Department of Statistics

Stockholms universitet

Correction sheet

Date: 28/11/2017

Room: Brunnsvikssalen

Course: Econometrics (eng)

Exam: Econometrics I (eng)

Anonymous code: EK1-HEC-XK2

☒ I authorise the anonymous posting of my exam, in whole or in part, on the department homepage as a sample student answer.

NOTE! ALSO WRITE ON THE BACK OF THE ANSWER SHEET

Mark answered questions

1	2	3	4	5	6	7	8	9	Total number of pages
X	X	X	X	X					5
Teacher's notes 22	32	6	12	26					

Points	Grade	Teacher's sign
98	A	BK

SU, DEPARTMENT OF STATISTICS

Room: Brunnsviksalen Anonymous code: EKHIEC-XKZ Sheet number: 1

1. $Y_t = \beta t + u_t$, $t = 1,2,\dots,T$ $\quad b = \frac{\sum_{t=1}^{T} Y_t}{\sum_{t=1}^{T} t} = \frac{\bar{Y}}{\bar{t}}$ $\quad b^* = \frac{Y_T}{T}$

$E(b) = \beta \quad V(b) = \sigma_b^2 = \frac{\sigma^2}{T\bar{t}^2}$

A. $E(b^*) = E\left(\frac{Y_T}{T}\right) = \frac{1}{T} E(Y_T) = \frac{1}{T} E\left[\beta T + u_T\right] = \frac{1}{T} E(\beta T) + E(u_T) =$ ($E(u_T) = 0$)

$= \frac{1}{T} \beta T = \beta$

(6)

Eftersom $E(b^*) = \beta$ så är den unbiased.

B. $V(b^*) = V\left(\frac{Y_T}{T}\right) = V\left(\frac{\beta T + u_T}{T}\right) = \frac{1}{T^2} V\left(\beta T + u_T\right) = \frac{V(\beta T) + V(u_t)}{T^2} =$

$= \frac{V(u_t)}{T^2} = \frac{\sigma^2}{T^2}$ svar: $V(b^*) = \frac{\sigma^2}{T^2}$

(5)

C. b är effektivare, eftersom den har mindre varians (utom då T=1, då de har samma varians):

$T=1$ ger: $V(b) = \frac{\sigma^2}{1 \cdot \frac{2^2}{4}} = \sigma^2 \quad V(b^*) = \frac{\sigma^2}{1^2} = \sigma^2$

(5)

Då $T > 1$ är b mer effektiv. till exempel:

$T = 5$ ger: $V(b) = \frac{\sigma^2}{5 \cdot \frac{6^2}{4}} = \frac{\sigma^2}{45} \quad V(b^*) = \frac{\sigma^2}{5^2} = \frac{\sigma^2}{25} > \frac{\sigma^2}{45} = V(b)$.

Eftersom $V(b) < V(b^*)$ är b en mer effektiv estimator.

D. Vi antar att $V(u_t) = \sigma^2 = t \cdot \sigma_k^2$ där σ_k^2 är en konstant.

För att uppfylla antagandet om konstant varians i feltermen dividerar vi modellen med $\sqrt{t}$:

$\frac{Y_t}{\sqrt{t}} = \frac{\beta t}{\sqrt{t}} + \frac{u_t}{\sqrt{t}} = \beta \sqrt{t} + \frac{u_t}{\sqrt{t}}$

(6)

Feltermen har nu variansen: $V\left(\frac{u_t}{\sqrt{t}}\right) = \frac{1}{t} V(u_t) = \frac{\sigma^2}{t} = \frac{t \cdot \sigma_k^2}{t} = \sigma_k^2$

D.v.s en konstant. Denna transformerade modell uppfylles således antagandet.

22

SU, DEPARTMENT OF STATISTICS

Room: Brunnsvikssalen Anonymous code: EKI-HEC-XKZ Sheet number: 2

2 Model: $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$

A. Vi genomför ett F-test med hypoteserna:

$H_0: \beta_{11} = \beta_{12}, \beta_{21} = \beta_{22}, \beta_{31} = \beta_{32}$

H_1: Åtminstone någon skiljer sig.

