Binary Choice Model

After finish this session, you should understand:

Example STATA

$Y=0$
$=1$

$P(Y=1)$

- Concept of Logit/Probit model.
- Index Function, Latent variable, and Probability Model.

Logistic Cum. Normal

MLE

- Differences between Logit and Probit model.
- Interpretation of estimated results of Logit/Probit model

mfx

1. Sign & Meaning. (Nonlinear)
2. Overall Test. – LR χ^2-Test
3. ~~χ^2~~
4. Ind. Test – z-Test

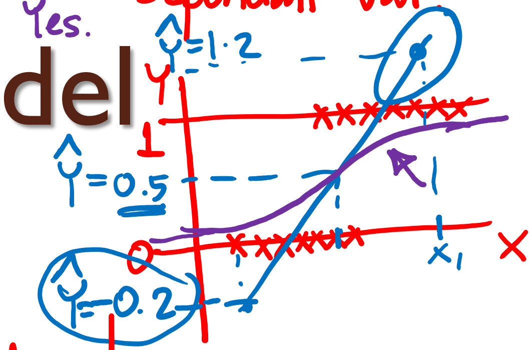
Asymp. Normal

- Marginal Effects
- Pseudo R-square vs Counted R-square

Binary Choice Model

$Y = 0$ No
$= 1$ Yes.

Dummy Variable
Dependent Var.



Nonlinear

Concept of Choice Model

<u>Discrete Choice Model</u>

Decision making process:

Net benefit.

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

Latent Variable

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

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

$$\boxed{\begin{matrix} y = 1 \checkmark \\ y = 0 \checkmark \end{matrix}} \qquad \begin{matrix} \text{if } U > 0 \\ \text{if } U \le 0 \end{matrix}$$

Logit Model

Random Utility Models

1970s

$\beta_{a1} + \beta_{a2}X_2 + \cdots + \beta_{ak}X_k$

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

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

$Y=1$ $Y=0$

$$P_i = \Pr(y=1|X) = \Pr\left[U^a > U^b\right]$$

$\cancel{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\varepsilon^2}{2\sigma^2}}}$

$$= \Pr\left[X\beta_a + \varepsilon_a - X\beta_b - \varepsilon_b > 0|X\right]$$

$$= \Pr\left[X(\beta_a - \beta_b) + \varepsilon_a - \varepsilon_b > 0|X\right]$$

$$\boxed{P(Y=1|X) = \Pr\left[X\beta + \varepsilon > 0|X\right]}$$

Normal Dist^{n}

By Tatre Jantarakolica

Logit Model

$Y = 0$
$ = 1$

Index function ① $\hat{I}_i = X\hat{\beta}$

$P(Y=0) + P(Y=1) = 1$

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

② $$\boxed{\hat{P}_i = \Pr(y = 1 \mid X) = \frac{1}{1+e^{-X\hat{\beta}}}}$$

$P(Y=0 \mid X) = 1 - \left(\frac{1}{1+e^{-X\beta}}\right)$

③ $\hat{Y}_i = 0$ if $\hat{P} \leq 0.5$

$\hat{Y}_i = 1$ if $\hat{P}_i > 0.5$

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

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

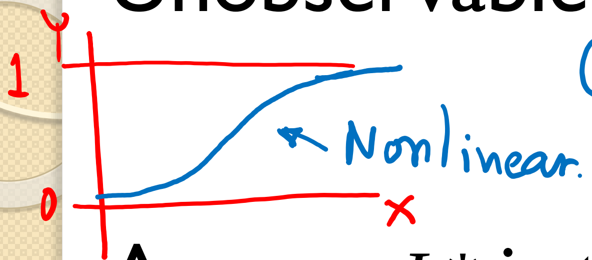
$P(1\ 0\ 2\ 0 \ldots a_n)$

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

By Tatre Jantarakolica

Probit Model

Unobservable utility index (I_i) or latent var.



① $\hat{I}_i = X\hat{\beta}$

$$\int_{-\infty}^{X\beta} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{X\beta}{2\sigma^2}}$$

Assume I_i* is threshold or critical level.

