

TEMA 7: Dummy variables

- ① Motivasjon og hovedkjennetegn
- ② Både kvalitative og kvantitative forklaringsvariabler
- ③ Kvalitative forklaringsvariabler med mer enn to verdier
- ④ Stigningsfall dummyer
- ⑤ Oppgavesett 7

1 Motivasjon og hovedkjennetegn

- Kjønn: Mann vs. kvinne
- Hudfarge: Hvit vs ikke-hvit
- Alder: Myndig vs. ikke-myndig
- Bosfed: Nord, sør, vest og øst

Def: Dummy-variabel: En variabel som kan anta to verdier; 1 og 0

(MERK: Samme som indikatorvariable)

Eksempel. Y_i = inntekten til person i
 D_i = kjønn til person i , dvs. $D_i = 0$
 hvis mann og $D_i = 1$ hvis kvinne
 Anta at vi ønsker å estimere:

$$Y_i = \beta_1 + \beta_2 D_i + u_i$$

Det betyr:

→ Gj. snittlig inntekt for menn:

$$b_1 + b_2 \cdot 0 = b_1$$

→ Gj. snittlig inntekt for kvinner:

$$b_1 + b_2 \cdot 1 = b_1 + b_2$$

Wage = time lønn i USD

female = 1 hvis kvinne, 0 ellers

EV views: wage \subset female

$$\widehat{wage}_i = 7,0995 - 2,5118 \text{ female}$$

Gj. snittlig time lønn for menn

$$\begin{aligned} \widehat{wage}_i &= 7,0995 - 2,5118 \cdot 0 \\ &= \underline{\underline{7,0995 \text{ USD}}} \end{aligned}$$

Gj. snittlig time lønn for kvinner

$$\begin{aligned} \widehat{wage}_i &= 7,0995 - 2,5118 \cdot 1 \\ &= \underline{\underline{4,5887 \text{ USD}}} \end{aligned}$$

Er forskjellen signifikant?

- (2) Både kvalitative og kvantitative forklaringsvariabler (ANCOVA modeller)

Eksempel (s. 185). Y_i = matkonsum for individ i , X_i = disponibel inntekt for individ i , D_i = kjønn til individ i (1 hvis kvinne, 0 hvis mann)

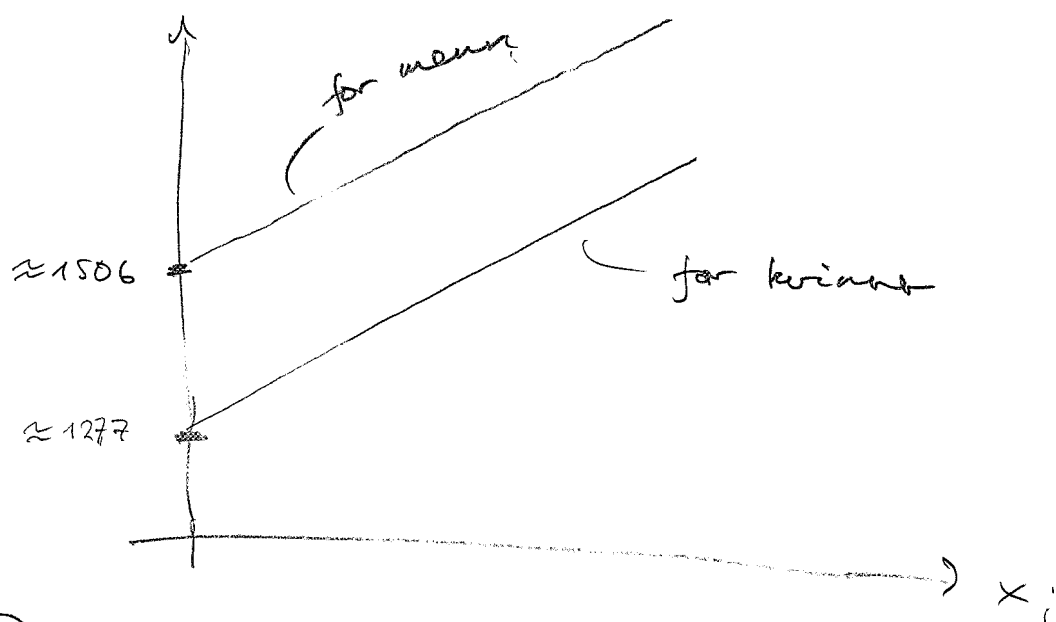
$$Y_i = \beta_1 + \beta_2 D_i + \beta_3 X_i + u_i$$

$$\hat{Y}_i = 1506,244 - 228,9868 D_i + 0,0589 X_i$$

Implikasjon:

-> For kvinner: $\hat{Y}_i = 1277,2574 + 0,0589 X_i$

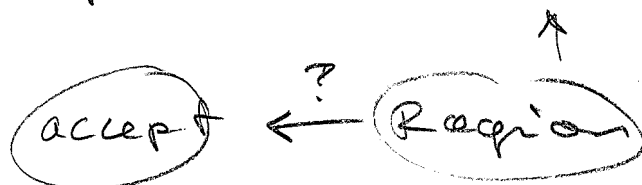
-> For menn: $\hat{Y}_i = 1506,2440 + 0,0589 X_i$



③ Kvalitative forklaringsvariable med mer enn to verdier

Eksempel: Region = Nord, sør, vest, øst

Eksempel (s. 187). Accept_i = %-andelen av søkerne som har kommet inn på skole i, Region = vest, nord, sør



Definer: $D_{2i} = 1$ hvis skolen ligger
i nord, 0 ellers

$D_{3i} = 1$ hvis skolen ligger
i sør, 0 ellers

$D_{4i} = 1$ hvis skolen ligger
i vest, 0 ellers

$$Y_i = B_2 D_{2i} + B_3 D_{3i} + B_4 D_{4i} + u$$

Skolene i nord: $\hat{Y}_i = b_2 \cdot 1 + b_3 \cdot 0 + b_4 \cdot 0$
 $= \underline{b_2}$

Skolene i sør: $\hat{Y}_i = b_2 \cdot 0 + b_3 \cdot 1 + b_4 \cdot 0$
 $= \underline{b_3}$

Skolene i vest: $\hat{Y}_i = b_2 \cdot 0 + b_3 \cdot 0 + b_4 \cdot 1$
 $= \underline{b_4}$

Hva med B_1 ?

→ Hvis vi har en kvalitativ variabel med r ulike verdier, så kan vi ikke inkludere konstanten hvis vi inkluderer r dummy-variabler

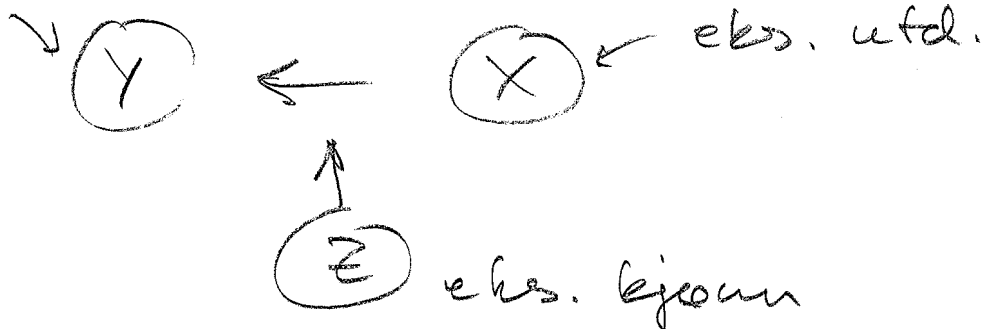
↳ stikkord: multikolinearitet

→ Hvis vi en kvalitativ variabel med r ulike verdier, så kan vi

vi inkluderer konstanten hvis vi
kun inkluderer $(k-1)$ dummy-
variabler

(4) Stigningsfalls dummyer

eks.
innt.



Eksempel. Datasett: Time lønn i USA

wage = time lønn i USA

female = 1 hvis kvinne, 0 ellers

educ = utdanning i ant. år

Anta at vi ønsker å estimere:

$$\text{wage} = \beta_1 + \beta_2 \text{educ} + \beta_3 \text{educ} \cdot \text{female} + u$$

For kvinner:

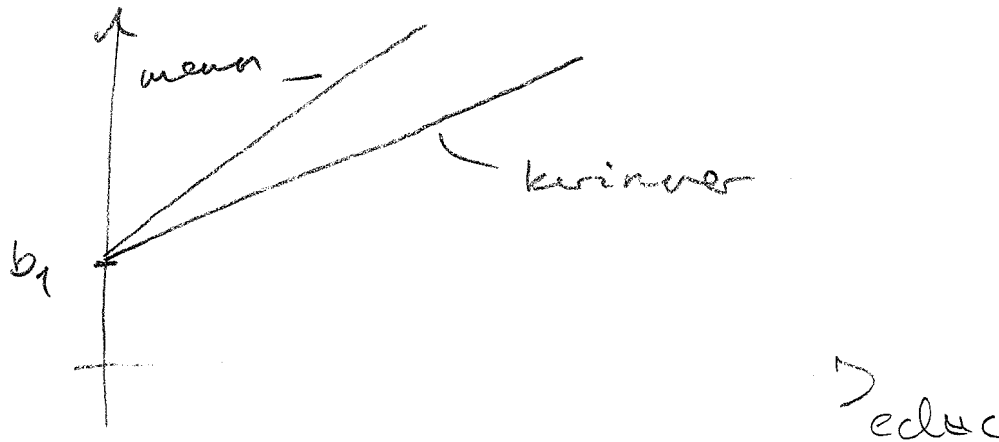
$$\widehat{\text{wage}} = b_1 + (b_2 + b_3) \cdot \text{educ}$$

$$= -0,2853 + \underbrace{(0,5754 - 0,1776)}_{0,3978} \text{educ}$$

For menn:

$$\widehat{\text{wage}} = b_1 + b_2 \text{educ}$$

$$= -0,2853 + 0,5754 \cdot \text{educ}$$



Tolkning av eksempelet: Et år mer i utdanning for menn gir en større økning i lønnsinntekten enn for kvinner (i gjennomsnitt)