Teststatistika: $F = \frac{(RSS_R - RSS_{UR})/m}{RSS_{UR}/(n-k)} \sim F(m; n-k)$ om H_0 är sann. ($m=3$, $n-k=37$)

$RSS_{UR} = RSS_1 + RSS_2 = 15000 + 12000 = 27000$ $\alpha = 0{,}05$

$RSS_R = RSS_{1+2} = 35000$ $m = 3$ $n-k = 40-3 = 37$

Beslutsregel: Om $F_{obs} > F_{0,05}(3, 37) \approx F_{0,05}(3;40) = 2{,}84$ så ska H_0 förkastas.
$\approx F_{0,05}(3;30) = 2{,}92$ (−2)

Resultat: $F_{obs} = \frac{(35000-27000)/3}{27000/37} = 3{,}65 > F_{0,05}(3;37)$

H_0 förkastas på signifikansnivån 5%.

(10)

Slutsats: Det finns stöd för att parametrarna skiljer sig åt mellan de två perioderna.

B. Vi genomför Goldfeld-Quandts test av lika varians:

$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 \neq \sigma_2^2$. $\alpha = 0{,}10$

Statistika: $F = \frac{S_1^2}{S_2^2} \sim F(n_1 - k_1; n_2 - k_2)$ om H_0 är sann. ($=17$, $=17$)

$n_1 = n_2 = 20$ $k_1 = k_2 = 3$

$S_2^2 = \frac{RSS_2}{n_2 - k_2} = \frac{12000}{17} = 705{,}88235$ $S_1^2 = \frac{RSS_1}{n_1 - k_1} = \frac{15000}{17} = 882{,}35294$

Beslutsregel: H_0 förkastas om $F_{obs} > F_{0,05}(17;17)$ $\approx F_{0,05}(15;17) = 2{,}31$ $\approx F_{0,05}(20;17) = 2{,}23$

Resultat: $F_{obs} = \frac{882{,}35294}{705{,}88235} \approx 1{,}25 < F_{0,05}(17;17)$

H_0 förkastas ej. Det finns ej stöd för olika varianser.

Slutsats: Antagandet om lika varians verkar vara uppfyllt.

Vänd.

R

C. Denna modell kommer att ha samma TSS som $TSS_{1+2} = 109000$ eftersom alla obs är med. Men den kommer att ha $RSS = RSS_1 + RSS_2$ eftersom Den genom D_t kommer att bilda olika ekvationer för period 1 resp. 2.

$RSS = RSS_1 + RSS_2 = 27000$. $k = 6$ $n = 40$

$ESS = TSS - RSS = 81000$.

Vi genomför ett F-test på hela modellen:

$H_0: \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ $\alpha = 0{,}05$.

H_1: Så är ej fallet.

Statistika: $F = \frac{ESS/5}{RSS/34} \sim F(5;34)$ om H_0 är sann.

Beslutsregel: H_0 förkastas om $F_{obs} > F_{0,05}(5;34)$ $\approx F_{0,05}(5;30) = 2{,}53$ $\approx F_{0,05}(5;40) = 2{,}45$.

Resultat: $F_{obs} = \frac{81000/5}{27000/34} = 20{,}4 > F_{0,05}(5;34)$.

H_0 förkastas på 5% signifikansnivå.

Slutsats: Denna modell förklarar åtminstone en del av variationen i Y_t.

32

SU, DEPARTMENT OF STATISTICS

Room: Brunnsvikssalen Anonymous code: EK1-HEC-XKZ Sheet number: 3

3. Det falska påståendet är b) $\Sigma u_i = 0$ (alltid).
Delta eftersom vi inte kan veta u_i exakt (avståndet till populationens regressionslinje), eftersom vi inte vet ekvationen för sambandet i populationen. Vi kan bara skatta det (därav $\Sigma \hat{u}_i = 0$).

Svar: b)

6

SU, DEPARTMENT OF STATISTICS

Room: Brännvinkssalen Anonymous code: EKI-HEC-XKZ Sheet number: 4.

4.

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -2{,}40 + 0{,}30X \iff \hat{p} = \frac{1}{1+e^{2{,}40-0{,}30X}}$$

A. b) verkar mest sannolikt. Vi utforskar därför det först.

För att reda ut hur mycket ODDS:et ökar för varje dag tar vi ODDS-kvoten för X=A och X=A+1, d.v.s: Hur förändras ODDS:et då det går en dag och X ökar med 1?

$$\frac{\left(\frac{\hat{p}}{1-\hat{p}}\right)\Big|X=A+1}{\left(\frac{\hat{p}}{1-\hat{p}}\right)\Big|X=A} = \frac{e^{-2{,}40+0{,}30(A+1)}}{e^{-2{,}40+0{,}30A}} = e^{-2{,}40+0{,}30A+0{,}30+2{,}40-0{,}30A} =$$

$$= e^{0{,}30} = 1{,}35$$

ODDS:et blir alltså 35% större om X ökar med 1.

