Choice Models

Covered Topics

- Binary Choice
  - LPM
  - logit
  - logistic regresion
  - probit
- Multiple Choice
  - Multinomial Logit

Binary Choice

- Yes or No
- Buy or Not Buy
- Join or Not Join
- Own or Not Own
- Switch or Stay

Multiple Choice

- Yes, No, Abstain
- Buy, Sell or No Action
- Buy Brand A, B, C or None
- Join Plan X, Y or Z

Mutual Exclusiveness

Note that all the choices must be mutually exclusive and exhaustive. One and only one choice or event will occur.

Choice Model (1)

Question: What determines the choice selection?

Model to determine the probability of an event under a given condition (value of independent variables)

$$\text{Pr(choice\#}j) = F_j(X_1, X_2, \ldots, X_K)$$

where X's are determinants for the probability.

Choice Model (2)

Note that

1) $\sum_j \Pr(\text{choice\# } j) = 1$

2) function $F_j()$ must return a value between 0 and 1

Quantification of Binary Choices

<u>Example</u>

JOIN=1 if the observation will join the government-run health insurance program

= 0, otherwise

Quantification of Multiple Choices

JA=1 if the observation will join Plan A
= 0, otherwise
JB=1 if the observation will join Plan B
= 0, otherwise
JC=1 if the observation will join Plan C
= 0, otherwise
Note that JA+JB+JC=1 always.

Binary Choice Model

<u>General Structure</u>

$$\Pr(JOIN = 1) = F(X_1, X_2, ..., X_K)$$

$$\Pr(JOIN = 0) = 1 - F(X_1, X_2, ..., X_K)$$

Note that

$$0 \le F(X_1, X_2, ..., X_K) \le 1$$

Linear Probability Model (1)

Define $P = \Pr(JOIN = 1)$

<u>Assumption of LPM</u>

Linearity of F(.)

$$P = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K$$

Note that there is no error term

Linear Probability Model (2)

<u>Formulation of LPM</u>

E(JOIN)=(1)P+(0)(1-P)=P

==> JOIN=P+v

where v is an error term. E(v)=0

$$JOIN = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K + v \quad \text{----(1)}$$

=> OLS is valid but not the best. Why?

Linear Probability Model (3)

Note that

$V(\nu) = V(JOIN)$ but

$V(JOIN) = (1\text{-}P)^2P + (0\text{-}P)^2(1\text{-}P)$

$= P(1\text{-}P)$

==> $V(\nu)$ is not constant. It depends on the independent variables (X's)

==> Violation of a CLRM assumption or ν is heteroscedastic

Linear Probability Model (4)

Define $w = \sqrt{\dfrac{1}{P(1-P)}}$

$$JOIN^* = \beta_1 X_1^* + \beta_2 X_2^* + \cdots + \beta_K X_K^* + \nu^* \quad \text{----}(2)$$

where $JOIN^* = wJOIN$

$X_k^* = wX_k$ for $k = 1,\ldots,K$

$\nu^* = w\nu$

Linear Probability Model (5)

Note that

$$V(\nu^*) = w^2 V(\nu)$$

$$= \frac{1}{P(1-P)} P(1-P)$$

$$= 1$$

==> OLS is BLUE for Model (2)

Linear Probability Model (6)

<u>Estimation of LPM</u>

Step 1 run OLS for unweighted model (1)

$$\Longrightarrow \widehat{JOIN} = \mathbf{X}\hat{\boldsymbol{\beta}}$$

Note that $\widehat{JOIN}$ is the estimate for P

Linear Probability Model (7)

Step 2 compute the weight

$$w = \sqrt{\frac{1}{\widehat{JOIN}(1-\widehat{JOIN})}}$$

Step 3 compute $JOIN^*, X_1^*, X_2^*, ..., X_K^*$

Step 4 estimate the weighted model (2) using OLS

Linear Probability Model (8)

Step 5 re-compute $\widehat{JOIN}$ using the new set of $\hat{\beta}$.

