STAT 4/579 STATISTICAL LEARNING Quiz 5 Name: ____________________

1. In Quiz 3 you fit 6 models to the data. Create a table with 6 rows and 5 columns giving R^2, Adj R^2, C_p, AIC, and BIC for each model. [Given your output from Quiz 3, you could potentially do this on a hand calculator.]

2. How do C_p, AIC, and BIC behave differently? (How do they differ in their choices among the 6 models?)

3. There are $n = 35$ observations in the Quiz 3 data set. Take a random sample **with replacement** of size $n = 35$ of the pairs (x_i, y_i). On this sample, again fit the Haar wavelet (step function) model with the same 8 categories as before. From this fitted model obtain a first set of predictions, say, $\hat{y}_{i1}$ for all 35 x_i values in the original data. Compute and give, say, R_1^2 from these predictions. (R_1^2 is the squared sample correlation between $(y_i, \hat{y}_{i1})$.) [Sorry, but there is nothing like this in any of my computing books. You need to randomly sample the rows of $[Y, X]$ to get a new $[Y_1, X_1]$. Fit the model to the new data, but then get predictions for all of the values in the old X. *I could be talked out of requiring this and the follow-up questions for anyone who is not a graduate student in Statistics, Math, or the School of Engineering.*]

4. Repeat part 2 using two different random samples to obtain predictions $\hat{y}_{i2}$ and $\hat{y}_{i3}$ and from these give new R^2s, say, R_2^2 and R_3^2

5. Combine $\hat{y}_{i1}$, $\hat{y}_{i2}$, and $\hat{y}_{i3}$ into the **Bagged** predictions $\hat{y}_{iB} = (\hat{y}_{i1}+\hat{y}_{i2}+\hat{y}_{i3})/3$. Compute R_B^2 based on the bagged predictions.

6. How does R_B^2 compare to R_1^2, R_2^2 and R_3^2 but most importantly to the comparable R^2 in problem 1? (It should not be systematically different.)

7. I will not ask you to do this but a more interesting bagging problem would be as follows. In addition to the 8 indicator variables associated with Haar wavelets (or the wavelets themselves), also create the 16 indicators that come from partitioning each variable in two. Starting with an intercept, do forward selection on the 24 indicator variables. Then take random samples, repeat the forward selection, and use bagging to combine the results. Hopefully, bagging would do better than a single forward selection but I expect it would do worse than a single backward elimination on the 24 variables.