Week 2 - AYU - Pod

Contents

Questions

Adding code chunks (Ctrl + Alt + I) below of each questions to answer the questions.

1. Run the all codes in the Section 1 and show all the results.

In the next questions, we will examine the hospital cost of patients in Wisconsin in 2003 using patients'
information such as their age, gender, and their length of stay at the hospital. The data and the variables
description can be downloaded at the below links.

- Hospital Costs

- Variable Description

2. Regress the total charge (`TOTCHG`) on age and length of stay. What is the R-squared of the regression?
   With significant level of 5%, is there any variable not significant based on coefficient p-values of the
   model?

3. Use the model to predict the total charge of a 13-year-old that stays a week at a hospital.

4. Adding variable `GENDER` to the regression. Write the equations of this new linear model. Use the model
   to predict the total charge of a 13-year-old female that stays a week at a hospital.

5. Does adding `AGE` and `APRDRG` to the model in Question 4 make it more significant? Use the F-test to
   address the question.

6. Could you improve the model in 3 using the methods discussed for the term life dataset?