Note that LPM does not assure that

$$0 \leq \widehat{JOIN} \leq 1$$

or $$0 \leq F(X_1, X_2, ..., X_K) \leq 1$$

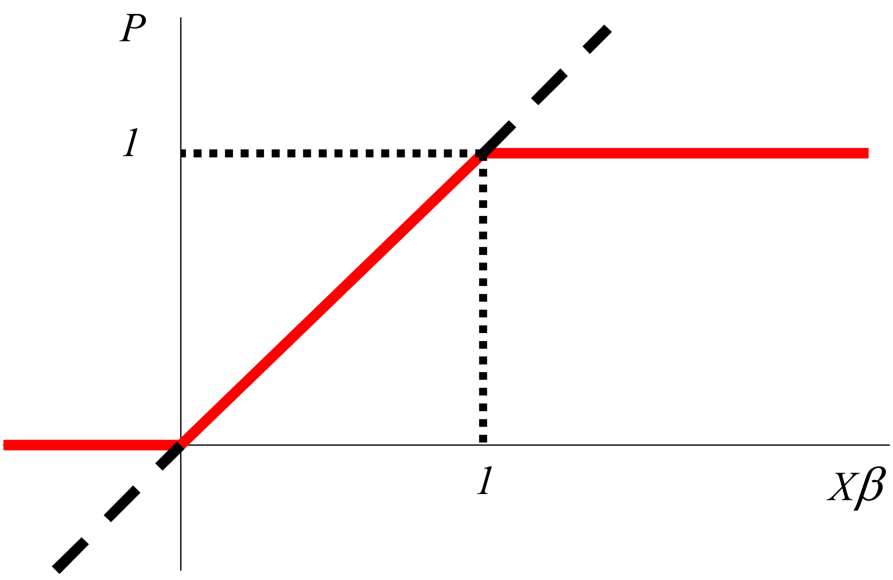
Linear Probability Model (9)

<u>Correction</u>

If $X\hat{\beta}<0$, set $\widehat{JOIN}=0$

If $X\hat{\beta}>1$, set $\widehat{JOIN}=1$

Linear Probability Model (10)



Linear Probability Model (11)

- Less expensive in computer time. No non-linear equations
- $\frac{\partial P}{\partial X_k} = \beta_k$ is the effect of X on the probability. In general, the explanatory variables should be unitless or are expressed in percentage

Logit Model (1)

Assumption of Logit

F() is a logistic function

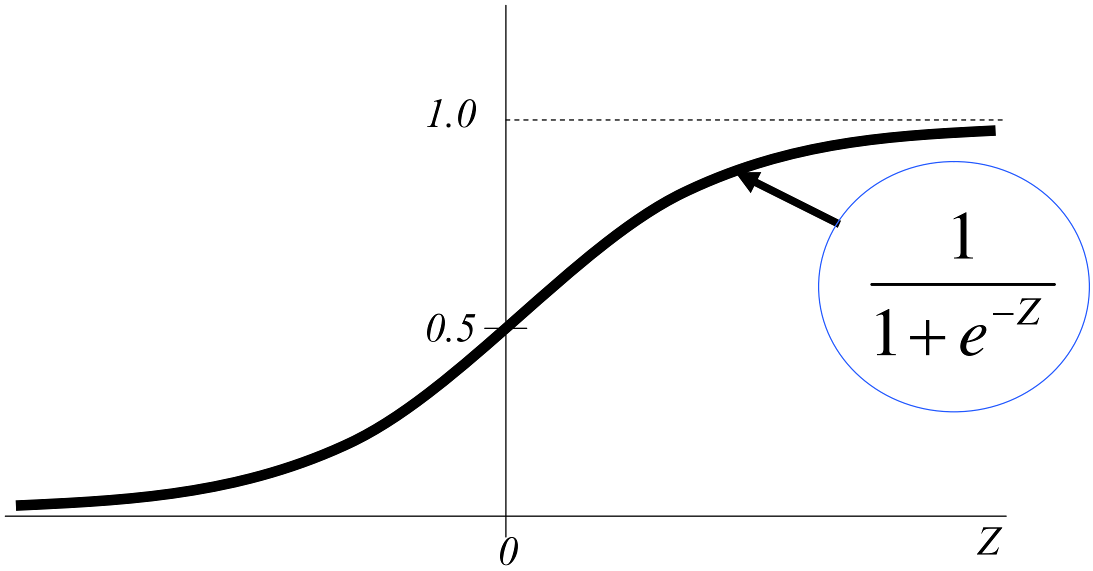
No error term

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

$$Z = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K$$

Note that $0 \leq F(Z) \leq 1$ always.

Logit Model (2)



Logit Model (3)

Note that OLS does not apply

<u>ML Estimation of Logit model</u>

$$\max_{\boldsymbol{\beta}} L = \prod_{i=1}^{n} (P_i)^{Y_i} (1-P_i)^{(1-Y_i)}$$

or $$\max_{\boldsymbol{\beta}} \ln L = \sum_{i=1}^{n} [Y_i \ln(P_i) + (1-Y_i)\ln(1-P_i)]$$

Note that Y=JOIN

Logit Model (4)

Note that $$1 - P = \frac{1}{1 + e^{Z}}$$

First-order conditions

For k=1,…,K

$$\frac{\partial \ln L}{\partial \beta_k} = \sum_{i=1}^{n} [X_{ki} Y_i \frac{e^{-Z_i}}{1 + e^{-Z_i}}]$$

$$- \sum_{i=1}^{n} [X_{ki} (1 - Y_i) \frac{e^{Z_i}}{1 + e^{Z_i}}] = 0$$

Logit Model (5)

Solving FOC for ML estimates.

