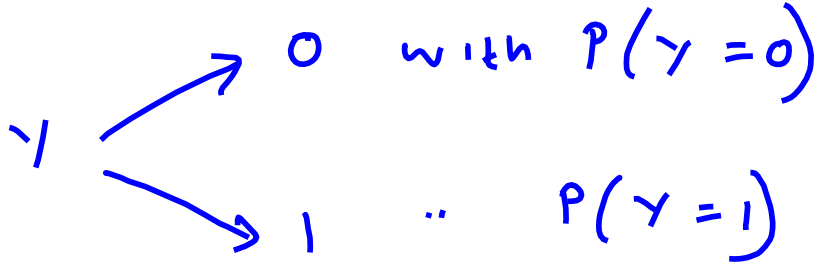
The Linear Probability Model

Dep var → dummy e.g. employment status
graduation "
sign treaty



$$E(Y) = [0 \times P(Y=0)] + [1 \times P(Y=1)] = P(Y=1)$$

$$y = \beta_0 + \beta_1 x_1 + \quad + \beta_k x_k + u$$

$$E(y \mid x_1, \quad, x_k) = \beta_0 + \beta_1 x_1 + \quad + \beta_k x_k$$

$$\downarrow$$

$$P(y = 1 \mid x_1, \quad, x_k)$$

slope coeffs

$$\frac{\Delta P(y = 1 \mid x_1 \quad x_k)}{\Delta x_j} = \beta_j$$

e.g. $P(\text{sign} = 1 \mid \text{GDPPC}, \text{democracy})$

$P(\text{empl} = 1 \mid \text{educ}, \text{job training})$

MROZ

$$inlf = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 kids1t6 + u$$

inlf ↓ : labor force participation status

kids1t6 ↓ : # kids < 6

$\hat{\beta}_3 = -0.303$

One addl child ↓ prob of in LF by 0.303

4 children ↓

$0.303 \times 4 = 1.212$

Shortcomings of the LPM

- fitted / predicted values can be outside $[0, 1]$

- constant $\hat{\beta}$ may $\Rightarrow$ unrealistic interpretn

  - May perform well near avg values in the data

- Heteroskedasticity

$$\text{var}(y|x) = P(y=1|x)\left[1 - P(y=1|x)\right]$$