

ECON 8320 – Tools for Data Analysis

Assignment 5 [25 points]

Ordinary Least Squares regressions are problematic at best when exploring binary dependent variables. The logistic regression is typically used in its place. In order to solve for the optimal coefficients in a logistic regression, practitioners will implement a maximum likelihood estimation (MLE), in which the model is stated as an optimization problem. The MLE equation for this problem is written as follows:

$$\ln(L) = \sum_i \left[y_i \ln \left(\frac{\exp(x_i \beta)}{1 + \exp(x_i \beta)} \right) + (1 - y_i) \ln \left(\frac{1}{1 + \exp(x_i \beta)} \right) \right]$$

Your assignment is to modify your class from last week's homework to include an option to specify that logistic regression be performed instead of OLS.

1. Take the inputs necessary to estimate a Logistic regression.
2. Include a method to fit a logistic regression model.
3. Include a method to print the coefficients of the logistic regression (you can ignore the variance estimates as well as t-stats and p-values for this assignment)
4. Store regression results as attributes that can be easily accessed
5. Include a method to predict the probability of success when given new observations

Note: the likelihood function needs to be MAXIMIZED