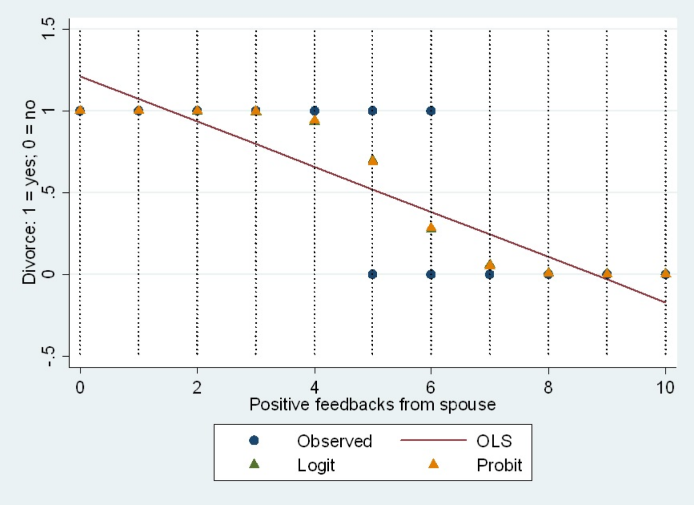
Binary Dependent Variables

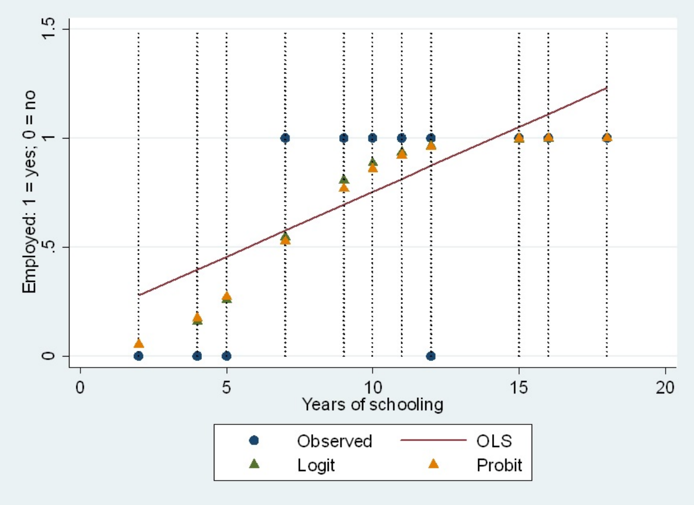
1. Examples
2. Latent variable framework
3. Probit
4. Logit
5. Maximum likelihood estimation
6. Interpretation

Examples



Divorce: 1 = yes; 0 = no
-.5
0
.5
1
1.5
0
2
4
6
8
10
Positive feedbacks from spouse
Observed
OLS
Logit
Probit

Examples (cont.)



Latent Variable Framework

- Latent (unobserved) variable

$$y^* = \beta_0 + \beta_1 x + u$$

- Such that

$$\begin{aligned} y &= 0 \text{ if } y^* < 0 \\ &= 1 \text{ if } y^* \geq 0 \end{aligned}$$

Latent Variable Framework (cont.)

