

1: Regression (think about)

(a) Reason:

When asked to state the simple linear regression model, a student wrote:

$$E(Y_i|X_i) = \beta_0 + \beta_1 X_i + u_i$$

Do you agree?

(b) Reason:

On a sample of $n = 50$ families:

- the average length of the fathers was 67.7 inches; the standard deviation was 2.74 inches.
- the average height of the sons was 68.7 inches; the SD was 2.81 inches.

Let us suppose you have estimated the regression line of sons height on fathers height. What is the regression prediction when the height of the father is 77.7? Explain.

(c) Reason:

Compare the following regression equations:

- $Y_i = \beta_0 + \beta_1 X_i + u_i$
- $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i$
- $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i$
- $Y_i = \beta_0 + \beta_1 X_i$
- $\hat{Y}_i = \beta_0 + \beta_1 X_i$
- $E(Y_i|X_i) = \beta_0 + \beta_1 X_i$
- $E(Y_i|X_i) = \hat{\beta}_0 + \hat{\beta}_1 X_i$
- $E(\hat{Y}_i|X_i) = \hat{\beta}_0 + \hat{\beta}_1 X_i$
- $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + u_i$
- $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i$
- $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$

Are they all correct? Comment your answer for each equation.

(d) Reason:

The regression function relating production output by an employee after taking a training program (Y) to the production output before the training program (X) is:

$$E(Y|X) = 20 + .95X$$

where X ranges from 40 to 100. An observer concludes that the training program does not raise production average on the average because β_1 is not greater than 1. Comment.

2: Regression (compute)

(e) Solve:

Pearson and Lee in a paper published in 1908, studying the relation between son's height on father's height, observed the following statistics: the average length of the fathers was 67.7 inches with a standard deviation equal to 2.74 inches; the average height of the sons was 68.7 inches with a standard deviation equal to 2.81 inches; the correlation was 0.501.

- estimate the regression line of son's height on father's height;
- interpret the estimated coefficients;
- estimate the variance of the regression;
- compute the R^2 index.