Slutsats: ODDS:et för att ha hört talas om varumärket ökar med c.a 35% för varje dag. b) är sant.

Svar: b).

(6)

B. $\hat{p} = 0{,}5 \rightarrow \ln\left(\frac{0{,}5}{1-0{,}5}\right) = \ln 1 = 0 = -2{,}40 + 0{,}30X \rightarrow$

$\rightarrow 2{,}40 = 0{,}3X \rightarrow X = 2{,}40/0{,}3 = 8$

Svar: Efter 8 dagar är den skattade sannolikheten för en person att ha hört talas om varumärket 50%.

(6)

12

5. A. Om det inte finns någon säsongsvariation så är $\beta_3=\beta_4=\beta_5=0$ i modell 2. För att se om så är fallet gör vi ett F-test där vi jämför den enkla modell 1 med den utökade modell 2.

$H_0: \beta_3=\beta_4=\beta_5=0$ H_1: så är ej fallet. $\alpha=0{,}05$

Teststatistika: $F=\frac{(RSS_R-RSS_{UR})/m}{RSS_{UR}/(n-k)} \sim F(m;n-k)$ om H_0 är sann.

$RSS_R = RSS_1 = 4110{,}19301$ $RSS_{UR} = RSS_2 = 1302{,}75351$

$m=3$ $n-k=48-5=43$

Beslutsregel: H_0 förkastas om $F_{obs} > F_{0,05}(3;43)$ $\approx F_{0,05}(3;40)=2{,}84$ $\approx F_{0,05}(3;60)=2{,}76$

Resultat: $F_{obs}=\frac{(4110{,}19301-1302{,}75351)/3}{1302{,}75351/43}=30{,}894 > F_{0,05}(3;43)$

H_0 förkastas på 5% sign. nivå.

R

Slutsats: Det finns stöd för säsongsvariation.

b) Modellen i Durbin-Watson-testet är:

$u_t=\rho u_{t-1}+e_t$, där ρ är korrelationen mellan u_t och u_{t-1}. H_0 är: $\rho=0$. Statistikan $d=\frac{\sum_{t=2}^{n}(\hat{u}_t-\hat{u}_{t-1})^2}{\sum_{t=1}^{n}\hat{u}_t^2}$ går mellan 0 och 4 där värden nära 0 indikerar positiv autokorrelation och värden nära 4 negativ.
D-W testar för första ordningens autokorrelation.

Modellen i Breusch-Godfrey-testet är:

$u_t=\rho_1 u_{t-1}+\rho_2 u_{t-2}+\rho_3 u_{t-3}+ \ldots +\rho_k u_{t-k}$

H_0 är: $\rho_1=\rho_2=\rho_3=\ldots=\rho_k=0$ och testas med statistikan nR^2 (från regressionen av $\hat{u}_t$ på regressorerna som givit residualerna och $\hat{u}_{t-1}, \hat{u}_{t-2}, \ldots, \hat{u}_{t-k}$) är approx $\chi^2(k)$-fördelad.
B-G testar för autokorrelation upp till ordning k.

I detta fall finns inget stöd för första ordningens autokorrelation enl. D-W ($d_L=1{,}585$ resp. $d_U=1{,}721$). Däremot visar B-G att det finns stöd ($\alpha=0{,}05$) för autokorrelation vid test upp till femte ordningen. (p-value $=0{,}0347$).

/26

Stockholm University

Department of Statistics

Econometrics I

WRITTEN EXAMINATION

Tuesday November 28 2017, 10 am – 15 pm

Tools allowed: Pocket calculator

Passing rate. 50% of overall total, which 100 points. For detailed grading

Criteria, see the course description.

The exam will be handed back: not decided

For the maximum number of points on each problem detailed and clear solutions are required.

If not indicated otherwise, the disturbance tem u_i in the models are assumed to fulfill the usual requirements of normality, homoscedasticity and independence.

Task 1 (22 points)

Assume the model $Y_t = \beta t + u_t$, where t= 1, 2, 3,...,T and that you want to estimate β. You are suggested the following estimators:

$$b = \frac{\sum_1^T Y_t}{\sum_1^T t} = \frac{\bar{Y}}{\bar{t}} \quad \text{and} \quad b^* = \frac{Y_T}{T} \ .$$

The estimator b is known to be unbiased with variance: $\sigma_b^2 = \frac{\sigma^2}{T\bar{t}^2} = \frac{\sigma^2}{T\frac{(T+1)^2}{4}}$.