- Observe $y = 1$ if

$$\underbrace{\beta_0 + \beta_1 x + u}_{}\quad y^* \geq 0$$

$$\geq 0$$

$$u \geq \underbrace{-\beta_0 - \beta_1 x}$$

- Thus

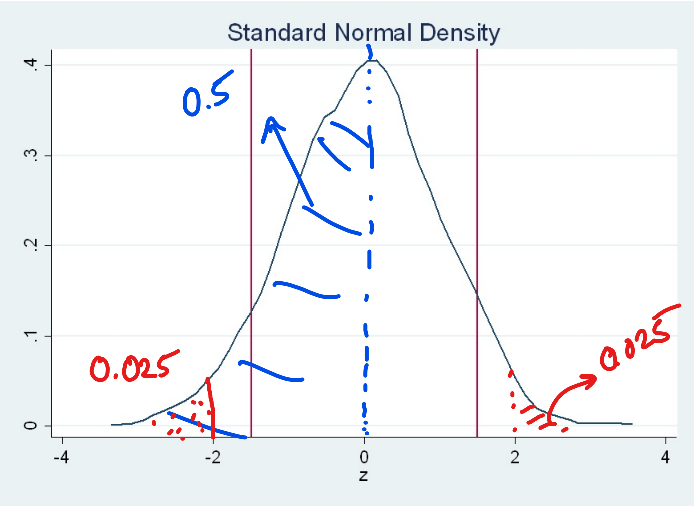
$$P(y = 1|x) = P(u \geq \;)$$

$$= P(u \geq -\beta_0 - \beta_1 x)$$

- $P(y = 1|x)$ bounded between 0 and 1

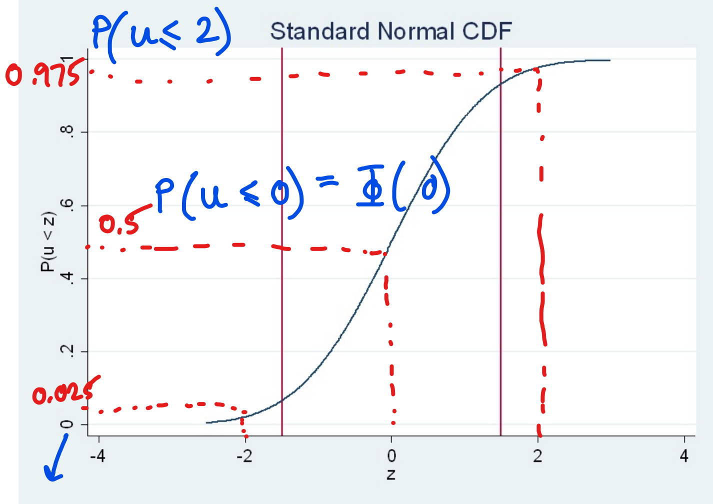
Probit

u follows $N(0, 1)$



Probit (cont.)

$\Phi(2)$



$P(u \leq -2) = \Phi(-2)$

Probit (cont.)

$-\beta_0 - \beta_1 x$

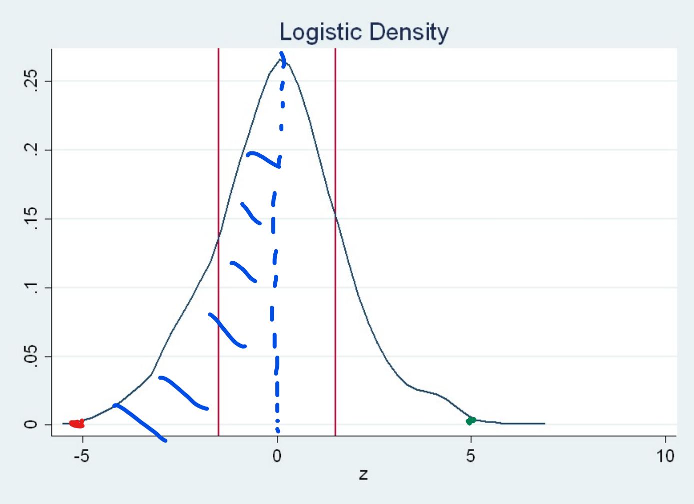
- Due to symmetry of $N(0,1)$

$$\begin{aligned} P(y=1|x) &= P(u \geq) \\ &= P(u \leq) \quad \beta_0 + \beta_1 x \\ &= \Phi(\beta_0 + \beta_1 x) \end{aligned}$$

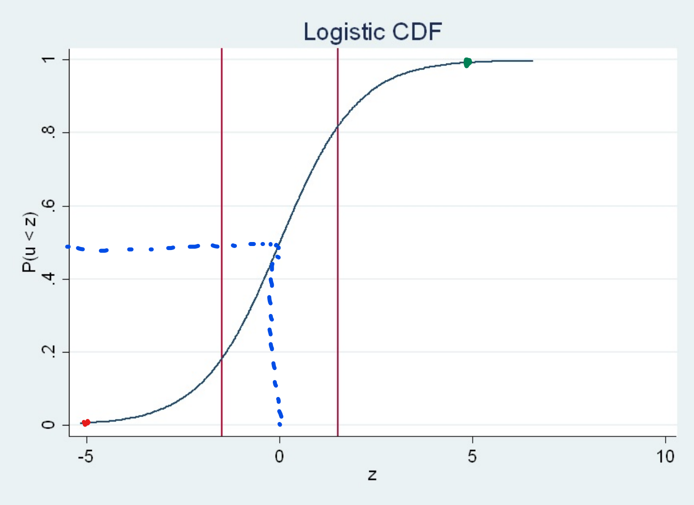
- $\widehat{\beta}_0$ and $\widehat{\beta}_1$ - from maximum likelihood estimation (MLE)

