

Ekonometri (3606)

Tentamen 18.3.2011

Skriftid: 4h

Hjälpmaterial: Kalkylator, formel- och tabellsamling

Examinator: Gunnar Rosenqvist

Provtexten får bortföras! Motivera dina svar! Lycka till!

1. Antag att $\hat{\theta}$ är en estimator av en (populations-)parameter θ . Värdet på estimatorn räknas alltså ut på basen av observationerna i ett stickprov och varierar från stickprov till stickprov.
 - (a) Vad avses med att $\hat{\theta}$ är en väntevärdesriktig estimator av θ .
 - (b) Vad avses med en estimators effektivitet.
 - (c) Antag att vi har oberoende observationer Y_1, Y_2, \dots, Y_n från en population med väntevärde μ och varians σ^2 . Visa att stickprovsmedeltalet är en väntevärdesriktig estimator av μ och beräkna dess varians.

(6 poäng)

2. Nedan anges en modell för efterfrågan på cigaretter. De variabler som används är

<i>cigs</i>	= antalet rökta cigaretter per dag
<i>income</i>	= årlig inkomst i USD
<i>cigprice</i>	= priset i cigarettpaket i cent
<i>educ</i>	= antal år av utbildning
<i>age</i>	= åldern i år
<i>restaurn</i>	= 1 om personen i fråga bor i en stat med förbud att röka i restaurang, 0 annars

Eftersom en majoritet av människorna inte röker har den beroende variabeln *cigs* ett stort antal nollar. En linär modell är därför inte idealisk eftersom den kan generera negativa anpassade värden. Icke desto mindre kan vi kanske dra några lärdomar även med en linär modell. Alltså anpassas följande teoretiska modell:

$$cigs = \beta_0 + \beta_1 \log(income) + \beta_2 \log(cigprice) + \beta_3 educ + \beta_4 age + \beta_5 age^2 + \beta_6 restaurn + u$$

Modellen estimeras med hjälp av data från USA. Materialet består av 807 observationer. Besvara följande frågor utgående från utskrift 1 i bilagan. Notera att variablerna *income* och *cigprice* används i logaritmerad form och *age* förekommer såväl i linär som i kvadratisk form.

- a) Tolka skattningarna av var och en av vinkelkoefficienterna.
- b) Testa på 5% nivån med ensidigt test signifikansen skilt för sig för var och en av variablerna $\log(income)$, $\log(cigprice)$ och *educ*.
- c) Vad säger i detta fall R^2 och justerade R^2 ? Varför finns det anledning att använda justerade R^2 ?

d) Testa hela modellens signifikans.

(6 poäng)

3. (i) Vad avses med homoskedasticitet och heteroskedasticitet?
(ii) Verkar vi ha homoskedasticitet eller heteroskedasticitet i modellen i uppgift 2?
(iii) Vid fall av heteroskedasticitet, hur inverkar det på resultaten av OLS-regressionen?
(iv) I utskrift 2 har man regresserat de kvadrerade residualerna från utskrift 1 mot modellens förklarande variabler. Genomför ett Breusch-Pagan test av heteroskedasticitet.
(v) Vilka åtgärder kan vidtas ifall heteroskedasticitet föreligger?

(6 poäng)

4. Utgå från den multipla modellen

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + u.$$

Då modellen estimeras vill man ofta ha så liten osäkerhet som möjligt gällande de estimerade parametrarna. Med andra ord bör estimatorernas varianser, $\text{Var}(\hat{\beta}_j)$, vara små. Givet vissa grundläggande antaganden får variansen som

$$\text{Var}(\hat{\beta}_j) = \frac{\sigma^2}{SST_j(1-R_j^2)} \quad \text{för } j = 1, 2, \dots, k$$

där σ^2 = feltermens varians
 $SST_j = \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2$
 R_j^2 = R^2 från x_j regisserad på de övriga oberoende variablerna.

