STA 5207 Assignment 7

Due Friday November 5

You do not need to turn in any computer output.

The data set stackloss in the faraway package gives the operational data of a plant for the oxidation of ammonia to nitric acid.

The columns are:

Air.Flow: Flow of cooling air

Water.Temp: Cooling water inlet temperature

Acid.Conc.: Concentration of acid

stack.loss: Stack loss

Use stack.loss as the response and Air.Flow, Water.Temp, and Acid.Conc. as predictors.

The SAS code Assignment7Data.sas (under Data Sets in SAS Studio) includes the data step for reading the data.

1. (45 points) Give the fitted equations using each of the following methods:

* Ordinary Least Squares
  + y = -39.9197 + 0.7156(Air.Flow) + 1.2953(Water.Temp) – 0.1521(Acid.Conc.)
* Huber’s Method
  + y = -41.0265 + 0.8294(Air.Flow) + 0.9261(Water.Temp) – 0.1278(Acid.Conc.)
* Least Absolute Deviations
  + y = -39.68986 + 0.83188(Air.Flow) + 0.57391(Water.Temp) -0.06087(Acid.Conc.)

1. (45 points) The last observation is highly influential. Remove this point and repeat question 1.

* Ordinary Least Squares
  + y = -43.7040 + 0.8891(Air.Flow) + 0.8166(Water.Temp) – 0.1071(Acid.Conc.)
* Huber’s Method
  + y = -42.8415 + 0.9184(Air.Flow) + 0.6854(Water.Temp) – 0.1078(Acid.Conc.)
* Least Absolute Deviations
  + y = -39.98645 + 0.83469(Air.Flow) + 0.56369(Water.Temp) – 0.05691(Acid.Conc.)

1. (10 points) Which method that is more resistant to this influential point than the others? Why?
   * Least absolute deviations is more resistant to this influential data point than the others because its original estimates were not affected much by that observation.