26 BE 7023 & 26 PH 7023: Advanced Biostatistics

Fall 2018

MB Rao

Homework Sheet No. 6 Due Date: October 31, 2018 Maximum Points: 30

Theme: Logistic Regression

Low birth weight is a harbinger of future ills. I want you to investigate risk factors of low birth weight. A good size data with the title ‘birthwt’ is available in the package ‘MASS.’ Download the data.

1. Describe the data. 4 points
2. Create a new folder by omitting the last column. Convert ‘race,’ ‘smoke,’ and ‘ht’ as factors. Obtain summary statistics. 4 points
3. Fit a logistic regression model with ‘low’ as the response variable using the new folder and covariates the rest. Identify the significant predictors. Interpret carefully the coefficients associated with ‘race.’ Check goodness-of-fit carefully defining what the null hypothesis is. 8 points
4. Fit a logistic regression model with ‘low’ as the response variable and predictors lwt, race, smoke, and ht. Write the prediction equation. Identify the most significant predictor. Find a way to plot P(low =1) as a function of lwt in the presence of various choices of race, smoke, and ht. Check goodness-of-fit. Obtain ‘confusion matrix.’ Calculate the misclassification rate. 14 points