Homework 6

STAT 697STA Spring 2023

Due March 29, 2023, 9:40am on Gradescope

1 Reading

- Read sections 3.1-3.3 HRW.

2 Questions

1. HRW 3.1

   hints: the file BostMortGAMfit.R is very helpful. Note that to plot the spline function as a function of lot size on the scale of the response (price), you need to choose values of the other predictors at which to plot it. There is example code in this file. For some reason, in part (a) they did not standardize the variables. You are welcome to standardize them or not do so given this provided code. When you tweak the example code, be sure to change the plot ranges accordingly, and you can remove part of the truncation of the lower bound of the fit.

2. On page 72, HRW lists 3 key assumptions of standard linear regression. There is a key standard assumption that they have not listed here. This assumption would not be met by data collected by sampling every student in 10 randomly chosen classrooms in Massachusetts. It would be satisfied by randomly sampling one student from each of 200 randomly selected classrooms in Massachusetts. What is this assumption? Is this assumption relaxes with standard GLMs?

3. In your own words (and/or equations) please describe:
   (a) What is meant by the *scaled deviance* of a model?
   (b) Why is the difference in scaled deviance between 2 nested models equal to the likelihood ratio test statistic?

4. (EXTRA CREDIT points) On page 72, equation 3.1 describes the general form of a one-parameter exponential family distribution. Write the Poisson distribution in this form. Be sure to be clear what the natural parameter θ (which multiplies y) is. Derive the expected value and variance according to the derivatives described and verify that they are the true expected value and variance of the Poisson distribution.