

UNIVERSITETET I OSLO ØKONOMISK INSTITUTT

Eksamen i: ECON3150/4150 – Elementær økonometri
Exam: ECON3150/4150 – Introductory econometrics

Eksamensdag: Fredag 20. mai 2005
Date of exam: Friday, May 20, 2005

Sensur kunngjøres: Mandag 13. juni
Grades will be given: Monday, June 13

Tid for eksamen: kl. 09:00 – 12:00
Time for exam: 9:00 a.m. – 12:00 noon

Oppgavesettet er på 7 sider
The problem set covers 7 pages

English version on page 5

Tillatte hjelpemidler:

- Alle trykte og skrevne hjelpemidler, samt lommekalkulator

Resources allowed:

- *All written and printed resources, as well as calculator*

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

The grades given: A-F, with A as the best and E as the weakest passing grade. F is fail.

Kjønn og Lønn

For et utvalg av norske lønnstakere, bestående av 75 kvinner og 75 menn, har vi observert variablene: kjønn, timelønn, antall år utdanning. Det blir ofte hevdet at kvinner blir underbetalt i arbeidslivet. Vi skal undersøke denne påstanden ved bruk av regresjonsanalyse på vårt datamateriale. I analysene benytter vi følgende variabler: W_i - timelønn for lønnstaker (i), E_i - antall år utdanning for lønnstaker (i), den kvalitative variable kjønn representeres ved dummy-variablene definert ved:

$$D_F(i) = \begin{cases} 1 & \text{hvis lønnstaker (i) er kvinne} \\ 0 & \text{hvis lønnstaker (i) mann} \end{cases}, \quad D_M(i) = \begin{cases} 1 & \text{hvis lønnstaker (i) er mann} \\ 0 & \text{hvis lønnstaker (i) er kvinne} \end{cases}$$

Første del av regresjonsanalysen er basert på regresjonslikningene:

$$(1) \quad W_i = \alpha_F D_F(i) + \alpha_M D_M(i) + \gamma_1 E_i + \varepsilon_i \quad i = 1, 2, \dots, 150$$

$$(2) \quad W_i = \beta_0 + \beta_1 D_F(i) + \gamma_1 E_i + \delta_i, \quad i = 1, 2, \dots, 150$$

Spesifikasjon (1) inkluderer begge dummy-variablene, men ekskluderer konstantleddet i regresjonen. Regresjon (2) inkluderer et konstantledd, men dummy-variablen for menn $D_M(i)$ er ekskludert fra regresjonen.

Resultatene av disse regresjonene anvendt på våre lønnsdata er vist i utskriftene (1) og (2)..

Spørsmål 1

Gjør rede for sammenhengene mellom parametrene α_F , α_M i regresjon (1) og parametrene β_0 , β_1 i regresjon (2). Drøft de empiriske resultatene slik de fremgår i utskrift 1 og utskrift 2.

Spørsmål 2

I regresjonsanalyser med dummy-variable som forklaringsvariabler foretrekker man gjerne spesifikkasjon (2) fremfor spesifikkasjon (1). Hvorfor?

I det følgende vil vi basere analysen på spesifikkasjon (2). For å få et mer fullstendig bilde av variabelen kjønns betydning for timelønnen har vi også beregnet regresjonene:

$$(3) \quad W_i = \beta_0 + \beta_1 D_F(i) + \gamma_1 E_i + \gamma_2 (D_F(i)E_i) + \delta_i \quad i = 1, 2, \dots, 150$$

der variabelen $(D_F(i)E_i)$ er produktet av variablene $D_F(i)$ og E_i , og

$$(4) \quad W_i = \beta_0 + \gamma_1 E_i + \delta_i \quad i = 1, 2, \dots, 150$$

Resultatene fra disse regresjonene er gitt i utskrift (3) og utskrift (4).

Spørsmål 3

Når du baserer deg på resultatene fra de vedlagte regresjonsberegninger, mener du at de gir støtte til en påstand om at kvinner blir underbetalt i norsk arbeidslivet? Begrunn svaret. Hvis du finner støtte for påstanden, drøft mulige årsaker til at man kan observere en slik lønnsforskjell.

