Erasmus School of Economics

MOOC Econometrics

Lecture 5.4 on Binary Choice: Evaluation

Richard Paap

Erasmus University Rotterdam

Residuals

Logit residuals:

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

Interesting cases:

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

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

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

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

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

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

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

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:

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

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.

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.

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:

$$\begin{aligned} m_{11} &= \textstyle\sum_{i=1}^{m} y_{n+i}\hat{y}_{n+i} && \text{data=1 \& prediction=1} \\ m_{00} &= \textstyle\sum_{i=1}^{m} (1 - y_{n+i})(1 - \hat{y}_{n+i}) && \text{data=0 \& prediction=0} \\ m_{10} &= \textstyle\sum_{i=1}^{m} y_{n+i}(1 - \hat{y}_{n+i}) && \text{data=1 \& prediction=0} \\ m_{01} &= \textstyle\sum_{i=1}^{m} (1 - y_{n+i})\hat{y}_{n+i} && \text{data=0 \& prediction=1} \end{aligned}$$

Prediction-realization table

Classify predictions in right and wrong:

observed	predicted $\hat{y}=0$	predicted $\hat{y}=1$	sum
$y=0$	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

Prediction-realization table

Classify predictions in right and wrong:

observed	predicted $\hat{y}=0$	predicted $\hat{y}=1$	sum
$y=0$	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.

Prediction-realization table

Classify predictions in right and wrong:

observed	predicted $\hat{y}=0$	predicted $\hat{y}=1$	sum
$y=0$	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.

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).