x$, where x is the square footage of the house. Interpret the residual for a house with 1800 square feet that recently sold for $200,000. 9) _____
A) The house sold for $16,232 more than was to be expected from the regression equation.
B) The house sold for $16,064 more than was to be expected from the regression equation.
C) The house sold for $16,400 less than was to be expected from the regression equation.
D) The house sold for $16,232 less than was to be expected from the regression equation.
E) The house sold for $16,400 more than was to be expected from the regression equation.

Select the most appropriate answer.

10) Among the possible lines that can go through data points in a scatterplot, the regression line results from the least squares method and has the smallest value for the _____. 10) _____

A) residual sum of squares
B) residual sum
C) slope
D) correlation
E) intercept

The scatterplot below shows the relationship between the ages of women when they first married and the ages when they had their first child. The correlation coefficient between the values is 0.9315.



The regression equation for the data is

$\widehat{\text{Age-first child}} = 6.404 + 0.955 \cdot (\text{Age-married})$

11) Would it be appropriate to say that the age at which a woman first marries causes her to have her first child after that age? 11) _____

A) Yes, the scatterplot shows a strong linear association, so the older a woman is when she first marries, the older she will be when she has her first child.
B) No, since the correlation coefficient does not equal 1, we cannot conclude causation.
C) No, correlation never implies causation.
D) Yes, since the correlation coefficient is close to 1, we can conclude causation.

12) Can we use the regression equation to predict the age at which a woman has her first child for a woman who first married at the age of 50? 12) _____

A) Yes, we can make a prediction because the scatterplot shows a strong positive linear relationship.
B) No, we cannot make a prediction because the correlation coefficient does not equal 1.
C) Yes, we can always make predictions once we have a regression equation.
D) No, we cannot make a prediction because a woman who marries at age 50 is outside the range of our data. We would be extrapolating.