Statistical and Machine Learning
ECON576
Problem Set 4

1. (10 points) This question involves the `Default` dataset, which is included in the package ISLR2.

   (a) Fit a logistic regression that uses `income` and `balance` to predict `default` (Model 1). Report the estimates. [Hint: it should include an intercept.]

   (b) Suppose we classify an individual to the default category with a threshold probability of 0.5. Compute the LOOCV test error estimate for Model 1.

   (c) Now consider a logistic regression that predicts `default` using `income`, `balance`, and a dummy variable for `student` (Model 2). Compute the LOOCV test error estimate for this model. Should we include a dummy variable for `student`?

2. (10 points) This question uses the bootstrap to estimate the standard errors under Model 2 in Exercise 1.

   (a) Write a function, `boot.estimates`, that takes as input the `Default` dataset and an index of the observations, and outputs the coefficient estimates of `income`, `balance` and `student` in a logistic regression (with an intercept).

   (b) Generate 1000 bootstrapped samples and use `boot.estimates` to estimate the standard errors of the coefficients corresponding to `income`, `balance` and `student`.