Regression analysis quiz

CMSC498 Spring 2015

April 2, 2015

Name(s):
UID(s):

1. Suppose you have data for CMSC undergraduate students including $X_1 =$ `grade in 351` (categorical: A,B,C,D, or F) and $X_2 =$ hours spent learning Java. You want to build a linear regression model of response $Y =$ value of stock options at maturation in first job out of school.

a. Since `grade in 351` is categorical, you need to define a set of dummy predictors to include in the linear regression model. How would you do that? How many dummy predictors would you add, and what would be their value be for students with A, B, C, D, or, F in 351.

b. Write out the full model (without interactions) in the form $Y = \beta_0 + \beta_1 X_1 + \cdots$. How do you interpret the estimate derived for β_0 and each of the parameters corresponding to the dummy predictors you added in part a.

2. Consider the following confusion matrix.

	Observed +	Observed -	Total
Predicted +	80	20	100
Predicted -	30	70	100
Total	110	90	

a. How many True Positives are there?

b. How many False Negatives are there?

c. What is the recall?

d. What is the precision?

e. What is the True Positive Rate?

f. What is the False Positive Rate?