Second-order Conditions

$$\frac{\partial^2 \ln L}{\partial \beta_j \partial \beta_k} = - \sum_{i=1}^{n} [X_{ji} X_{ki} Y_i \frac{e^{Z_i}}{(1 + e^{Z_i})^2}]$$

$$- \sum_{i=1}^{n} [X_{ji} X_{ki} (1 - Y_i) \frac{e^{-Z_i}}{(1 + e^{-Z_i})^2}]$$

yields Variance-covariance matrix of $\hat{\boldsymbol{\beta}}$

Logit Model (6)

Variance-Covariance Matrix for $\hat{\boldsymbol{\beta}}$

$$V(\hat{\beta}) = \left[-\frac{\partial^2 \ln L}{\partial \beta_j \partial \beta_k} \right]^{-1}$$

Note that it is not the estimated VC matrix. Do Z-test or Chi-square test instead of t-test or F-test on parameters

Logit Model (7)

Interpretation

$$\frac{\partial P}{\partial X_k} = \frac{e^{-Z_i}}{(1+e^{-Z_i})^2} \beta_k = \{+\} \beta_k$$

sign of β_k ==>direction of the effect of X_k on the probability to JOIN.

Logit Model (8)

No R^2 for a logit model since there is no error term.

Define $$pseudo-R^2 = \frac{\#\,\text{correct prediction}}{\text{sample size (n)}}$$

It is a measure for goodness-of-fit.

$\widehat{JOIN}>0.5$ ==> predict that JOIN=1

$\widehat{JOIN}<0.5$ ==> predict that JOIN=0

Logistic Regression (1)

Assumption of Logistic Regression

F(.) is a logistic function but the observation(experiment) for each given set of independent variables (**X**) will be repeated several times. Only the proportion of JOIN=1 can be observed.

Logistic Regression (2)

From Logit Model

$$\ln\left(\frac{P}{1-P}\right) = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K$$

Note that P is the expected proportion of population **JOIN**ing given X's

Logistic Regression (3)

Define

R_i=observed proportion of observation with the same value of $\mathbf{X}_i$ that **JOIN**.

<u>Derived Model</u>

$$\ln\left(\frac{R_i}{1-R_i}\right) = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_K X_{Ki} + \nu_i$$

$$V(\nu_i) = \frac{1}{N_i R_i (1-R_i)}$$ Why?

Logistic Regression (4)

Define $w = \sqrt{N_i R_i (1 - R_i)}$

<u>Estimation</u>

$$R_i^* = \beta_1 X_{1i}^* + \beta_2 X_{2i}^* + + \beta_K X_{Ki}^* + \nu_i^*$$

where $R_i^* = w_i \ln\left(\dfrac{R_i}{1 - R_i}\right)$

$$X_{ki}^* = w_i X_{ki} \quad \text{for } k = 1, \ldots, K$$

$$\nu_i^* = w_i \nu_i$$

Logistic Regression (5)

=> OLS is BLUE

<u>Interpretation of the parameters</u> same as those for logit model as the underlying function is also logistic

Probit Model (1)

<u>Assumption of Probit</u>

No error term

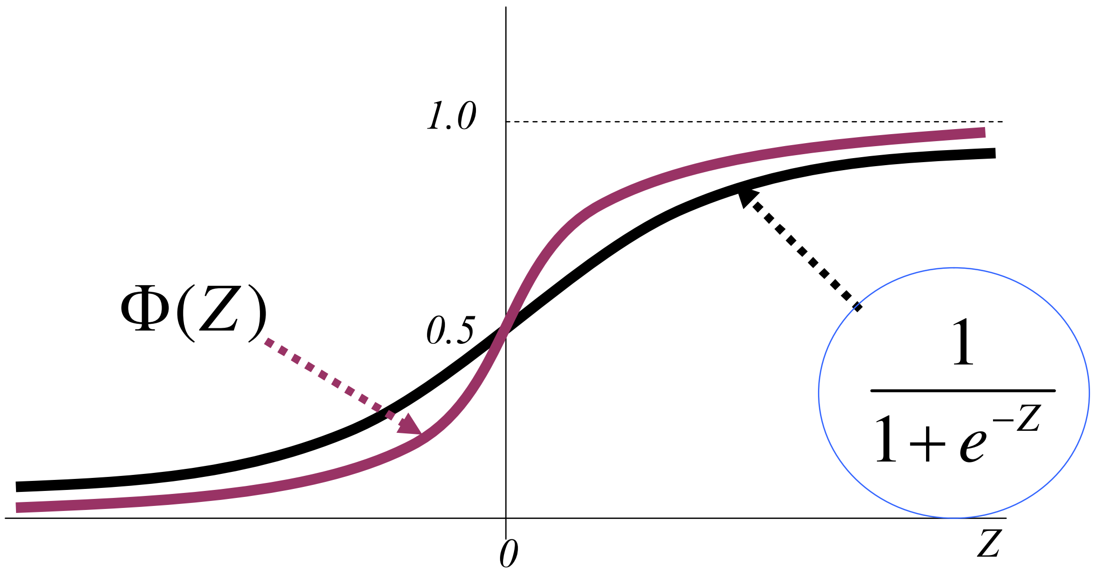
F() is a cumulative distribution function of a standard normal.