② $$\frac{\partial \hat{P}_i}{\partial X} \qquad \hat{P}_i = \Pr(y = 1 | X) = P(I_i^* \le I_i) = F(X\beta) = \Phi(X\hat{\beta})$$

③ $\hat{Y}=0 \leftarrow \hat{P} \le 0.5$

$\hat{Y}=1 \leftarrow \hat{P} > 0.5$

Assume Φ(.) is cumulative normal distribution function.

Normit

$$L = \prod_{i=1}^{n} \Phi(\cdot)^{Y_i=1} (1-\Phi(\cdot))^{Y_i=0}$$

MLE

1. Sign & Meaning mfx
2. Overall Test
3. R^2 ?
4. Ind. Test z-test

By Tatre Jantarakolica

Marginal Effect

at Mean
at Median

mfx. at (value of X_s)

Mean
$\bar{X}$

Logit Model

Logistic $\frac{1}{1+e^{-X\hat{\beta}}}$

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

$$\frac{\partial \Pr\left(Y=1|X\right)}{\partial x} = \Lambda\left(X\beta\right)\left[1-\Lambda\left(X\beta\right)\right]\beta$$

$$= \hat{p}\left[1-\hat{p}\right]\beta$$

Probit Model

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

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

By Tatre Jantarakolica

Overall Significance Test

MLE

②

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

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

where LR χ^2 Test

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

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

③ $\cancel{R^2}$

④

Individual Test — Z-test

Asymp. Normal $n \to \infty$

<u>Z-test:</u> $H_0: \beta_i = 0$

<u>Test Statistic</u> $Z - test = \dfrac{\hat{\beta}_i}{s_{\hat{\beta}_i}} \sim N(0,1)$

Measure Goodness of Fit

MLE

$\log 0.1 = -1$

$\log 0.01 = -2$

Pseudo (McFadden) R^2

$R^2 = 1 - \frac{RSS}{TSS}$

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

< 1

$\to 0.$

Make Comparison

$R^2 = 0.9$

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.

By Tatre Jantarakolica

Measure Goodness of Fit

- Cox-Snell R^2

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

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.

0 – 0.75

- Cragg – Uhler (Nagelkerke) R^2 –

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

0 – 1

By Tatre Jantarakolica

Measure Goodness of Fit

McKelvey – Zavoina R^2

$$McKelvey - Zavoina\ R^2 = \frac{\hat{V}ar(\hat{I})}{\hat{V}ar(\hat{I}) + \hat{V}ar(\hat{\varepsilon})}$$

where

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

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

Measure Goodness of Fit

$(0-0.5)^2$

$(1-0.9)^2$

Efron R^2

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

where

$\boxed{\text{Making Comparison}}$

y_i = Actual value

$n_1 + n_2 = n$

$\hat{P}_i$ = 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 > 0.5

$\hat{I} = X\hat{\beta}$

$\hat{P} = \frac{1}{1+e^{-\hat{I}}}$

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

where

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

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

Actual

$Y=0$

$Y=1$

Example 1

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

$\hat{Y}=0$ $\hat{Y}=1$

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

By Tatre Jantarakolica

Forecasting Error Index

Counted R^2

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

$\hat{Y}=0$ $\hat{Y}=1$

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

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

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

By Tatre Jantarakolica

Forecasting Error Index

Adjusted Counted R^2

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

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

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

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

5 5

By Tatre Jantarakolica

Evaluation Criteria

$\frac{\partial Y}{\partial X_2} = \beta_2$

★★★

1. Sign and meaning of the Coefficients.
   - Whether the estimated results are according to the theory. (−) (+)

     $\frac{\partial P}{\partial X_2} \neq \beta_2$
   - Meaning – Marginal Effects at …
2. Overall Test – LR-Chi-squares-test. MLE
   - Whether all explanatory variables can be used in explaining the dependent variable. $Y=0,1$
3. GOF and Forecasting Error Index.
   - Pseudo R^2 - How well does the estimated results contribute to the likelihood of the model? – $\boxed{\text{Making comparison}}$
   - Counted R^2

     Overall $Y=0$ $Y=1$

     > 0.5

Evaluation Criteria

4. Individual Test – z-test.

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