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

$$y = 1 \qquad \text{if } U > 0$$

$$y = 0 \qquad \text{if } U \leq 0$$

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 :

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

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

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

$$= \Pr\left[X\beta + \varepsilon > 0 \mid X\right]$$

Logit Model

Index function $I_i = X\beta$

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

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

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

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

Probit Model

Unobservable utility index (I_i) or latent var.

$$I_i = X\beta$$

Assume I_i* is threshold or critical level.

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

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

Marginal Effect

Logit Model

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

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

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

Probit Model

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

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

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 — Z-test

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

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

Measure Goodness of Fit

Psuedo (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

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

Measure Goodness of Fit

McKelvey – Zavoina R^2

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

where

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

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

Measure Goodness of Fit

Efron R²

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

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 R²

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

where

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

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

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

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

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

Forecasting Error Index

Adjusted Counted R²

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

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

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

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