A. Show that the estimator b^* also is unbiased.

B. Derive the variance of b^* assuming constant error variance (σ^2)

C. Which estimator of the two is the most efficient one? Demonstrate for T=5.

D. The estimator b is the GLS-estimator when the error variance σ^2 is proportional to t. Present under this assumption the transformed model for which the assumption of homoscedastic error term is fulfilled. This task is independent of task A-C.

Task 2. (34 points)

From a regression analysis we collect the following information.

Model: $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$

	Period 1	Period 2	Period 1+2
TSS	46000	42000	108000
RSS	15000	12000	35000
N	20	20	40

The estimation results for the period 1 and period 2 are based on two independent regressions. The results from period 1+2 are obtained using all 40 observations.

A. Are the parameters of the model the same for the two periods? Perform a formal test. Your solution should include null- and alternative hypothesis in terms of the beta parameters, test statistic, its distribution with specification of degrees of freedom, decision rule, results and a conclusion.

B. Is the assumption of equal disturbance variance fulfilled in task A? Perform a formal test. Your solution should include null- and alternative hypothesis, test statistic, its distribution with specification of the degrees of freedom, decision rule, results and a conclusion.

C. Suppose we estimate the Model:
$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 D_t + \beta_5 X_{2t} D_t + \beta_6 X_{3t} D_t + u_t$
on the same data, where $D_t = 0$ for period 1 and equal to 1 for period 2. Make a formal test of whether this model explain at least some of the variation in Y_t.

The tasks A, B and C can be solved independently of each other.

Task 3 (6 points)

Which of the following statements is not correct for a simple linear regression model with an intercept:

a) $\sum \hat{u}_i =$ is always equal to 0

b) $\sum u_i =$is always equal to 0

c) $\sum \hat{Y}_i = \sum Y_i$ (always)

d) $\sum X_i \hat{u}_i =$ is always equal to 0.

e) $\sum \hat{Y}_i \hat{u}_i =$is always equal to 0.

Task 4 (12 points)

A new brand is introduced on a market. This is followed up by an advertising campaign. During a period after the introduction the market penetration is measured in the target group. Let Y=1 if a person have heard of the brand and let Y=0 if not. Let X be the number of days after the introduction. On the basis of collected data a logistic regression were performed with the following result:

$$\ln\left(\frac{\hat{P}}{1-\hat{P}}\right) = -2{,}40 + 0{,}30X \qquad P = P(Y=1)$$

A. Which of the following statements is correct?

a) The ODDS that a person has heard of the brand is estimated to increase with 0.30 units for each additional day since the introduction.

b) The ODDS that a person has heard of the brand is estimated to increase with 35% for each additional day since the introduction.

c) The probability that a person has heard of the brand is estimated to increase with 0.30 units for each additional day since the introduction.

d) The probability that a person has heard of the brand increases with 0.30% for each additional day since the introduction.

e) The probability that a person has heard of the brand is estimated to increase with 35% for each additional day since the introduction.

B. According to the estimated model, after how many days is the probability that a person has heard of the brand 50%?

Task 5 (26 points)

Given: The quarterly data below and the regression outputs on the next page

time	D1t	D2t	D3t	D4t	Yt
1	1	0	0	0	110.2
2	0	1	0	0	134.9
3	0	0	1	0	128.4
4	0	0	0	1	117.7
5	1	0	0	0	117.5
6	0	1	0	0	137.9
7	0	0	1	0	144.4
8	0	0	0	1	126.9
9	1	0	0	0	115.7
10	0	1	0	0	142.5
11	0	0	1	0	143.8
12	0	0	0	1	125.1
13	1	0	0	0	138.6
14	0	1	0	0	135.3
15	0	0	1	0	149.9
16	0	0	0	1	138.2
17	1	0	0	0	143
18	0	1	0	0	150.8
19	0	0	1	0	153.7
20	0	0	0	1	147.6
21	1	0	0	0	145
22	0	1	0	0	161.4
23	0	0	1	0	152.1
24	0	0	0	1	142.8
25	1	0	0	0	144.9
26	0	1	0	0	163.4
27	0	0	1	0	167.7
28	0	0	0	1	160.4
29	1	0	0	0	157.9
30	0	1	0	0	164.6
31	0	0	1	0	167.2
32	0	0	0	1	154.2
33	1	0	0	0	156.7
34	0	1	0	0	174
35	0	0	1	0	190.5
36	0	0	0	1	169.1
37	1	0	0	0	151.8
38	0	1	0	0	174.6
39	0	0	1	0	179.1
40	0	0	0	1	166.1
41	1	0	0	0	170.4
42	0	1	0	0	192.5
43	0	0	1	0	187.3
44	0	0	0	1	170.3
45	1	0	0	0	171.2
46	0	1	0	0	194.2
47	0	0	1	0	198.8
48	0	0	0	1	184

