

## Week 13 Quiz: Correlation and Linear Regression

### Problem 1

A researcher collects data on the number of hours students studied for an exam and their final exam scores. After calculating the correlation coefficient, she finds  $r = 0.85$ . What is the correct interpretation?

- a) Studying for more hours causes a student to get a higher exam score.
- b) There is a strong positive linear relationship between study hours and exam scores.
- c) 85% of the variation in exam scores is explained by the number of study hours.
- d) There is a weak positive relationship between the two variables.

**Answer: b**

### Problem 2.

An automotive engineer analyzes a dataset of car models and finds a correlation of  $r = -0.78$  between a car's weight and its fuel efficiency (miles per gallon). This means:

- a) Heavier cars tend to have lower fuel efficiency.
- b) Heavier cars tend to have higher fuel efficiency.
- c) The weight of a car explains 78% of its fuel efficiency.
- d) There is no relationship between weight and fuel efficiency.

**Answer: a**

### Problem 3

A study finds a significant positive correlation between the number of firefighters at a fire and the amount of damage the fire causes ( $r = 0.65$ ). A news headline reads "Firefighters Cause More Damage!" The flaw in this conclusion is:

- a) The correlation coefficient is not strong enough.
- b) The correlation is positive, but it should be negative.
- c) The correlation does not imply causation; a larger fire requires more firefighters and causes more damage.
- d) The correlation was calculated incorrectly.

**Answer: c**

#### Problem 4

A regression analysis is performed to predict the price of a textbook (in \$) based on its number of pages. The slope is 0.08. What is the correct interpretation?

- a) For each additional page, the price increases by \$0.08.
- b) For each additional dollar, the book gains 0.08 pages.
- c) The book's base price is \$0.08.
- d) For every 8 pages, the price increases by \$1.

**Answer: a**

#### Problem 5

A botanist fits a regression line to predict the height of a certain tree species (in feet) from its age (in years). The slope is 1.5. Which statement is correct?

- a) A 3-year-old tree is predicted to be 4.5 feet tall.
- b) For every additional year, the tree's height increases by 1.5 feet.
- c) The correlation between age and height is 1.5.
- d) A 10-foot tree is predicted to be 15 years old.

**Answer: b**

#### Problem 6

A real estate agent builds a model to predict a house's selling price based on its age. The slope is -\$2,000. This implies:

- a) For every year older a house is, its predicted selling price decreases by \$2,000.
- b) For every \$1,000 less a house costs, it is 2 years older.
- c) Newer houses are harder to sell.
- d) The correlation between price and age is positive.

**Answer: a**

#### Problem 7

A model predicting a person's shoe size from their monthly income has an  $R^2$  value of 0.03. This means:

- a) The correlation between shoe size and income is strong.
- b) Only 3% of the variation in shoe size is explained by income; the model is essentially useless.
- c) Income causes a 3% change in shoe size.
- d) The model makes correct predictions 3% of the time.

**Answer: b)**

### **Problem 8**

The regression equation to predict a car's fuel efficiency (mpg) from its weight (lbs) is: Predicted mpg =  $40 - 0.008(\text{Weight})$ . What is the predicted mpg for a car that weighs 3000 lbs?

- a) 16 mpg
- b) 24 mpg
- c) 40 mpg
- d) 64 mpg

**Answer: a**

### **Problem 9**

In a regression predicting house price (in thousands of dollars) from square footage, the slope is 0.15. If square footage increases by 200, what is the predicted change in price?

- a) \$30
- b) \$300
- c) \$3,000
- d) \$30,000

**Answer: d)**

### **Problem 10**

A 95% confidence interval for the slope of the regression of blood pressure on medication dosage is  $[-3.5, -1.2]$ . What can we conclude?

- a) There is no significant relationship between dosage and blood pressure.

b) We are 95% confident that for each unit increase in dosage, blood pressure decreases between 1.2 and 3.5 units.

c) 95% of dosages will lower blood pressure by between 1.2 and 3.5 units.

d) The correlation coefficient is negative.

**Answer: b)**

### **Problem 11**

If the correlation between two variables is zero, what is the slope of the best-fit regression line?

a) Positive

b) Negative

c) Zero

d) Undefined

**Answer: c)**

### **Problem 12**

A regression is run to see if restaurant bill amount predicts tip amount. The t-statistic for the slope is 4.56 and the p-value is 0.0001. What is the conclusion at  $\alpha=0.05$ ?

a) Fail to reject  $H_0$ ; bill amount is not a significant predictor of tip.

b) Reject  $H_0$ ; there is significant evidence that bill amount is a useful predictor of tip.

c) Reject  $H_0$ ; bill amount causes the tip amount to increase.

d) The R-squared value must be very high.

**Answer: b)**

### Problem 13

For a simple linear regression, the t-test statistic for the slope ( $H_0: \beta_1=0$ ) is directly related to:

- a) The y-intercept.
- b) The sample size and the correlation coefficient.
- c) The mean of the x-variable.
- d) The confidence level.

**Answer B.**

### Problem 14

In a hypothesis test for the slope in a regression model, what is the typical null hypothesis ( $H_0$ )?

- a)  $H_0: \beta_1 > 0$
- b)  $H_0: \beta_1 = 0$
- c)  $H_0: \beta_1 < 0$
- d)  $H_0: r = 0$

**Answer: B**

### Problem 15

A 95% CI for the slope relating hours of video games played to GPA is  $[-0.1, 0.3]$ . What does this tell you about the relationship?

- a) Playing video games causes a slight increase in GPA.
- b) Playing video games causes a slight decrease in GPA.
- c) There is no evidence of a significant linear relationship (the slope could be zero).
- d) The R-squared value is very high.

**Answer C**