Spørsmål 4

Ta utgangspunkt i regresjon (2) (utskrift 2) og beregn årslønnen for en mann med 10 års utdanning og som arbeider 1800 timer i året. Beregn usikkerheten ved ditt estimat ved hjelp av estimatets standardavvik. Det blir oppgitt at den estimerte kovariansen mellom $\hat{\beta}_0$ og $\hat{\gamma}_1$ er -29.082 .

Spørsmål 5

De empiriske resultatene tyder på at utdanning er en svak forklaringsvariabel. Vi er derfor interessert i å finne ut hvordan estimatene for β_0 og β_1 i regresjon (2) vil bli endret hvis vi ekskluderer utdanning (E) i regresjon (2). Du får oppgitt at i utvalget er gjennomsnittlig utdanning for menn 12.3466 år og gjennomsnittlig utdanning for kvinner 11.9866 år. Bruk dette til å beregne de nye estimatene for β_0 og β_1 når variabelen E er fjernet fra regresjonen.

Utskrift 1 (Regresjon (1))

	Coefficient	Std.Error	t-value	t-prob
D_F	82.3653	19.28	4.27	0.000
D_M	124.549	19.80	6.29	0.000
E	2.65070	1.535	1.73	0.086

sigma 49.8728 RSS 365632.519
 R^2 0.174338 $F(2,147) = 15.52 [0.000]**$

no. of observations 150 no. of parameters 3
 mean(timelønn) 135.707 var(timelønn) 2952.24

Utskrift 2 (Regresjon (2))

	Coefficient	Std.Error	t-value	t-prob
Constant	124.549	19.80	6.29	0.000
D_F	-42.1839	8.163	-5.17	0.000
E	2.65070	1.535	1.73	0.086

sigma 49.8728 RSS 365632.519
 R^2 0.174338 $F(2,147) = 15.52 [0.000]**$

no. of observations 150 no. of parameters 3
 mean(timelønn) 135.707 var(timelønn) 2952.24

Utskrift 3 (Regresjon (3))

	Coefficient	Std.Error	t-value	t-prob
Constant	129.266	25.03	5.16	0.000
D_F	-54.0309	39.13	-1.38	0.169
E	2.26865	1.973	1.15	0.252
$(D_F E)$	0.976874	3.155	0.310	0.757

sigma 50.0269 RSS 365392.546
 R^2 0.17488 $F(3,146) = 10.31 [0.000]**$

no. of observations 150 no. of parameters 4
 mean(timelønn) 135.707 var(timelønn) 2952.24

Utskrift 4 (Regresjon (4))

	Coefficient	Std.Error	t-value	t-prob
Constant	96.9259	20.66	4.69	0.000

E	3.18752	1.659	1.92	0.057
sigma	54.0306	RSS	432057.218	
R^2	0.0243395	$F(1,148) =$	3.692	[0.057]
no. of observations	150	no. of parameters	2	
mean(timelønn)	135.707	var(timelønn)	2952.24	

Gender and Wages

For a sample of Norwegian workers consisting of 75 women and 75 men, we have observed the variables: gender, wages per hour, years of education. It is often said that women are underpaid in the labour market. We wish to investigate this assertion by applying regression analysis to our data-set. We use the following variables in our analysis: W_i - wage per hour paid to worker (i), E_i - worker (i)'s years of education, the qualitative variable gender is represented by the dummy variables defined by:

$$D_F(i) = \begin{cases} 1 & \text{if worker (i) is a woman} \\ 0 & \text{if worker (i) is a man} \end{cases}, \quad D_M(i) = \begin{cases} 1 & \text{if worker (i) is a man} \\ 0 & \text{if worker (i) is a woman} \end{cases}$$

The first part of the regression analysis is based on the equations:

$$(1) \quad W_i = \alpha_F D_F(i) + \alpha_M D_M(i) + \gamma_1 E_i + \varepsilon_i \quad i = 1, 2, \dots, 150$$

$$(2) \quad W_i = \beta_0 + \beta_1 D_F(i) + \gamma_1 E_i + \delta_i, \quad i = 1, 2, \dots, 150$$

The specification (1) includes both dummy variables in addition to education, but excludes the intercept term. The regression equation (2) includes an intercept term, but excludes the dummy variable for men $D_M(i)$.