Logit

u follows logistic distribution



Logit (cont.)



Logit (cont.)

- Due to symmetry of the logistic distribution

$$
\begin{aligned}
P(y = 1|x) &= P(u \geq -\beta_0 - \beta_1 x) \\
&= P(u \leq \beta_0 + \beta_1 x) \\
&= \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)}
\end{aligned}
$$

- $\widehat{\beta}_0$ and $\widehat{\beta}_1$ - from MLE

Maximum Likelihood Estimation

- Example
  - 20% of population - below 15 years
  - Random sample of 3 people
  - Joint probability or likelihood (L) of 2 under 15 and 1 over 15

$$L = 0.2 \times 0.2 \times 0.8$$

$= 0.03$

Maximum Likelihood Estimation (cont.)

MLE : finds an estimate of $p_{ins.}$ that maximizes the likelihood of observing the data that we actually observe.

- Example
  - $p_{insured}$: probability of insured
  - Random sample of 3 people: 2 insured and 1 uninsured
  - Joint probability or likelihood (L) of observing this

$$L = p_{insured} \times p_{insured} \times (1 - p_{insured})$$
$$= p_{insured}^2 - p_{insured}^3$$

  - MLE finds $p_{insured}$ that maximizes L

$p_{ins.} = 0 \rightarrow L = 0$

$= 0.5 \rightarrow L = 0.125$

$= 0.7 \rightarrow L = 0.147$

Try diff. values or use calculus.

value of $p_{ins.}$ that max. $L \rightarrow p_{ins.} = 2/3$ Makes sense!

Maximum Likelihood Estimation (cont.)

x_i : x for obs i

- In case of probit or logit with $y = 1$ and $y = 0$ for insured and uninsured
- $P(y = 1|x) = G(\beta_0 + \beta_1 x)$ and

likelihood

$$L = G(\beta_0 + \beta_1 x_1) \times G(\beta_0 + \beta_1 x_2) \times [1 - G(\beta_0 + \beta_1 x_3)]$$

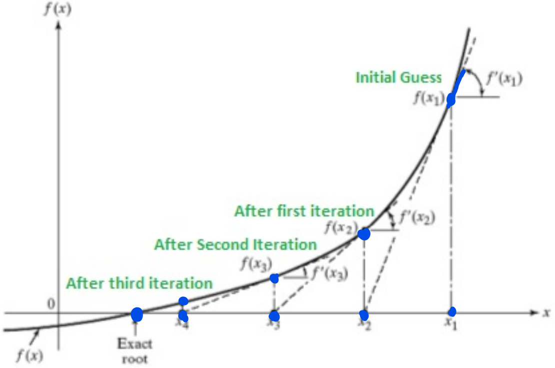
- MLE finds $\widehat{\beta}_0$ and $\widehat{\beta}_1$ that maximizes L or $\log(L)$

$$\log(L) = \log(\ldots x_1) + \log(\ldots x_2) + \log(\ldots x_3)$$

log likelihood

Maximum Likelihood Estimation (cont.)

Nonlinear optimization



`geeksforgeeks.org`

Interpretation

- Probit - continuous x

$$\frac{\Delta P(y=1|x)}{\Delta x} = \left(\text{std. normal density at } x\right) \times \beta_1$$

- Logit - continuous x

$$\frac{\Delta P(y=1|x)}{\Delta x} = \left(\text{logistic density at } x\right) \times \beta_1$$

- The effects depend on x

Interpretation (cont.)

- Typically calculate
  - Effect at the average value of x
  - The average of the effects across all values of x