Discrete Response Model
Lecture 4

datascience@berkeley

Nominal Response Model

Introduction

Suppose there are J **categories** for the response variable with corresponding probabilities π_1, π_2, …, π_J. Using the first category as a **"baseline,"** we can form **"baseline category logits"** as $\boldsymbol{\log(\pi_j/\pi_1)}$ for j = 2, …, J.

When J = 2, we have $\log(\pi_2/\pi_1) = \log(\pi_2/(1-\pi_2))$, which is equivalent to $\log(\pi/(1-\pi))$ in logistic regression with $\pi = \pi_2$.

When there is only one explanatory variable x, we can form the <u>multinomial regression model</u> of

$$\boxed{\log(\pi_j/\pi_1) = \beta_{j0} + \beta_{j1}x \text{ for } j = 2, \ldots, J}$$

One can easily compare other categories so that category 1 is not always used. For example, suppose you would like to compare category 2 to 3. Then

$$\log(\pi_2/\pi_1) - \log(\pi_3/\pi_1) = \log(\pi_2) - \log(\pi_3) = \log(\pi_2/\pi_3)$$

and

$$\boxed{\beta_{20} + \beta_{21}x - \beta_{30} - \beta_{31}x = \beta_{20} - \beta_{30} + x(\beta_{21} - \beta_{31})}$$

Introduction

What is π_j only? Consider the case of one explanatory variable x again:

We can re-write the model as $\pi_j = \pi_1 e^{\beta_{j0}+\beta_{j1}x}$

Noting that $\sum_{j=1}^{J}\pi_j = 1$, we have

$$\pi_1 + \underbrace{\pi_1 e^{\beta_{j0}+\beta_{21}x}}_{\pi_2} + \dots + \underbrace{\pi_1 e^{\beta_{j0}+\beta_{J1}x}}_{\pi_J} = 1$$

Thus,

$$\pi_1 = \frac{1}{1+\sum_{j=2}^{J} e^{\beta_{j0}+\beta_{j1}x}}$$

Also, we can now find that

$$\pi_j = \frac{e^{\beta_{j0}+\beta_{j1}x}}{1+\sum_{j=2}^{J} e^{\beta_{j0}+\beta_{j1}x}}.$$

for j = 2, …, J.

Introduction

- Parameters are estimated using maximum likelihood estimation.
- For a sample of size m, the likelihood function is simply the product of m multinomial distributions with probability parameters as given above.
- Iterative numerical procedures are used then to find the parameter estimates.
- The **multinom() function** from the **nnet package** (within the default installation of R) performs the necessary computations.
- The covariance matrix for the parameter estimates follows from using standard likelihood procedures as outlined in Appendix B (of the text).
- Wald and LR-based inference methods are performed in the same ways as for likelihood procedures in earlier weeks.

Berkeley SCHOOL OF INFORMATION