A. Perform a formal test of whether there is any seasonal variation in Y_t.

B. In connection with the estimation results you find information about potential autocorrelation. Present the models for u_t for the Durbin Watson test and the Breusch Godfrey test and the corresponding null hypothesis (in terms of the parameters of the models) for the two test. Comment shortly (at most two sentences per test) on the results of the tests.

. regress Yt time

```
      Source |       SS          df       MS          Number of obs  =       48
-------------+------------------------------          F(1, 46)       =   203.71
       Model |  18202.1186        1   18202.1186      Prob > F       =   0.0000
    Residual |  4110.19301       46   89.3520219      R-squared      =   0.8158
-------------+------------------------------          Adj R-squared  =   0.8118
       Total |  22312.3116       47   474.730035      Root MSE       =   9.4526

------------------------------------------------------------------------------
          Yt |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   1.405672   .0984862    14.27   0.000      1.20743    1.603914
       _cons |   120.0256   2.771938    43.30   0.000      114.446    125.6052
------------------------------------------------------------------------------
Durbin-Watson d-statistic =  1.950222.
```

regress Yt time D2t D3t D4t

```
      Source |       SS          df       MS          Number of obs  =       48
-------------+------------------------------          F(4, 43)       =   173.37
       Model |  21009.5581        4   5252.38953      Prob > F       =   0.0000
    Residual |  1302.75351       43   30.2965932      R-squared      =   0.9416
-------------+------------------------------          Adj R-squared  =   0.9362
       Total |  22312.3116       47   474.730035      Root MSE       =   5.5042

------------------------------------------------------------------------------
          Yt |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   1.399847   .0575359    24.33   0.000     1.283815    1.515879
         D2t |   15.53349   2.247831     6.91   0.000      11.0003    20.06667
         D3t |   17.20031   2.250039     7.64   0.000     12.66267    21.73794
         D4t |   2.425461   2.253714     1.08   0.288    -2.119586    6.970508
       _cons |   111.3785   2.067827    53.86   0.000     107.2083    115.5487
------------------------------------------------------------------------------

Durbin-Watson d-statistic =  1.865191

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
      K=5    |        12.007               5                   0.0347
---------------------------------------------------------------------------
                        H0: no serial correlation
```

Prel. Solutions to exam in Econometrics I November 2017.

Task 1

A) $b^* = \frac{Y_T}{T} = \frac{\beta T + u_T}{T} = \beta + \frac{u_T}{T} \quad E(b^*) = \beta + \frac{E(u_T)=0}{T} = \beta$ Hence unbiased

B) $V(b^*) = V\left(\beta + \frac{u_T}{T}\right) = V\left(\frac{u_T}{T}\right) = \frac{\sigma^2}{T^2}$

C) $\sigma_b^2 = \frac{\sigma^2}{T\frac{(T+1)^2}{4}} = \frac{\sigma^2}{5\frac{(5+1)^2}{4}} = \frac{\sigma^2}{45} \qquad V(b^*) = \frac{\sigma^2}{5^2} = \frac{\sigma^2}{25}$ Of the two estimators b has the smallest variance and is therefore more efficient than b*.

