Introducing linear regression

StatPREP Class Lesson

Showing various sorts of relationships for discussion

Open up the linear regression Little App. (See footnote[1]). In this activity we'll explore three different data sets available through the app. For each data set, the variable indicated below plays the role of the response variable. Some explanatory variables are listed as sub-points.

[1] https://dtkaplan.shinyapps.io/LA_linear_regression/

1. Data set: NHANES2. Response variable: BMI. It's important for students to know what BMI is. Explanation from the CDC & BMI calculator for students.

    - age (r = 0.5 reasonable scatterplot to assume linearity)
    - income (r = -0.07) shows a very diffuse scatter plot but also helps demo the app to students.
    - pulse: weak relationship
    - systolic: weak-to-moderate relationship
    - diastolic: has outliers
    - sleep_hour: weak-to-moderate. But has a negative relationship

2. Data set: CPS85. Response variable: wage

    - age
    - education

3. Data set: Births_2014 Response variable: mother's age

    - father's age. Moderate size correlation. Ask what it means

Open-ended exploring

Consider systolic blood pressure from the NHANES2 data.

Background: Explain to students what is the difference between the systolic and diastolic blood pressure. Each time the heart beats, the blood pressure in the arteries goes up. It quickly rises to a maximum and then decays until the next beat. Systolic is the maximum blood pressure each beat, diastolic the minimum. The "pulse pressure" is the difference between the two. See this site on blood pressure.

Tasks

1. Determinine three explanatory variables that are predictive of systolic blood pressure.
2. For each of the three, list the strength of the relationship both as a fraction of the variation explained as as the change in systolic blood pressure per unit change of the explanatory variable.

3. Then check whether those three explanatory variables explain diastolic blood pressure as well. Which of systolic or diastolic blood pressure is better explained by the explanatory variables?

Diamonds

Similar to the above, but predict the price of a diamond.

Version 0.2, 2019-05-28, Helen Burn, Word version