Ch 7 Dummy / binary indep var

y	x	u
trade / env performance	trade / env agreement	political preferences
wage	marital status	work ethic, reliability

Single dummy variable

$$y = \beta_0 + \delta_0 x + u$$

x: 0/1

$$E(y|x) = \beta_0 + \delta_0 x$$

effect of $x=1$ rel to $x=0$ (base/reference group)

$$\Rightarrow \delta_0 = E(y|x=1) - E(y|x=0)$$

$$\hat{\delta}_0 = \bar{y}_{x=1} - \bar{y}_{x=0}$$

$$= \overline{wage}_{marr} - \overline{wage}_{not\ marr}$$

if $y \rightarrow$ wage

$x \rightarrow 1$ for marr

0 o w

$$y = \beta_0 + \beta_1 x_1 + \delta_0 x_2 + u$$

x_2: 0/1

if y wage, x_1 educ, x_2 1 for marr, 0 o w

$$E(y \mid x_1, x_2) = \beta_0 + \beta_1 x_1 + \delta_0 x_2$$

$$\delta_0 = E(y \mid x_1, x_2 = 1) - E(y \mid x_1, x_2 = 0)$$

effect of $x_2 = 1$ rel to $x_2 = 0$ after controlling for x_1

$x_2 = 0$ → base / reference group

Note 2 groups → denoted by a single dummy / binary var

x_3 1 → not marr

0 → marr

Not req'd · x_2 and x_3 are perfectly collinear

x_2	x_3
1	0
0	1

→ Dummy variable trap

If y log (wage)

approximate effect of $x_2 = 1$ $\longrightarrow$ $100\,\hat{\delta}_0$ /

exact effect $\longrightarrow$ $100\left[\exp(\hat{\delta}_0) - 1\right]$ /

e g $\hat{\delta}_0 = 0\,26 \longrightarrow$ 26 / (approx)

297 / (exact)

Multiple Categories

M ⟶ 1 married

0 not "

W ⟶ 1 (western region)

0 (not " ")

4 groups ⟶ choose 1 as base / reference group & include dummies for rest

$$y = \beta_0 + \beta_1 x_1 + \delta_0 x_2 + \delta_1 x_3 + \delta_2 x_4 + u$$

wage — educ

$x_2 = 1$ if $M = 1$, $W = 1$

$= 0$ ow

$x_3 = 1$ if $M = 1$, $W = 0$

$= 0$ ow

$x_4 = 1$ if $M = 0$, $W = 1$

$= 0$ ow

base/reference group $\rightarrow M = 0, W = 0$

δ_0 effect of $x_2 = 1$ rel to base

δ_1 " $x_3 = 1$ " Effect of $x_2 = 1$ rel to

δ_2 " $x_4 = 1$ " $x_3 = 1$

$= (\delta_0 - \delta_1)$

$\hat{\delta}_0 = 272$ $\hat{\delta}_1 = 1.3$ $\hat{\delta}_2 = 002$

Effect of marr & western region rel to base = $272

" " " & not " " " = $13

Interactions Among Dummy Vars

$$y = \beta_0 + \beta_1 x_1 + \delta_M M + \delta_W W + \delta_{MW} M \times W + u$$

base/ref group $M = 0$ & $W = 0$ $\longrightarrow = \$272$

effect of $M = 1$ & $W = 1$ rel to base

$\delta_M + \delta_W + \delta_{MW}$

δ_M $\longrightarrow$ \$1.3

effect of $M = 1$ & $W = 0$ rel to base

effect of $M = 0$ & $W = 1$ $\cdot \delta_W$

$\hat{\delta}_M = 13$

$\hat{\delta}_W = 002$

$\delta_{MW} = 14$

Allow for different slopes

$$y = \beta_0 + \beta_1 x_1 + \delta_0 M + \delta_1 M \times x_1 + u$$

- y → wage
- x_1 → educ
- M → 1 → married, 0 → otherwise

eg $\hat{\beta}_1 = 0.46$, $\hat{\delta}_0 = 0.33$, $\hat{\delta}_1 = 0.1$

	Intercept	Slope on educ
Marr	$\beta_0 + \delta_0$	$\beta_1 + \delta_1$
Not marr	β_0	β_1
Diff	δ_0	δ_1

The Linear Probability Model (LPM)

Dep var → dummy e g employment status,
graduation "
signing treaty

Y → 0 with $P(Y=0)$

Y → 1 with $P(Y=1)$

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

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

$$P(y=1 \mid x_1 \quad x_k) = \beta_0 + \beta_1 x_1 + \quad + \beta_k x_k$$

↳ instead of $E(y \mid x_1 \quad x_k)$

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

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

Slope coefficients (continuous x_j) $= \dfrac{\Delta P(y=1 \mid x_1 \quad x_k)}{\Delta x_j} = \beta_j$

MRO2

$inlf = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 kids16 + u$

$inlf$ ↓ labor force participation status

$\hat{\beta}_3 = -0303$ (four addl children)

one addl ch'd ↓ prob of being in LF by 0303

$kids16$ ↓ # kids < 6

$(4 \times 0303) = 1212$

shortcomings of the LPM

- Fitted / predicted values can be outside $[0, 1]$
- Constant $\hat{\beta}$ may $\Rightarrow$ unrealistic interpretation

- Heteroskedasticity

$$Var(y|x) = P(y=1|x)\left[1 - P(y=1|x)\right]$$

- May perform well near avg values in data