STATS 500 - Homework 4

Due **Wednesday**, October 14, 2015

Based on Chapter 6, problem 2 (p. 97)

Using the `teengamb` dataset, fit a model to predict gambling expenditure from all other available variables. Perform regression diagnostics on this model to answer the following questions. Display **only** those plots that are relevant to the questions below. Present your diagnostics in a logical order.

- Check the constant variance assumption for the errors. Modify the model if necessary (see below).
- Check the normality assumption.
- Check for large leverage points.
- Check for outliers.
- Check for influential points.

Hints: You should start with a linear regression of `gamble` on `sex, status, income` and `verbal`. A diagnostic plot will reveal heteroscedasticity in residuals. A standard solution for the type of heteroscedasticity that you will see is to take the log or the square root of the response. Since in this case the `gamble` variable has some zero values, take the square root of `gamble` and fit a new model for this response. Then do all your diagnostic analysis for the new model.

Solutions to this homework should be about 4-5 pages long.