Regression Inference Applications

YOUR NAME

09 November, 2020

Exercises

1. We noticed that the 95% prediction interval was much wider than the 95% confidence interval. In words, explain why this is.

2. Beer and blood alcohol content

Many people believe that gender, weight, drinking habits, and many other factors are much more important in predicting blood alcohol content (BAC) than simply considering the number of drinks a person consumed. Here we examine data from sixteen student volunteers at Ohio State University who each drank a randomly assigned number of cans of beer. These students were evenly divided between men and women, and they differed in weight and drinking habits. Thirty minutes later, a police officer measured their blood alcohol content (BAC) in grams of alcohol per deciliter of blood. The data is in the `openintro` package under `bac`.

   a. Create a scatterplot for cans of beer and blood alcohol level.
   b. Describe the relationship between the number of cans of beer and BAC.
   c. Write the equation of the regression line. Interpret the slope and intercept in context.
   d. Do the data provide strong evidence that drinking more cans of beer is associated with an increase in blood alcohol? State the null and alternative hypotheses, report the p-value, and state your conclusion.
   e. Build a 95% confidence interval for the slope and interpret it in the context of your hypothesis test from part d.
   f. Suppose we visit a bar, ask people how many drinks they have had, and also take their BAC. Do you think the relationship between number of drinks and BAC would be as strong as the relationship found in the Ohio State study?
   g. Predict the average BAC after two beer and build a 90% confidence interval around that prediction.
   h. Repeat except build a 90% prediction interval and interpret.
   i. Plot the data points with a regression line, confidence band, and prediction band.

3. Suppose I build a regression fitting a response variable to one predictor variable. I build a 95% confidence interval on β_1 and find that it contains 0, meaning that a slope of 0 is feasible. Does this mean that the response and the predictor are independent?