D) $\frac{Y_t}{\sqrt{t}_t} = \beta\sqrt{t} + \frac{u_t}{\sqrt{t}} \qquad (V\left(\frac{u_t}{\sqrt{t}}\right) = \frac{\sigma^2 t}{t} = \sigma^2)$

Task 2

A) Model period 1: $Y_t = \beta_{11} + \beta_{21}X_{2t} + \beta_{31}X_{3t} + u_t$

Model period 2: $Y_t = \beta_{12} + \beta_{22}X_{2t} + \beta_{32}X_{3t} + u_t$

H_0: $\beta_{11} = \beta_{12}$, $\beta_{21} = \beta_{22}$, $\beta_{31} = \beta_{32}$
H_1: At least one of the restrictions in H_0 is false   Significance level: 5%

Test statistic: $\frac{(RSS_R - RSS_{UR})/k}{RSS_{UR}/(n-2k)}$ F-distributed with k=3 and 40-2*3=34 d.f. given u IN(0,σ^2) for both periods.,

Decision rule: Reject H_0 if $F_{obs} > F_{0.05,3,34}$ = ?? $F_{3,30}$=2,92 $F_{3,40}$=2,84

F_{obs}= (35000-27000)/3/(27000/(40-6)) = 3,36

The result is significant. All parameters do not seen to be the same for the two periods.

B) Denote the two periods error variances by $\sigma_1^2$and σ_2^2.
H_0: $\sigma_1^2 = \sigma_2^2$ H_1: $\sigma_1^2 \neq \sigma_2^2$ Significance level: 10%
Test statistic: $\frac{s_1^2}{s_2^2}$ F-distributed with n_1-k=17 and n_2-k=17 d.f. given u IN(0,σ^2)
Since we have put the largest residual variance in the numerator we only need to study the rejection region to the right.
Reject H_0 if F_{obs} is larger than $F_{0.05,17,17}$=?? $F_{15,17}$=2,31 $F_{20,17}$=2,23

F_{obs}= 15000*17/(12000*17)=1,25 non sign .
We don't get any support for unequal error variance on 10% significance level.

C. Model: $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 D_t + \beta_5 X_{2t} D_t + \beta_6 X_{3t} D_t + u_t$

H_0: $\beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ H_1: At least one of β_2, β_4 to β_6 differs from zero

Significance level: say 5%

Test statistic: $\frac{ESS/(k-1)}{RSS/(n-k)}$ that is F-distributed with k-1=5 and n-k=34 df if H_0 is true and u_t IN(0, σ^2).

Decision rule: Reject H_0 if $F_{obs} > F_{0,05,\,5,\,34} =$?? $F_{5,30} = 2,53$ $F_{5,40} = 2,45$.

ESS=TSS-RSS=108000-15000-12000=81000, RSS=15000+12000=27000

F_{obs}=81000*34/(27000*5)=20,4. The result is significant and strongly indicates that the model has at least some explanatory power.

Task 3 Answer: Alternative b

Task 4 A) Alternative b

B) In order for P(Y=1) must $e^{-(-2,4+0,30X)} = 1$, that is -2,4+0,3X=0, which givesr X = 8.

Task 5

A: Model: $Y_t = \beta_1 + \beta_2 time + \beta_3 D2_t + \beta_4 D3_t + \beta_5 D4_t + u_t$

H_0: $\beta_3 = \beta_4 = \beta_5 = 0$ H_1: At least one of β_3, β_4 and β_4 differs from zerol

Signifikanslevel: say 5%

Test statistic: $F=((RSS(R) - RSS_{UR})/m)/s^2_{UR}$

F-distributed with m=3 och 48-5=43 d.f. given u IN($0,\sigma^2$),

Reject H_0 if $F_{obs} > F_{0,05,\,3,43}$ = ?? $F_{3,40}=2,84$

Result: $F_{obs} = \frac{(4110.19301-1302.75351)/3}{5,5042^2} = 30,9$ Significant result. We can be pretty sure of existence of seasonality in the data.

B) Durbin Watson test:

$$u_t = \rho u_{t-1} + \varepsilon_t$$

H_0: $\rho=0$ H_1: $\rho>0$ If we test for positive autocorrelation of first order.

Reject H_0 if $d<d_L$ do not reject if $d>d_U$

Breusch Godfreys test:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \rho_3 u_{t-3} + \rho_4 u_{t-4} + \rho_5 u_{t-5} + \varepsilon_t$$

H_0: $\rho_1=\rho_2=\rho_3=\rho_4=\rho_5=0$ against H_1: At least one of $\rho_k \neq 0$, k=1, 2,...,5.

Comments on the d-test:
The value on the Durbin Watson test statistic is close to 2 (1,95 and 1,87) and hence is clearly non-significant ($d_{U,k'=1}$=1,6; $d_{U,k'=4}$=1,72. We have not detected any autocorrelation of order 1 in the two models.

Comments on the Breusch Godfreys test:
The p-value of the test is equal to 0.0347, which means significant results on the 5% level. This suggests the existence of higher order autocorrelation up to order 5 and that the model 2 can be improved.