

Binary Choice Model

After finish this session, you should understand:

- Concept of Logit/Probit model.
- Index Function, Latent variable, and Probability Model.
- Differences between Logit and Probit model.
- Interpretation of estimated results of Logit/Probit model
- Marginal Effects
- Pseudo R-square vs Counted R-square

Binary Choice Model

Concept of Choice Model

Discrete Choice Model

Decision making process:

Individual makes a marginal benefit-marginal cost calculation based on the utilities achieved by each choice.

Model shows factors that influence decision making through the index function.

Index function determines net cost-benefit of each choice individual make.

Concept of Choice Model

Index function

$$U = X\beta + \varepsilon$$

Where U is unobserved variable of the difference between benefit and cost
 ε is normally or logistic distributed with mean 0 and variance 1

The observation is

$$\begin{array}{ll} y = 1 & \text{if } U > 0 \\ y = 0 & \text{if } U \leq 0 \end{array}$$

Logit Model

Random Utility Models

$$U^a = X \beta_a + \varepsilon_a \quad \text{and} \quad U^b = X \beta_b + \varepsilon_b$$

$Y=1$ if consumer's choice of alternative a :

$$\begin{aligned} P_i &= \Pr(y = 1 | X) = \Pr[U^a > U^b] \\ &= \Pr[X \beta_a + \varepsilon_a - X \beta_b - \varepsilon_b > 0 | X] \\ &= \Pr[X (\beta_a - \beta_b) + \varepsilon_a - \varepsilon_b > 0 | X] \\ &= \Pr[X \beta + \varepsilon > 0 | X] \end{aligned}$$

Logit Model

Index function $I_i = X\beta$

Assume that the probability function that the choice will be chosen is logistic distribution:

$$P_i = \Pr(y = 1 | X) = \frac{1}{1 + e^{-X\beta}}$$

$$P_i = \frac{1}{1 + e^{-X\beta}} = \frac{e^{X\beta}}{1 + e^{X\beta}} = \Lambda(X\beta)$$

where: $\Lambda(.)$ is logistic distribution function.

$$L = \prod_{i=1}^n \left(\frac{1}{1 + e^{-X_i\beta}} \right)^{Y_i=1} \cdot \left(1 - \frac{1}{1 + e^{-X_i\beta}} \right)^{Y_i=0}$$

Probit Model

Unobservable utility index (I_i) or latent var.

$$I_i = X\beta$$

Assume I_i^* is threshold or critical level.

$$P_i = \Pr(y = 1 | X) = P(I_i^* \leq I_i) = F(X\beta) = \Phi(X\beta)$$

Assume $\Phi(.)$ is cumulative normal distribution function.

Marginal Effect

Logit Model

$$\Pr(Y = 1|X) = \Lambda(X\beta)$$

$$\begin{aligned}\frac{\partial \Pr(Y = 1|X)}{\partial x} &= \Lambda(X\beta)[1 - \Lambda(X\beta)]\beta \\ &= \hat{p}[1 - \hat{p}]\beta\end{aligned}$$

Probit Model

$$\Pr(Y = 1|X) = \Phi(X\beta)$$

$$\frac{\partial \Pr(Y = 1|X)}{\partial x} = \phi(X\beta)\beta$$

Overall Significance Test

$$H_0: \beta_2 = \beta_3 = \dots = \beta_k = 0$$

Test Statistic $2(\ln L_{UR} - \ln L_R) \sim \chi^2_{(k-1)}$

where

$\ln L_{UR}$ = Log-likelihood value of estimated model

$\ln L_R$ = Log-likelihood value of restricted model

Individual Test — Z-test

Z-test: $H_0: \beta_i = 0$

Test Statistic $Z - test = \frac{\hat{\beta}_i}{s_{\hat{\beta}_i}} \sim N(0,1)$

Measure Goodness of Fit

Pseudo (McFadden) R^2

$$McFadden\ R^2 = 1 - \frac{\ln L_{UR}}{\ln L_R}$$

where

$\ln L_{UR}$ = Log-likelihood value of estimated model

$\ln L_R$ = Log-likelihood value of restricted model
(model with only intercept term).

