# Dummy Variables

$$y_i = \beta_0 + \beta_1 \cdot c_i + \beta_2 \cdot x_i + \varepsilon_i$$

$$c_i = \begin{cases} 1, & \text{if bad} \\ 2, & \text{if medium} \\ 3, & \text{if good} \end{cases}$$

$$y_i = \beta_0 + \beta_1 \cdot d_b + \beta_2 \cdot d_g + \beta_3 \cdot x_i + \varepsilon_i$$

$$+ \beta_4 \cdot d_m$$

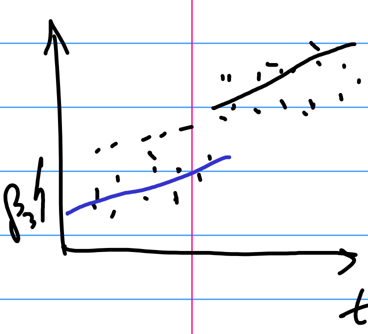
$$\begin{bmatrix} 1 & 1 & 0 & 0 & x_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & 1 & 0 & \vdots & \vdots \\ \vdots & 0 & 1 & 0 & \\ \vdots & \vdots & 1 & 1 & \\ \vdots & \vdots & 0 & \vdots & \\ 1 & 0 & 0 & 1 & x_n \end{bmatrix}$$

dummy variable trap

$Q_1$. a) Change in intercept:

$$Y_i = \beta_1 + \beta_2 \cdot P_i + \beta_3 D_i + \varepsilon_i$$

$Y_i$: consump. $P_i$: price $D_i$: before / after crisis

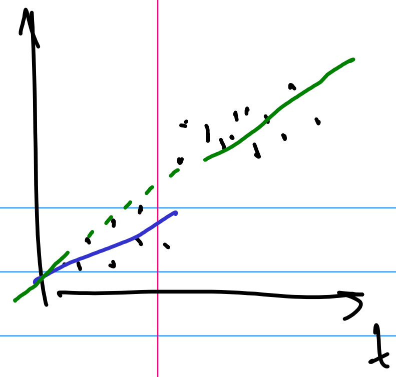



Before crisis : $Y_i = \beta_1 + \beta_2 \cdot P_i + \varepsilon_i$

After crisis : $Y_i = (\beta_1 + \beta_3) + \beta_2 P_i + \varepsilon_i$

b) Change in slope:

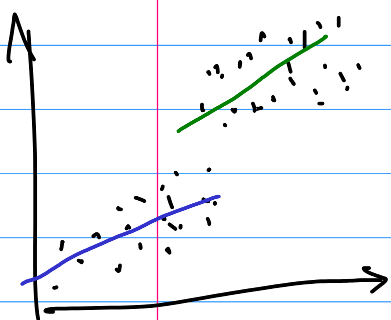
$$Y_i = \beta_1 + \beta_2 P_i + \beta_3 \cdot P_i \cdot D_i + \varepsilon_i$$




Before crisis : $Y_i = \beta_1 + \beta_2 \cdot P_i + \varepsilon_i$

After crisis : $Y_i = \beta_1 + (\beta_2 + \beta_3) P_i + \varepsilon_i$

c) Changes in both intercept and slope



$$Y_i = \beta_1 + \beta_2 \cdot D_i + \beta_3 P_i + \beta_4 \cdot P_i D_i + \varepsilon_i$$

Before crisis: $Y_i = \beta_1 + \beta_3 \cdot P_i + \varepsilon_i$

After crisis : $Y_i = (\beta_1 + \beta_2) + (\beta_3 + \beta_4) P_i + \varepsilon_i$

F - test:  $H_0: \beta_2 = \beta_4 = 0$

$H_a: \ldots$

$$F = \frac{(RSS_R - RSS_{UR})/2}{RSS_{UR}/(N-4)}$$

Chow test:

Pooled regression: $y_i = \beta_1 + \beta_2 x_{1i} + \ldots + \beta_k x_k + \varepsilon_i$

$n = n_A + n_B$

Sub. A: $y_i = \beta_1^A + \beta_2^A x_{1i} + \ldots + \beta_k^A x_{ki} + \varepsilon_i$

Sub B: $y_i = \beta_1^B + \beta_2^B x_{1i} + \ldots + \beta_k^B x_{ki} + \varepsilon_i$

**Problem 2. (ICEF exam)** A student decided to investigate the market of private mathematics teachers in Moscow, with particular interest to those who can teach in English. He took a random sample of 30 profiles of teachers who provide private teaching in math (taken from population of 300 profiles registered in certain internet site) and run some regressions trying to find factors influencing the prices of teaching ($PRICE_i$ - price of a standard two-hour lesson in thousands of roubles, $DIST_i$ - distance in the number of metro stations from the center of Moscow to the teacher's place, $HOME_i$ - dummy variable indicating visit of the tutor to the client, $ENG_i$ - dummy variable indicating ability to teach the subject in English):

Home - 0 (A)

Home = 1 (B)

$$\widehat{PRICE}_i = \underset{(0.49)}{6.59} - \underset{(0.06)}{0.16} DIST_i \qquad R^2 = 0.185 \tag{1}$$

$$\widehat{PRICE}_i = \underset{(0.40)}{4.51} + \underset{(0.58)}{2.54} HOME_i \qquad R^2 = 0.40 \tag{2}$$

$$\widehat{PRICE}_i = \underset{(0.64)}{5.13} - \underset{(0.06)}{0.08} DIST_i + \underset{(0.95)}{1.95} HOME_i + \underset{(0.07)}{0.07} DIST * HOME_i \qquad R^2 = 0.437 \tag{3}$$

? ?

$$\widehat{PRICE}_i = \underset{(0.61)}{4.52} - \underset{(0.06)}{0.08} DIST_i + \underset{(0.75)}{2.18} HOME_i + \underset{(0.76)}{1.58} ENG_i - \underset{(1.09)}{0.39} HOME * ENG_i \qquad R^2 = 0.553 \tag{4}$$

{he He}
{hE HE}

a) $4,51 + 2,54 \approx 7,05$

b) if $Home_i = 1$ $\quad P_i = 7,08 - 0,01 D_i + \hat{a}_i$

c) is (3) significant:

$$F = \frac{0,437/3}{(1-0,437)/(30-4)} = 6,73 \qquad F_{crit,\,1\%} = 4,6$$

is "Dist" significant:

R: $\quad P_i = \beta_0 + \beta_2 \cdot H_i + \varepsilon_i$

UR: ...

$H_0: \beta_1 = \beta_3 = 0$

is factor "teaching at home" is sign.:

F-test $\quad H_0: \beta_2 = \beta_3 = 0$

$H_0: \beta_1^A = \beta_1^B, \ldots\ldots, \beta_k^A = \beta_k^B$

$H_a:$ at least one isn't held

d) $$F = \frac{(RSS_P - RSS_A - RSS_B)/k}{(RSS_A + RSS_B)/(n-2k)} \sim F(k, n-2k)$$

$$F = \frac{(\text{improvement in fit})/(\text{extra d.o.f.})}{(\text{remaining RSS})/(\text{remaining d.o.f.})}$$

for (4):

$$F = \frac{\left(RSS_P - (RSS_{he} + RSS_{He} + RSS_{hE} + RSS_{HE})\right)/2\cdot 3}{(RSS_{he} + RSS_{He} + RSS_{hE} + RSS_{HE})/(30 - 2\cdot 4)}$$