$$P = \Phi(Z)$$

$$Z = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K$$

Note that $0 \le \Phi(Z) \le 1$ always.

Probit Model (2)



Multinomial Logit Model (1)

Assumption of Multinomial Logit

Define $PA_i = \Pr(JA_i=1)$

$PB_i = \Pr(JB_i=1)$

$PC_i = \Pr(JC_i=1)$

Choose the choice of plan C as the reference.

Multinomial Logit Model (2)

$$\frac{PA_i}{PC_i} = e^{ZA_i}$$

where $ZA_i = \alpha_1 X_{1i} + \alpha_2 X_{2i} + \cdots + \alpha_K X_{Ki}$

$$\frac{PB_i}{PC_i} = e^{ZB_i}$$

where $ZB_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_K X_{Ki}$

Multinomial Logit Model (3)

$$\frac{PA_i + PB_i}{PC_i} = e^{ZA_i} + e^{ZB_i}$$

$$1 + \frac{PA_i + PB_i}{PC_i} = 1 + e^{ZA_i} + e^{ZB_i}$$

$$PC_i = \frac{1}{1 + e^{ZA_i} + e^{ZB_i}}$$

Multinomial Logit Model (4)

$$PA_i = \frac{e^{ZA_i}}{1 + e^{ZA_i} + e^{ZB_i}}$$

$$PB_i = \frac{e^{ZB_i}}{1 + e^{ZA_i} + e^{ZB_i}}$$

Multinomial Logit Model (5)

ML Estimation of Multinomial Logit model

$$\max_{\boldsymbol{\beta}} L = \prod_{i=1}^{n} (PA_i)^{JA_i} (PB_i)^{JB_i} (1 - PA_i - PB_i)^{(1 - JA_i - JB_i)}$$

$$\text{or} \quad \max_{\boldsymbol{\beta}} \ln L = \sum_{i=1}^{n} [JA_i \ln(PA_i) + JB_i \ln(PB_i) + (1 - JA_i - JB_i) \ln(1 - PA_i - PB_i)]$$

Solving FOC yields $\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}$

Multinomial Logit Model (6)

Interpretation

$$\frac{\partial PA_i}{\partial X_{ki}} = \frac{e^{ZA_i}}{1 + e^{ZA_i} + e^{ZA_i}} \alpha_k - \frac{e^{ZA_i}}{(1 + e^{ZA_i} + e^{ZA_i})^2} \left(e^{ZA_i} \alpha_k + e^{ZB_i} \beta_k \right)$$

$$= \frac{e^{ZA_i}}{1 + e^{ZA_i} + e^{ZA_i}} \left(1 - \frac{e^{ZA_i}}{1 + e^{ZA_i} + e^{ZA_i}} \right) \alpha_k - \frac{e^{ZA_i}}{1 + e^{ZA_i} + e^{ZA_i}} \frac{e^{ZB_i}}{1 + e^{ZA_i} + e^{ZA_i}} \beta_k$$

Multinomial Logit Model (6)

Interpretation

$$\frac{\partial PA_i}{\partial X_{ki}} = \overset{+}{PA_i(1-PA_i)}\alpha_k - \overset{+}{PA_i PB_i}\beta_k$$

own-effect cross-effect

sign of α_k ==>direction of the own-effect of X_k on the probability to JOIN A.

sign of β_k ==>direction of the cross-effect of X_k on the probability to JOIN A.

Other Choice Models

- Nested Logit /Serial Logit
- Ordered Logit
- Generalized Extreme-Value (GEV)

LIMDEP

Models for Limited Dependent Varaibles

- Censored Regression
- Tobit Models