

MOOC Econometrics

Lecture 5.4 on Binary Choice: Evaluation

Richard Paap

Erasmus University Rotterdam



Test question

Test

Suppose that we have perfect fit for all n observations, that is,

$$y_i - \frac{\exp(x_i' \beta)}{1 + \exp(x_i' \beta)} \approx 0$$

for all i . What is the numerical value of the likelihood function

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

For all observations equal to 1 (or 0) the likelihood contribution is very close to 1. Hence, the likelihood function equals about 1.



Residuals

Logit residuals:

$$\begin{aligned} y_i - E[y_i] &= y_i - (0 \times \Pr[y_i = 0] + 1 \times \Pr[y_i = 1]) \\ &= y_i - \Pr[y_i = 1] \\ &= y_i - \frac{\exp(x_i' b)}{1 + \exp(x_i' b)} \end{aligned}$$

Interesting cases:

- Lower bound: $y_i - E[y_i] \approx -1$
- Upper bound: $y_i - E[y_i] \approx 1$
- Perfect fit $y_i - E[y_i] \approx 0$



Measures of fit

Define

- $L(b)$: the maximum value of the likelihood function of the model under consideration
- $L(b_1)$: maximum value of the likelihood function in case the model only contains an intercept.

Perfect fit corresponds to $L(b) \approx 1$ or $\log(L(b)) \approx 0$.

Two popular pseudo R^2 measures are:

- McFadden R^2 :

$$R^2 = 1 - \frac{\log(L(b))}{\log(L(b_1))}$$

- Nagelkerke R^2 :

$$R^2 = \frac{1 - \left(\frac{L(b_1)}{L(b)} \right)^{2/n}}{1 - L(b_1)^{2/n}}$$



Prediction probability

If the value of x_{n+1} is available, one can predict the value of y_{n+1} using

$$\begin{aligned} E[y_{n+1}] &= 0 \times \Pr[y_{n+1} = 0] + 1 \times \Pr[y_{n+1} = 1] \\ &= \Pr[y_{n+1} = 1] \\ &= \frac{\exp(x'_{n+1}\beta)}{1 + \exp(x'_{n+1}\beta)} \end{aligned}$$

To estimate this probability we replace β by its estimate b and obtain $\hat{\Pr}[y_{n+1} = 1]$.



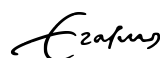
Lecture 5.4, Slide 5 of 12, Erasmus School of Economics

Test question

Test

Does a higher value of the cut-off value c generate more, the same or less predictions which are equal to 1?

A higher value of c means that less (or the same number of) prediction probabilities are above c and hence you forecast less (or the same number of) ones.



Lecture 5.4, Slide 7 of 12, Erasmus School of Economics

0/1 Prediction

The prediction is a probability and never exactly equal to 0 or 1.

Transform the prediction probability into 0/1 forecast \hat{y}_{n+1} by the rule:

$$\begin{aligned} \hat{y}_{n+1} &= 1 \text{ if } \hat{\Pr}[y_{n+1} = 1] > c \\ \hat{y}_{n+1} &= 0 \text{ if } \hat{\Pr}[y_{n+1} = 1] \leq c. \end{aligned}$$

Many statistical packages use $c = 0.5$. However, one may also consider

$$c = \frac{1}{n} \sum_{i=1}^n y_i,$$

that is, the fraction of observations in the sample equal to one.



Lecture 5.4, Slide 6 of 12, Erasmus School of Economics

Evaluation of predication accuracy

Suppose one has m out-of-sample predictions for y_i denoted by \hat{y}_i .

Count the number of correct and incorrect predictions:

$$m_{11} = \sum_{i=1}^m y_{n+i} \hat{y}_{n+i} \quad \text{data}=1 \text{ \& prediction}=1$$

$$m_{00} = \sum_{i=1}^m (1 - y_{n+i})(1 - \hat{y}_{n+i}) \quad \text{data}=0 \text{ \& prediction}=0$$

$$m_{10} = \sum_{i=1}^m y_{n+i}(1 - \hat{y}_{n+i}) \quad \text{data}=1 \text{ \& prediction}=0$$

$$m_{01} = \sum_{i=1}^m (1 - y_{n+i})\hat{y}_{n+i} \quad \text{data}=0 \text{ \& prediction}=1$$



Lecture 5.4, Slide 8 of 12, Erasmus School of Economics

Prediction-realization table

Classify predictions in right and wrong:

observed	predicted		sum
	$\hat{y} = 0$	$\hat{y} = 1$	
$y = 0$	m_{00}/m	m_{01}/m	$(m_{00} + m_{01})/m$
$y = 1$	m_{10}/m	m_{11}/m	$(m_{10} + m_{11})/m$
sum	$(m_{00} + m_{10})/m$	$(m_{01} + m_{11})/m$	1



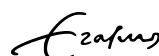
Lecture 5.4, Slide 9 of 12, Erasmus School of Economics

Prediction-realization table

Classify predictions in right and wrong:

observed	predicted		sum
	$\hat{y} = 0$	$\hat{y} = 1$	
$y = 0$	m_{00}/m	m_{01}/m	$(m_{00} + m_{01})/m$
$y = 1$	m_{10}/m	m_{11}/m	$(m_{10} + m_{11})/m$
sum	$(m_{00} + m_{10})/m$	$(m_{01} + m_{11})/m$	1

$m_{01}/m + m_{10}/m$ denotes the fraction of incorrect forecasts.



Lecture 5.4, Slide 11 of 12, Erasmus School of Economics

Prediction-realization table

Classify predictions in right and wrong:

observed	predicted		sum
	$\hat{y} = 0$	$\hat{y} = 1$	
$y = 0$	m_{00}/m	m_{01}/m	$(m_{00} + m_{01})/m$
$y = 1$	m_{10}/m	m_{11}/m	$(m_{10} + m_{11})/m$
sum	$(m_{00} + m_{10})/m$	$(m_{01} + m_{11})/m$	1

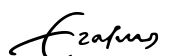
The fraction $m_{00}/m + m_{11}/m$ is called the hit rate.



Lecture 5.4, Slide 10 of 12, Erasmus School of Economics

Training Exercise 5.4

- 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 5.4, Slide 12 of 12, Erasmus School of Economics