- a) Bör σ^2 vara stort eller litet för att få en liten varians för estimatoren? Vad kan man beakta då man planerar sin undersökning och då man formulerar sin modell för att påverka σ^2 i "rätt" riktning?
b) Bör SST_j vara stort eller litet för att få en liten varians för estimatoren? Vad kan man beakta då man planerar sin undersökning och då man formulerar sin modell för att påverka SST_j i "rätt" riktning?
c) Bör R_j^2 vara stort eller litet för att få en liten varians för estimatoren? Vad kan man beakta då man planerar sin undersökning och då man formulerar sin modell för att påverka R_j^2 i "rätt" riktning?

(6 poäng)

5. Förklara kortfattat i ord vad följande tidsserie-ekonometriska begrepp betyder och ge ett exempel på var och ett av dem
a) statisk modell, b) autokorrelation, c) tidstrend. (6 poäng)

BILAGA 1

Utskrift 1

smoke.xls loaded from C:\Documents and Settings\rosenqvi\My Documents\Ekonometri\Ekonometri 2011\swbook1_3e\excelfiles\smoke.xls

Ox Professional version 6.00 (Windows/U) (C) J.A. Doornik, 1994-2009

---- PgGive 13.0 session started at 18:32:19 on 16-03-2011 ----

EQ(1) Modelling cigs by OLS-CS

The dataset is: C:\Documents and Settings\rosenqvi\My Documents\Ekonometri\Ekonometri 2011\swbook1_3e\excelfiles\smoke.xls

The estimation sample is: 1 - 807

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-3.63987	24.08	-0.151	0.8799	0.0000
Lincome	0.880269	0.7278	1.21	0.2268	0.0018
Lcigprice	-0.750854	5.773	-0.130	0.8966	0.0000
Educ	-0.501498	0.1671	-3.00	0.0028	0.0111
Age	0.770694	0.1601	4.81	0.0000	0.0281
Agesq	-0.00902280	0.001743	-5.18	0.0000	0.0324
Restaurn	-2.82508	1.112	-2.54	0.0112	0.0080

sigma	13.4048	RSS	143750.658
R^2	0.0527369	F(6,800) =	7.423 [0.000]**
Adj.R^2	0.0456325	log-likelihood	-3236.23
no. of observations	807	no. of parameters	7
mean(cigs)	8.68649	se(cigs)	13.7215

Normality test: Chi^2(2) = 623.67 [0.0000]**

Hetero test: F(10,796) = 3.7326 [0.0001]**

Hetero-X test: F(20,786) = 2.4787 [0.0003]**

RESET23 test: F(2,798) = 1.9343 [0.1452]

Utskrift 2

Algebra code for smoke.xls:
residualssq = residuals^2;

EQ(2) Modelling residualssq by OLS-CS

The dataset is: C:\Documents and Settings\rosenqvi\My
Documents\Ekonometri\Ekonometri 2011\swbook1_3e\excelfiles\smoke.xls
The estimation sample is: 1 - 807

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-636.304	652.5	-0.975	0.3298	0.0012
Lincome	24.6385	19.72	1.25	0.2119	0.0019
Leigprice	60.9768	156.4	0.390	0.6968	0.0002
Educ	-2.38423	4.528	-0.527	0.5986	0.0003
Age	19.4175	4.339	4.48	0.0000	0.0244
Agesq	-0.214790	0.04723	-4.55	0.0000	0.0252
Restaurn	-71.1814	30.13	-2.36	0.0184	0.0069
sigma	363.249		RSS	105559908	
R^2	0.0399733		F(6,800) =	5.552 [0.000]**	
Adj.R^2	0.0327731		log-likelihood	-5898.9	
no. of observations	807		no. of parameters	7	
mean(Y)	178.13		se(Y)	369.352	

Normality test: Chi^2(2) = 6405.4 [0.0000]**

Hetero test: F(10,796) = 0.95985 [0.4773],

Hetero-X test: F(20,786) = 0.75127 [0.7733]

RESET23 test: F(2,798) = 2.3930 [0.0920]