

Statistics 110/201 Homework Assignment for Monday, November 16, 2009

Use the **Patient Satisfaction** data from Problem 6.15 for this homework.

1. There are three predictor variables available for this data set.
 - a. Examine all subsets of predictor variables, and determine the best model based on adjusted R^2 . (If you are using Stata, you will need to do this “manually” by running all 7 possible models using the regress command. Or, if you download “tryem” you can find the best model with 1 and 2 variables, and then include the model with all 3 variables, and compare those choices.)
 - b. For the final model you choose, give the regression results.
 - c. For the final model you choose, calculate C_p and state whether it has a reasonable value.
 - d. For the final model you choose, show an appropriate residual plot and comment on whether you think the model is a good fit.
2. (This is a modified version of Problem 9.17 in the book.) For this problem, use p-to-enter of .04 and p-to-remove of .05.
 - a. Determine the best model using forward *stepwise* regression. (Remember that “stepwise” allows variables to both enter and be removed.) Show your steps.
 - b. Determine the best model using forward *selection* (which enters variables but never removes variables once they are in the model).
 - c. Determine the best model using backward *elimination* (which removes variables but never brings them back in once they have been removed).
3. Did you choose the same final model in Problem 1 and the 3 parts of Problem 2? If not, explain which method you think gave you the best model. If you did choose the same model, explain whether that would always happen, and if not, which method would give you the best model in general.