Discrete Response Model
Lecture 2

datascience@berkeley

Binomial Logistic Regression Model

The Logit Transformation and the Logistic Curve

The Most Common Transformation

- There are a number of solutions to prevent i from being outside the range of a probability.
- Most solutions rely on non-linear transformations to prevent these types of problems from occurring.
- The most commonly used transformation results in the logistic regression model

$$\begin{aligned}\pi_i &= P(y=1|\mathbf{x})\\ &= \frac{exp(\beta_0+\mathbf{x}\beta)}{1+exp(\beta_0+\mathbf{x}\beta)}\\ &= \frac{exp(\beta_0+\beta_1 x_1+\cdots+\beta_k x_p)}{1+exp(\beta_0+\beta_1 x_1+\cdots+\beta_p x_p)}\end{aligned}$$

- Note that $exp(\beta_0+\mathbf{x}\beta)>0$ so that the numerator is always less than the denominator, forcing the response probability to go fall within the zero and one range: $0<\pi_i<1$

The Logistic regression can be written as log(Odd Ratio):

$$log\left(\frac{\pi_i}{1-\pi_i}\right)=\beta_0+\boldsymbol{x}_i\boldsymbol{\beta}$$

Berkeley SCHOOL OF INFORMATION