Lab 7 – Practice Questions

Q1. Consider the use of a logistic regression model to predict the probability of default using income and balance on the Default data set. Set a random seed before beginning your analysis.

i. Using the summary() and glm() functions, determine the estimated standard errors for the coefficients associated with income and balance in a multiple logistic regression model that uses both predictors.
ii. Write a function that outputs the coefficient estimates for income and balance in the multiple logistic regression model.
iii. Use the boot() function together with the function in (ii) to estimate the standard errors of the logistic regression coefficients for income and balance.
iv. Comment on the estimated standard errors obtained using the glm() function and using your bootstrap function.

Q2. Compute the LOOCV error for a simple logistic regression model on the Weekly data set. Recall that in the context of classification problems, the LOOCV error is given in equation (5.4). (Refer to ISLR for 5.4)

i. Fit a logistic regression model that predicts Direction using Lag1 and Lag2.
ii. Fit a logistic regression model that predicts Direction using Lag1 and Lag2 using all but the first observation.
iii. Use the model from (ii) to predict the direction of the first observation. Was this observation correctly classified?
iv. Write a for loop from i=1 to i=n, where n is the number of observations in the data set, that performs each of the following steps:
    i. Fit a logistic regression model using all but the ith observation to predict Direction using Lag1 and Lag2.
    ii. Compute the probability of the market moving up for the ith observation.
    iii. Use the probability for the ith observation in order to predict whether or not the market moves up.
    iv. Determine whether or not an error was made in predicting the direction for the ith observation. If an error was made, then indicate this as a 1, and otherwise indicate it as a 0.
v. Take the average of the n numbers obtained in (iv) in order to obtain the LOOCV estimate for the test error. Comment on the results.