Discrete Response Model
Lecture 2

datascience@berkeley

Introduction to Binary Response Models and Linear Probability Model

Shortcoming of Linear Probability Model

In LPM, β_j measures the chnage in probability of success when x_j changs by 1 unit, holding other explanatory variables constant:

$$\triangle P(y = 1|\mathbf{x}) = \beta_j \triangle x_j$$

I am going to refer you to the textbook for examples, because we will not study further LPM in this course.

The mechanics of applying the OLS to binary response variable is the same as those covered in $w203$. However, there are a few shortcomings of LPM which makes it very unattractive in modeling binary response variable.

1. The predicted probabilities can go out of the range of 0 and 1.
2. A probability is linearly related to the explanatory variables for all their possible values. This is not realistic. (See the example in the book for more details.)
3. LPM violates the homoskedastic assumption underneath the Classical Linear Regression Model. That is,

$$Var(y|\mathbf{x}) = p(\mathbf{x})[1 - p(\mathbf{x})]$$

Berkeley SCHOOL OF INFORMATION