McFadden's Adjusted R^2

$$McFadden's\ Adjusted\ R^2 = 1 - \frac{\ln L_{UR} - k}{\ln L_R}$$

where k = Number of independent variables.

Measure Goodness of Fit

Cox-Snell R^2

$$Cox - Snell R^2 = 1 - \exp\left(-2\left[\ln L_{UR} - \ln L_R\right]/n\right)$$

Cox-Snell R^2 cannot attain a value of 1 which is the disadvantage of this measure. When the model perfectly fits, this $R^2=0.75$.

Cragg – Uhler (Nagelkerke) R^2

$$Cragg - Uhler R^2 = \frac{(Cox - Snell R^2)}{(1 - \exp(2[\ln L_R]/n))}$$

Measure Goodness of Fit

McKelvey – Zavoina R^2

$$\text{McKelvey – Zavoina } R^2 = \frac{\hat{Var}(\hat{I})}{\hat{Var}(\hat{I}) + Var(\hat{\varepsilon})}$$

where

$\hat{Var}(\hat{I})$ = variance of predicted index value

$\hat{Var}(\hat{\varepsilon})$ = variance of predicted residuals.

Measure Goodness of Fit

Efron R^2

$$\text{Efron } R^2 = 1 - \left(\frac{n}{n_1 \cdot n_2} \right) \sum_{i=1}^n (y_i - \hat{P}_i)^2$$

where

y_i = Actual value

$\hat{P}_i$ = Predicted probability $Prob[y=1]$

n_1 = Number of observation $y_i = 0$

n_2 = Number of observation $y_i = 1$

n = Total number of observation.

Forecasting Error Index

Counted R^2

$$\text{Counted } R^2 = \frac{\text{No. of Correct Prediction}}{\text{Total No. of Observation}}$$

where

If $\text{Prob}[y=1]$ or $\hat{P} \leq 0.5$, then, $\hat{y} = 0$

If $\text{Prob}[y=1]$ or $\hat{P} > 0.5$, then, $\hat{y} = 1$

Example I $\text{Counted } R^2 = \frac{90 + 3}{100} = 0.93 \text{ or } 93\%$

	Predicted $Y=0$	Predicted $Y=1$	Total
Actual $Y=0$	90	5	95
Actual $Y=1$	2	3	5
Total	92	8	100

Forecasting Error Index

Counted R^2

Example 2

	<i>Predicted $Y=0$</i>	<i>Predicted $Y=1$</i>	<i>Total</i>
<i>Actual $Y=0$</i>	95	0	95
<i>Actual $Y=1$</i>	5	0	5
<i>Total</i>	100	0	100

$$\text{Overall Counted } R^2 = \frac{95 + 0}{100} = 0.95 \text{ or } 95\%$$

$$Y = 0 \text{ Counted } R^2 = \frac{95}{95} = 1.00 \text{ or } 100\%$$

$$Y = 1 \text{ Counted } R^2 = \frac{0}{5} = 0.00 \text{ or } 0\%$$

Forecasting Error Index

Adjusted Counted R^2

$$\text{Counted } R^2 = \frac{\text{No. of Correct Prediction} - n^*}{\text{Total No. of Observation} - n^*}$$

where n^* is number of most frequent outcome.

Example 2

	Predicted $Y=0$	Predicted $Y=1$	Total
Actual $Y=0$	95	0	95
Actual $Y=1$	5	0	5
Total	100	0	100

$$\text{Overall Counted } R^2 = \frac{95 + 0}{100} = 0.95 \text{ or } 95\%$$

$$\text{Adjusted Counted } R^2 = \frac{(95 + 0) - 95}{100 - 95} = 0 \text{ or } 0\%$$

Evaluation Criteria

1. Sign and meaning of the Coefficients.

- Whether the estimated results are according to the theory.
- Meaning – Marginal Effects at ...

2. Overall Test – LR-Chi-squares-test.

- Whether all explanatory variables can be used in explaining the dependent variable.

3. GOF and Forecasting Error Index.

- Pseudo R^2 - How well does the estimated results contribute to the likelihood of the model? – Making comparison
- Counted R^2

Evaluation Criteria

4. Individual Test – z-test.

- Whether each explanatory variables can explain the dependent variable.
- z-test – MLE assume Asymptotic Normal.