What is Linear Regression?

Worksheet

Linear regression models the relationship between x and y with the equation = a + bx, finding the line of best fit using the least-squares method to minimize prediction errors.

$$\hat{y} = b_0 + b_1 x$$

Questions

1. In = a + bx, what does b represent?

A) The y-intercept
B) The slope
C) The correlation
D) The residual

2. A regression line is = 3 + 2x. What is when x = 4?

A) 9
B) 11
C) 14
D) 8

3. What method finds the best-fit regression line?

A) Maximum likelihood
B) Least squares
C) Bayesian inference
D) Random sampling

4. What does a residual measure?

A) The slope of the line
B) The correlation coefficient
C) The difference between observed and predicted y
D) The x-intercept

5. A regression equation is = 2 + 3x. Predict y when x = 5.

6. Data points (1,3), (2,5), (3,7). Find the slope (b) and intercept (a).

7. Using = 1 + 2x from the example above, predict y when x = 4.

8. Define: What is linear regression?

9. Define: What does the slope b represent?

10. Define: What does the intercept a represent?

Answer Key

1. B) The slope - b is the slope - how much y changes per unit of x.

2. B) 11 - = 3 + 2(4) = 11.

3. B) Least squares - Least squares minimizes the sum of squared residuals.

4. C) The difference between observed and predicted y - Residual = observed y predicted y.

5. = 2 + 3(5) = 2 + 15 = 17

6. x = 2, = 5 Deviations: (-1,-2), (0,0), (1,2) (x-x)(y-) = 2 + 0 + 2 = 4 (x-x) = 1 + 0 + 1 = 2 b = 4/2 = 2 a = - bx = 5 - 2(2) = 1 Result: = 1 + 2x

7. = 1 + 2(4) = 1 + 8 = 9

8. A method for fitting a straight line = a + bx that best predicts y from x by minimizing squared errors.

9. The average change in y for each one-unit increase in x.

10. The predicted value of y when x = 0.