

MOOC Econometrics

Lecture 5.3 on Binary Choice: Estimation

Richard Paap

Construction of likelihood function

Likelihood contribution for

- observation $y_i = 1$:

$$\Pr[y_i = 1] = \frac{\exp(x_i' \beta)}{1 + \exp(x_i' \beta)}$$

- observation $y_i = 0$:

$$\Pr[y_i = 0] = \frac{1}{1 + \exp(x_i' \beta)}$$

Likelihood contribution for observation i :

$$\left(\frac{\exp(x_i' \beta)}{1 + \exp(x_i' \beta)} \right)^{y_i} \left(\frac{1}{1 + \exp(x_i' \beta)} \right)^{1-y_i}$$

Introduction

Logit model specification in vector notation:

$$\Pr[y_i = 1] = \frac{\exp(x_i' \beta)}{1 + \exp(x_i' \beta)},$$

where $x_i = (1, x_{2i}, \dots, x_{ki})'$ and $\beta = (\beta_1, \dots, \beta_k)'$

It is not possible to write this model in regression notation

$$y_i = x_i' \beta + \varepsilon_i$$

We use maximum likelihood for parameter estimation.

(Log)-likelihood function

Likelihood function of n independent observations:

$$L(\beta) = \prod_{i=1}^n \left(\frac{\exp(x_i' \beta)}{1 + \exp(x_i' \beta)} \right)^{y_i} \left(\frac{1}{1 + \exp(x_i' \beta)} \right)^{1-y_i}$$

Log-likelihood function:

$$\begin{aligned} \log(L(\beta)) &= \sum_{i=1}^n y_i \log \left(\frac{\exp(x_i' \beta)}{1 + \exp(x_i' \beta)} \right) + (1 - y_i) \log \left(\frac{1}{1 + \exp(x_i' \beta)} \right) \\ &= \sum_{i=1}^n y_i x_i' \beta - \log(1 + \exp(x_i' \beta)), \end{aligned}$$

where we use that $\log(ab) = \log(a) + \log(b)$, $\log(a^b) = b \log(a)$ and $\log(a/b) = \log(a) - \log(b)$.

Test question

Test

The maximum likelihood estimator [MLE] is the value of β that maximizes the log-likelihood function. Is the MLE also the value that maximizes the likelihood function?

As the log function is a monotonically increasing function in β , the answer is yes.



Lecture 5.3, Slide 5 of 1, Erasmus School of Economics

Test question

Test

Suppose that all observations on y_i are 0, that is, $y_i = 0$ for $i = 1, \dots, n$. What is the value of the maximum likelihood estimator in this case?

When all observations are 0, the first-order conditions imply that

$$\frac{1}{n} \sum_{i=1}^n \frac{\exp(x_i' b)}{1 + \exp(x_i' b)} = \frac{1}{n} \sum_{i=1}^n y_i = 0.$$

As the logit function is always larger than 0, there is no value of b for which the first-order conditions holds. The MLE does not exist.



Lecture 5.3, Slide 7 of 1, Erasmus School of Economics

Maximum likelihood estimation

The MLE b is obtained by maximizing $\log(L(\beta))$ with respect to β . First-order conditions:

$$\begin{aligned} \frac{\partial \log(L(\beta))}{\partial \beta} &= \frac{\partial \sum_{i=1}^n y_i x_i' \beta - \log(1 + \exp(x_i' \beta))}{\partial \beta} = 0 \\ &= \sum_{i=1}^n y_i x_i' - \frac{\exp(x_i' \beta) x_i'}{1 + \exp(x_i' \beta)} = 0 \end{aligned}$$

Use numerical methods to solve for β .

The first-order conditions imply that

$$\frac{1}{n} \sum_{i=1}^n \frac{\exp(x_i' b)}{1 + \exp(x_i' b)} = \frac{1}{n} \sum_{i=1}^n y_i$$



Lecture 5.3, Slide 6 of 1, Erasmus School of Economics

Properties of maximum likelihood estimator

It can be shown that under regularity conditions the maximum likelihood estimator [MLE] is

- ① Consistent
- ② Efficient for large n
- ③ Asymptotically normally distributed, and hence

$$b \approx N(\beta, V)$$

The (co)variance matrix V can be estimated by

$$\hat{V} = \left(\sum_{i=1}^n \left(\frac{\exp(x_i' b)}{1 + \exp(x_i' b)} \right) \left(\frac{1}{1 + \exp(x_i' b)} \right) x_i x_i' \right)^{-1}$$



Lecture 5.3, Slide 8 of 1, Erasmus School of Economics

Testing for single parameter restriction

We want to compare

- logit model without parameter restrictions
- logit model with a single $\beta_j = 0$

Hypothesis:

$$H_0: \beta_j = 0 \text{ versus } H_1: \beta_j \neq 0$$

You can use the t -test like in a linear regression: Test statistic:

$$z_j = \frac{b_j - 0}{SE(b_j)} \approx N(0, 1),$$

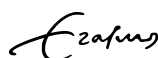
where $SE(b_j)$ is the standard error of b_j .



Lecture 5.3, Slide 9 of 1, Erasmus School of Economics

Training Exercise 5.3

- Train yourself by making the training exercise (see the website).
- After making this exercise, check your answers by studying the webcast solution (also available on the website).



Lecture 5.3, Slide 11 of 1, Erasmus School of Economics

Testing for a set of parameter restrictions

We want to compare

- logit model without parameter restrictions and estimates b_1
- logit model with m parameter restrictions and estimates b_0

The null hypothesis is that the m parameter restrictions are correct.

To compute the test statistic we need

- $L(b_1)$: maximum likelihood value in full model
- $L(b_0)$: maximum likelihood value in restricted model

Test statistic:

$$LR = -2(\log(L(b_0)) - \log(L(b_1))) \approx \chi^2(m),$$

where m is the number of restrictions.



Lecture 5.3, Slide 10 of 1, Erasmus School of Economics