Eksempel.

$$\text{wage} = B_1 + \underbrace{B_2 \text{ female}}_{\text{legges til}} + B_3 \text{ educ} + B_4 \cdot \text{educ} \cdot \text{female} + u$$

EViews: wage c female educ educ*female

$$\widehat{\text{wage}} = 0,2005 - 1,1985 \cdot \text{female} + 0,5395 \text{ educ} - 0,0860 \text{ educ} \cdot \text{female}$$

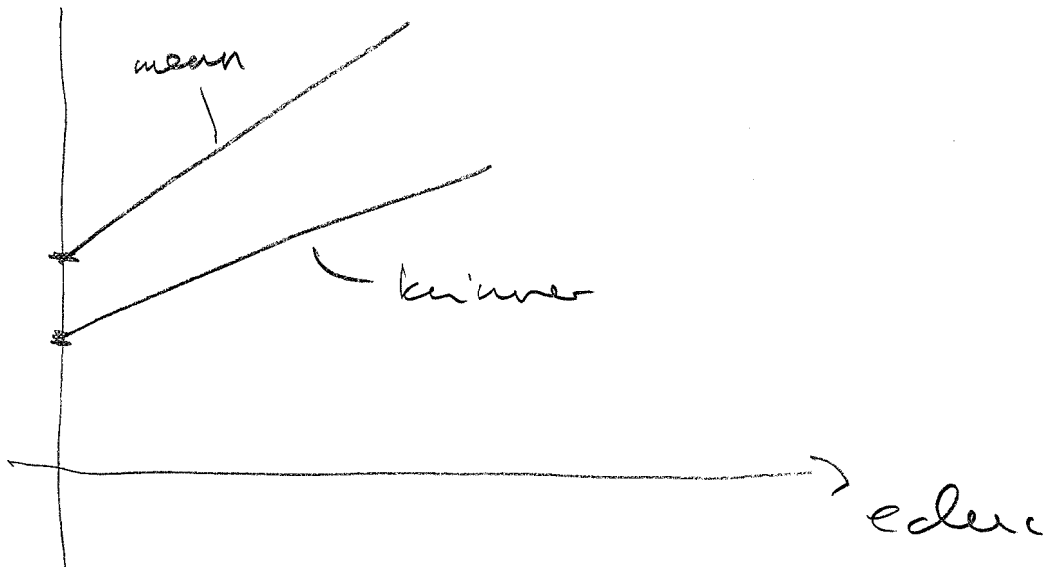
Kvinner:

$$\widehat{\text{wage}} = (0,2005 - 1,1985) + (0,5395 - 0,0860) \cdot \text{educ}$$

$$= -0,998 + 0,4535 \cdot \text{educ}$$

Menn:

$$\widehat{\text{wage}} = 0,2005 + 0,5395 \text{ educ}$$



Tolkningen: Kvinner tjener 1,1985 USD mindre (i gjennomsnitt) enn menn og effekten av 1 år mer i utdanning er 0,086 mindre for kvinner

⑤ Oppgavesett 7

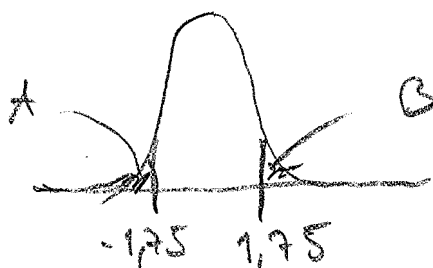
$$1. \hat{\beta}_{ie} = 0,114 - 0,021 \cdot 1$$

$$= \underline{0,093} = \underline{9,3\%}$$

B)

$$2. t = \frac{\hat{\alpha}_2}{se(\hat{\alpha}_2)} = \frac{-0,021}{0,012} = -1,75$$

P-verdi for tosidig test: $2 \cdot P(t > |-1,75|)$



$A = B$



$$\text{Ant. frihetsgrader} : n - k = 16$$
$$\begin{array}{ccc} & \uparrow & \uparrow \\ & 18 & 2 \end{array}$$

$$P(t > | -1,75 |) \approx 0,05$$

$$\Rightarrow P\text{-verdi} \approx 0,10 = 10\%$$

A)

$$3. E(\text{Income}) = \beta_1 + \beta_2 \cdot 1 = \underline{\underline{\beta_1 + \beta_2}}$$

B)