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.

By Tatre Jantarakolica

Concept of Choice Model

<u>Index function</u>

$$I = x'\beta + \varepsilon$$

Where I 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

$$y = 1 \qquad if \;\; I > 0$$
$$y = 0 \qquad if \;\; I \le 0$$

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 :

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

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

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

$$= \Pr\left[x'\beta + \varepsilon > 0 | x\right]$$

Logit Model

<u>Index function</u> $I_i = x'\beta$

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

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

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

Probit Model

Unobservable utility index (I_i) or latent var.

$$I_i = x'\beta$$

Assume I_j* is threshold or critical level.

$$P_i = \Pr(y=1 | x) = P(I_i^* \le I_i) = F(x'\beta) = \Phi(x'\beta)$$

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

Marginal Effect

Logit Model

$$\Pr\left(Y=1\middle|x\right)=\Lambda(x'\beta)$$

$$\frac{\partial E\left[y\middle|x\right]}{\partial x}=\Lambda\left(x'\beta\right)\left[1-\Lambda\left(x'\beta\right)\right]\beta$$

Probit Model

$$\Pr\left(Y=1\middle|x\right)=\Phi(x'\beta)$$

$$\frac{\partial E\left[y\middle|x\right]}{\partial x}=\phi\left(x'\beta\right)\beta$$

Overall Significance Test

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

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

where

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

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

Individual Test--Pseudo t-test

Pseudo t-test: $H_0: \beta_i = 0$

Test Statistic $Pseudo\ t-test\ or\ Z-test \approx \dfrac{\hat{\beta}_i}{s_{\hat{\beta}_i}} \sim N(0,1)$

Measure Goodness of Fit

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 of the overall test

Measure Goodness of Fit

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

y_i = Actual value

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

Forecasting Error Index

Counted R^2

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

where

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

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

Example 1

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

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

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

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

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

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