STA302 Lec5101, Methods of Data Analysis 1
Module 7: Problem set

October 22, 2024

1 Basic learning objective practice

Problem.

A multiple linear model is fit to a response using 7 predictors from a dataset with 30 observations. This model has an SST of 1500, and an SSreg of 900. Compute the adjusted and unadjusted coefficients of determination.

Problem.

To check whether there is a linear relationship among at least one predictor in a model using response, predictor x_j was chosen to be the random response variable X_j in a new regression model, using the remaining predictors as predictors. This predictor model was fit on 40 observations and the sample variance of X_j is 5.63. If the residual sum of squares for this predictor model for X_j is 40.82, what would be the variance inflation factor of predictor x_j?

Problem.

True or False: if the VIF for a predictor is 2, then the variance of the estimated coefficient for this predictor is not inflated.

Problem.

Run the following in R:

```
dataset = read.table("https://gattonweb.uky.edu/sheather/book/docs/datasets/defects.txt",
header = T)
fit = lm(Defective ~ Temperature + + Density + Rate, data = dataset)
summary(fit)
```

From the summary output why might there be co/multicolinearity in these predictors?

Problem.

Sheather: Chapter 3: Exercises 3 Part A, B, C

Use:

```
dataset = read.csv("https://gattonweb.uky.edu/sheather/book/docs/datasets/AdRevenue.csv",
header = T)
```

Problem.

Sheather: Chapter 5, Exercise 3 (using output provided)

Problem.

Sheather: Chapter 6, Exercise 1

Problem.

Sheather: Chapter 3: Exercise 8 (all parts)

Use:

```
dataset = read.table("https://gattonweb.uky.edu/sheather/book/docs/datasets/diamonds.txt",
header = T)
```