Quiz 9

Name:

Remember to state the logic behind your answers. If you use any variables, please state what the variables represent.

1. If k is small in a k-fold cross validation is the bias in the estimate of out-of-sample (test set) accuracy smaller or bigger?

2. What is k in leave one out cross validation?

3. For the following question select all that are correct. Stepwise regression is a method to:

a. eliminate extraneous variables from a multiple regression equation.
b. select the best regression equation with a single independent variable.
c. select independent variables to include in a multiple regression equation that increases the R^2 the most.

4. For the following question select all that are correct. The adjusted R^2

a. equals the ratio of $RSS = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ to $TSS = \sum_{i=1}^{n}(y_i - \bar{y})^2$
b. takes values in the range of -1 to 1
c. equals the ratio of $RSS = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ to $TSS = \sum_{i=1}^{n}(y_i - \bar{y})^2$, corrected for the number of independent variables (x's) in the regression equation.

5. For the following question select all that apply. Which of the following methods are appropriate for selecting the best model from among a choice of models:

a. Choose the model with the highest R^2.
b. Choose the model with the highest C_p.
c. Choose the model with the lowest BIC.
d. Choose the model with the lowest AIC.
e. Choose the model with lowest cross-validation mean square error.