

UNIVERSITETET I OSLO

ØKONOMISK INSTITUTT

Utsatt eksamen i: **ECON4150 – Introductory econometrics**

Eksamensdag: Onsdag 7. januar 2009

Tid for eksamen: kl. 09:00 – 12:00

Oppgavesettet er på 4 sider

Tillatte hjelpemidler:

- Alle trykte og skrevne hjelpemidler, samt lommekalkulator er tillatt

Eksamen blir vurdert etter ECTS-skalaen. A-F, der A er beste karakter og E er dårligste ståkarakter. F er ikke bestått.

For å ta hensyn til kvalitative forklaringsvariable i en regresjonsligning benytter vi gjerne dummyvariable.

Spørsmål 1

Gi eksempler på økonometriske anvendelser av dummyvariable.

På et utvalg av norske lønnstakere har vi estimert regresjonen

$$(1) \quad X_i = \beta_0 + \beta_1 K_i + \beta_2 E_i + \varepsilon_i \quad i = 1, 2, 3, \dots, 200$$

der X_i betegner timelønnen til lønnstaker (i), K_i betegner dummyvariabelen for kjønn, som er lik 1 hvis person i er en kvinne og lik 0 hvis person i er en mann, E_i betegner antall år utdanning til lønnstaker (i). Det er like mange kvinner som menn i utvalget. Variabelen ε_i betegner det stokastiske restledd i regresjonen, som antas å være uavhengige og identisk fordelte for alle i , med forventning 0 og varians σ^2 .

Utskrift 1 viser resultatet av regresjon (1) på vårt datasett.

Spørsmål 2

Kommenter de empiriske resultatene. Hva vil du trekke frem som spesielt interessant ved disse resultatene.

Utskrift 1 viser at effekten av utdanning på timelønnen er usikker. Vi har derfor også estimert lønnsrelasjonen når utdanning er ekskludert, dvs.

$$(2) \quad X_i = \beta_0 + \beta_1 K_i + \varepsilon_i$$

Minstekvadraters estimatorene på regresjonskoeffisientene β_0 og β_1 i regresjonen (2) er gitt ved formlene

$$(a) \quad \hat{\beta}_0 = \bar{X} - \hat{\beta}_1 \bar{K}$$

$$(b) \quad \hat{\beta}_1 = \frac{\sum_{i=1}^{200} (K_i - \bar{K}) X_i}{\sum_{i=1}^{200} (K_i - \bar{K})^2}$$

der \bar{X} betegner gjennomsnittlig timelønn i datasettet og \bar{K} betegner gjennomsnittet av dummyvariabelen K_i .

Spørsmål 3

Vis at formlene for minstekvadraters estimatene reduserer seg til

$$\hat{\beta}_0 = \bar{X}_M \quad \text{og} \quad \hat{\beta}_1 = \bar{X}_K - \bar{X}_M$$

der \bar{X}_M betegner gjennomsnittet av timelønnen for mennene i utvalget og \bar{X}_K betegner gjennomsnittet av timelønnen for kvinnene.

Mange som forsker på arbeidsmarkedsrelaterte problemer og lønnsdannelse, vil hevde at lønnsrelasjoner er ikke-lineære, slik at de lineære regresjonene i (1) og (2) passer dårlig for dette formålet. I stedet for disse blir følgende spesifikasjoner foreslått

$$(3) \quad \ln X_i = \gamma_0 + \gamma_1 K_i + \gamma_2 E_i + \varepsilon_i$$

$$(4) \quad \ln X_i = \alpha_0 + \alpha_1 \ln K_i + \alpha_2 \ln E_i + \varepsilon_i$$

der $\ln X_i$, $\ln K_i$ og $\ln E_i$ betegner den naturlige logaritmen til de angitte variable.

Spørsmål 4

Forklar hvorfor vi ikke kan benytte regresjon (4) med vårt valg av forklaringsvariable.

Utskrift 3 viser resultatet av regresjon (3) på datasettet.

Spørsmål 5

Kommenter kort resultatene i utskrift 3. Synes du det er grunn til å foretrekke regresjon (3) fremfor regresjon (1)?

Spørsmål 6

Hvordan vil du tolke parameteren γ_1 i regresjon (3)?

Spørsmål 7

Estimer timelønnen for en mannlig arbeider med 10 års utdanning når du: (a) benytter regresjon (1), og (b) benytter regresjon (3).

Utskrift 1

EQ(1) Modelling X_i by OLS-CS (using yngvedes08-1(1).xls)

The estimation sample is: 1 to 200

	Coefficient	Std.Error	t-value	t-prob
Constant	122.955	28.01	4.39	0.000
K_i	-34.0488	9.892	-3.44	0.001
E_i	1.23641	2.542	0.486	0.627
sigma	69.8621	RSS	961501.39	
R^2	0.0586661	F(2,197) =	6.139 [0.003]**	
		DW	1.86	
no. of observations	200	no. of parameters	3	
mean(X_i)	119.005	var(X_i)	5107.12	

Utskrift 2

EQ(2) Modelling X_i by OLS-CS (using yngvedes08-1(1).xls)

The estimation sample is: 1 to 200

	Coefficient	Std.Error	t-value	t-prob
Constant	136.147	6.973	19.5	0.000
K_i	-34.2837	9.861	-3.48	0.001
sigma	69.7273	RSS	962655.671	
R^2	0.057536	F(1,198) =	12.09 [0.001]**	
		DW	1.86	
no. of observations	200	no. of parameters	2	
mean(X_i)	119.005	var(X_i)	5107.12	

Utskrift 3

EQ(3) Modelling $\ln X_i$ by OLS-CS (using yngvedes08-1(1).xls)

The estimation sample is: 1 to 200

	Coefficient	Std.Error	t-value	t-prob
Constant	4.63374	0.1453	31.9	0.000
K_i	-0.276416	0.05131	-5.39	0.000
E_i	0.0181150	0.01319	1.37	0.171
sigma	0.362414		RSS	25.8747474
R^2	0.138615		F(2,197) =	15.85 [0.000]**
			DW	1.96
no. of observations	200		no. of parameters	3
mean($\ln X_i$)	4.6871		var($\ln X_i$)	0.150193