

ch. 17 : Binary Dep. Vars.

Latent Variable Framework

Latent (unobserved) var.

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

such that

$$y = 0 \quad \text{if } y^* < 0$$

$$= 1 \quad \text{if } y^* \geq 0$$

↓

$$\beta_0 + \beta_1 x + u \geq 0$$

$$\text{or, } u \geq -\beta_0 - \beta_1 x$$

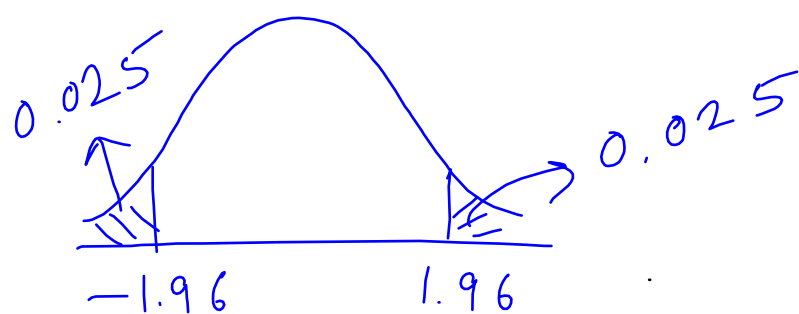
$$Pr(y=1|x) = Pr(u \geq -\beta_0 - \beta_1 x)$$

↓

bounded b/w 0 and 1

Probit

$$u \sim N(0, 1)$$



$$P(y=1|x) = P(u \geq -\beta_0 - \beta_1 x)$$

$$= P(u \leq \beta_0 + \beta_1 x)$$

$$= \Phi(\beta_0 + \beta_1 x)$$

$\hat{\beta}_0, \hat{\beta}_1$: from maximum likelihood estimation (MLE)

Logit

$u \sim \text{logistic distribution}$

$$P(u \leq \beta_0 + \beta_1 x) = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)}$$

$\hat{\beta}_0, \hat{\beta}_1$: from MLE

Maximum Likelihood Estimation

Example

$$P(\text{below } 15) = 0.2$$

Random sample of 3

It. prob. or likelihood of 2 under 15
& 1 at least 15

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

Example

$$P_{\text{insured}} = \text{prob. of insured}$$

Random sample of 3 $\rightarrow$ 2 insured
1 not "

It. prob. or likelihood of observing this

$$L = P_{\text{ins.}} \times P_{\text{ins.}} \times (1 - P_{\text{ins.}})$$

MLE finds value of $P_{\text{ins.}}$ that maximizes L .

$\hookrightarrow$ an estimate of $P_{\text{ins.}}$ that maximizes the likelihood of observing the data that we obtain

$P_{\text{ins.}}$	$\nearrow$	0	$\rightarrow$	$L = 0$	Try diff. values or use calculus.
	$\rightarrow$	0.5	$\rightarrow$	$L = 0.125$	
	$\searrow$	0.7	$\rightarrow$	$L = 0.147$	

$$\text{Value that max. } L \rightarrow P_{\text{ins.}} = 2/3$$

In case of Logit / Probit with

$y = 1$	and	0
$\downarrow$		$\downarrow$
ins.		not ins.

$$P(y=1 | x) = G(\beta_0 + \beta_1 x)$$

$$L = G(\cdot) \times G(\cdot) \times [1 - G(\cdot)]$$

MLE : finds $\hat{\beta}_0$ and $\hat{\beta}_1$ that
maximizes L or $\log(L)$.

Interpretation

Probit $\rightarrow$ continuous x

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

Logit $\rightarrow$ continuous x

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

Effects depend on x .

Typically calculate

- └ effect at avg. value of x
- └ avg. of effects across all values of x

$$\text{LPM} : \hat{\beta}_1 = \boxed{0.06}$$

$$\begin{array}{l} \text{Logit} \\ \text{Probit} \end{array} : \hat{\beta}_1 = 0.335 \quad \hat{\beta}_1 = 0.618$$

$$\boxed{0.030}$$

$$\text{effect at avg. } \bar{x} = \boxed{0.040}$$

$$\text{avg. of all effects} = 0.043$$

$$0.038$$