The results of these regressions applied to our data set are shown in output (1) and output (2).

Question 1

Clarify the relations between the parameters α_F , α_M in the regression equation (1) and the parameters β_0 , β_1 in the regression equation (2). Discuss the empirical results as they are shown in output (1) and in output (2).

Question 2

In regression analyses where dummy variables enter the set of explanatory variables, one usually prefers specification (2) to specification (1). Explain why.

In order to obtain a more complete picture of the importance of gender for salaries, we have also run the regressions:

$$(3) \quad W_i = \beta_0 + \beta_1 D_F(i) + \gamma_1 E_i + \gamma_2 (D_F(i)E_i) + \delta_i \quad i = 1, 2, \dots, 150$$

where the variable $(D_F(i)E_i)$ is the product of the variables $D_F(i)$ and E_i , and

$$(4) \quad W_i = \beta_0 + \gamma_1 E_i + \delta_i \quad i = 1, 2, \dots, 150$$

The results of these regressions are shown in output (3) and output (4).

Question 3

Do you think these regression results support the assertion that women are underpaid in the Norwegian labour market? Substantiate your answer.

If you find support for this assertion, discuss possible reasons explaining that we can observe differences in wages paid to females and males.

Question 4

Use regression (2) (output (2)) to calculate the annual income for a man with 10 years' of education and who works 1800 hours per year. Calculate the standard error of the estimate as a measure of the uncertainty of the estimate. You are informed that the estimated covariance between $\hat{\beta}_0$ and $\hat{\gamma}_1$ is -29.082 .

Question 5

Our empirical results indicate that education is a weak explanatory variable in the wage equations. Therefore, we are interested in finding out how the estimates of β_0 and β_1 in the regression equation (2), will be changed if we exclude education (E) from this regression. It is stated that in this sample the mean value of men's education is 12.3466 years and the mean value of women's education is 11.9866 years. Use this information to calculate the new estimates of β_0 and β_1 when the variable (E) has been removed from the regression.

Output 1 (Regression (1))

	Coefficient	Std.Error	t-value	t-prob
D_F	82.3653	19.28	4.27	0.000
D_M	124.549	19.80	6.29	0.000
E	2.65070	1.535	1.73	0.086
sigma	49.8728	RSS	365632.519	
R^2	0.174338	F(2,147) =	15.52 [0.000]**	
no. of observations	150	no. of parameters	3	
mean(wages per hour)	135.707	var(wages per hour)	2952.24	

Output 2 (Regression (2))

	Coefficient	Std.Error	t-value	t-prob
Constant	124.549	19.80	6.29	0.000
D_F	-42.1839	8.163	-5.17	0.000
E	2.65070	1.535	1.73	0.086
sigma	49.8728	RSS	365632.519	
R^2	0.174338	F(2,147) =	15.52 [0.000]**	

no. of observations 150 no. of parameters 3
 mean(wages per hour) 135.707 var(wages per hour) 2952.24

Output 3 (Regression (3))

	Coefficient	Std.Error	t-value	t-prob
Constant	129.266	25.03	5.16	0.000
D_F	-54.0309	39.13	-1.38	0.169
E	2.26865	1.973	1.15	0.252
$(D_F E)$	0.976874	3.155	0.310	0.757

sigma 50.0269 RSS 365392.546
 R² 0.17488 F(3,146) = 10.31 [0.000]**

no. of observations 150 no. of parameters 4
 mean(wages per hour) 135.707 var(wages per hour) 2952.24

Output 4 (Regression (4))

	Coefficient	Std.Error	t-value	t-prob
Constant	96.9259	20.66	4.69	0.000
E	3.18752	1.659	1.92	0.057

sigma 54.0306 RSS 432057.218
 R² 0.0243395 F(1,148) = 3.692 [0.057]

no. of observations 150 no. of parameters 2
 mean(wages per hour) 135.707 var(wages per hour) 2952.24