Assignment on Propensity Scores

March 2, 2022

1 Data Analysis: Propensity Scores

Use the dataset called "Teaching Hospital Outcome.txt". The outcome of interest is unfavorable discharge (discharge of a patient that is not to their home but to a nursing facility). The exposure of interest is whether the patient is in a teaching hospital or not. The covariates (potential confounders) are age, sex, race (white=referent, black, other), hispanic ethnicity, diabetes and hypertension.

1. Estimate the odds ratio between unfavorable discharge and teaching hospital controlling for the other variables provided (e.g. potential confounding variables) using a logistic regression model.

2. Develop a prediction model for whether or not patients recieved care at a teaching hospital using the method of random Forests. Use this model to calculate the propensity of receiving care at a teaching hospital.

3. Create 10 bins of the propensity score (e.g deciles of the propensity scores). Estimate the odds ratio between unfavorable discharge and teaching hospital controlling for this categorical bin variable. Note: The random forest approach may not yield propensities with a alot of distinct values, which may lead to an error when the bins are created. To create the bins use the code Bin = cut(PS.rf, unique(quantile(PS.rf, (0:10)/10, na.rm=TRUE))) .

4. Calculate inverse propensity weights (after trimming the propensities so none is smaller than 0.01). Estimate the odds ratio between unfavorable discharge and teaching hospital using these weights.