Stockholms universitet
Statistiska institutionen
Raul Cano

SKRIVNINGSDATUM: 10-06-2019

Skriftlig tentamen i **Regressionsanalys och tidsserieanalys** (4,5 hp), ingående som moment 1 i kursen **Regressionsanalys och undersökningsmetodik, 15 hp.**

Skrivtid: 5 timmar
Hjälpmedel: Miniräknare utan lagrade formler eller lagrad text. Vidhäftade formel- och tabellblad (obs! vidhäftas endast de tabellsidor som behövs för den här tentamen).
Återlämning av tentamen: hämtas på studentexpeditionen, plan 7 i B-huset fr.o.m. tisdagen den 18 juni. Kolla på vår hemsida studentexpeditionens mottagningstider under terminstid.

Tentamen består av fyra uppgifter som kan ge totalt 100 poäng. För betyget A gäller 90-100 p., för betyget B gäller 80-89 p., för betyget C gäller 70-79 p., för betyget D gäller 60-69 p., för betyget E gäller 50-59 p., för betyget Fx gäller 40-49 p. och för betyget F gäller 0-39 p. För detaljerade betygskriterier se kursbeskrivningen på kurshemsidan.

För full poäng på en uppgift krävs fullständiga och väl motiverade lösningar.

Obs! Använd minst 5 decimaltal i dina beräkningar. Om du avrundar t.ex. 5,43689 till 5,4 i dina beräkningar då blir det poängavdrag!
Förtydligande: Om du får t.ex. 0,000003645 då ska du använda 0,000003645 i dina beräkningar och inte avrunda till 0,00000 eller 0,000003.

Uppgift 1: (30 poäng)
I Dalaköping undersöker en fastighetsmäklare sambandet mellan försäljningspris på bostadsrätter i ett visst område (Y) och bostadsyta (X). Data från 5 bostadsrätter har erhållits (se nedan). Y är givet i miljoner kronor och X i kvadratmeter.

y	3	4	5	6	10
x	40	60	70	80	100

a). Anpassa regressionslinjen $\hat{y} = a + b\,x$ till materialet. (10 poäng)
b). Testa H_0: $\beta = 0$ mot H_1: $\beta \neq 0$ (med F-test). Använd signifikansnivå 5% ($\alpha = 0{,}05$). (5 poäng)
Obs! Testa med F-test, inte med t-test, om du testar med t-test då blir det noll poäng.
c). Testa H_0: $\beta = 0$ mot H_1: $\beta < 0$ (med t-test). Använd signifikansnivå 5% ($\alpha = 0{,}05$). (5 poäng)
d). Man har beräknat ett 95%-igt konfidensintervall för det förväntade y-värdet för ett visst x-värde och fått följande intervall (5,785 ; 9,385). Din uppgift är att beräkna det x-värde som man har använt för att skapa det konfidensintervallet. (10 poäng)

Uppgift 2: (20 poäng)
I Dalaköping vill ett företag göra en prognos om försäljning det kommande året. Följande försäljningssiffror (100 000-tals kronor) finns tillgängliga.

År	2014	2015	2016	2017	2018
Försälj.	2	3	5	7	10

a). Anpassa en exponentiell modell (exponentiell trendfunktion) till tidsserien. (10 poäng)

b). Enligt den anpassade modellen, gör en prognos för företagsförsäljning, år 2019. (10 poäng)

Uppgift 3: (20 poäng)
I Dalaköping har man undersökt sambandet mellan variablerna inkomst Y (månadslön i tusentals kronor), ålder X_1 och kön X_2 (dummyvariabel med värdet 1 för kvinnor och 0 för män.). Datamaterialet består av 20 observationer. Du har tillgång till följande information:

```
The regression equation is
Y = 14,64 + 0,97 X1 - 13,71 X2

Predictor   Coef     SE Coef
Constant               8,43
X1                     0,26
X2                     3,87

Analysis of Variance

Source              DF       SS
Regression                1397,64
Residual Error
Total                     2568,80
```

a) Pröva på 5 % signifikansnivå om modellen som helhet är signifikant. Ange de hypoteser som du testar. (10 poäng)

b) Hur mycket större är inkomsten i genomsnitt för en man jämfört med en kvinna i samma ålder, enligt undersökningen? (10 poäng)

Uppgift 4: (30 poäng)
Följande tabell visar förbrukningen av naturgas (i tusentals m^3) i Dalaköping, åren 2015-2018.

År	Kvartal 1	Kvartal 2	Kvartal 3	Kvartal 4
2015	7,5	0,5	1,5	4
2016	6,5	1	2	5,5
2017	9	3	3	7
2018	10	3,5	4,5	7,5

Nedan visas en SAS-utskrift från en regressionsanalys med värden på elförbrukning som beroende (undersöknings-) variabel och tid t (t = 1, 2, 3, …, 16), D1 (kodad 1 om kvartal 1 och 0 annars), D2 (kodad 1 om kvartal 2 och 0 annars) samt D3 (kodad 1 om kvartal 3 och 0 annars) som oberoende (förklarande) variabler.

a). Skatta trend samt säsongkomponenter i en additiv modell där trenden kan antas följa en linjär funktion. (10 poäng)
b). Tolka alla säsongkomponenter (fyra komponenter). (10 poäng)
c). Gör prognoser för andra och tredje kvartalen 2019. (10 poäng)

The REG Procedure
Model: MODEL1
Dependent Variable: Y

Number of Observations Read	16
Number of Observations Used	16

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	124.61250	31.15313	101.16	<.0001
Error	11	3.38750	0.30795		
Corrected Total	15	128.00000			

Root MSE	0.55494	R-Square	0.9735	
Dependent Mean		4.75000	Adj R-Sq	0.9639
Coeff Var	11.68287			

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	3.31250	0.41620	7.96	<.0001
t	1	0.26875	0.03102	8.66	<.0001
D1	1	3.05625	0.40328	7.58	<.0001
D2	1	-3.46250	0.39727	-8.72	<.0001
D3	1	-2.98125	0.39362	-7.57	<.0001

Stockholms universitet
Statistiska institutionen
Regressionsanalys och undersökningsmetodik
Vårterminen 2017
Jörgen Säve-Söderbergh

Formelsamling – regressionsanalys

$$r=\frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2\sum_{i=1}^{n}(y_i-\overline{y})^2}}$$

$$\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})=\sum_{i=1}^{n}x_i y_i-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right)$$

$$\sum_{i=1}^{n}(x_i-\overline{x})^2=\sum_{i=1}^{n}x_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2$$

Enkel lineär regression

$$b_1=\hat{\beta}_1=\frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$$

$$b_0=\hat{\beta}_0=\overline{y}-b_1\overline{x}$$

$$\underbrace{\sum_{i=1}^{n}(y_i-\overline{y})^2}_{\text{SST}}=\underbrace{\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2}_{\text{SSR}}+\underbrace{\sum_{i=1}^{n}e_i^2}_{\text{SSE}}.$$

$$\text{SSE}=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2=\sum_{i=1}^{n}(y_i-\overline{y})^2-b_1^2\sum_{i=1}^{n}(x_i-\overline{x})^2.$$

$$s_e^2=\text{MSE}=\frac{\text{SSE}}{n-2}=\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n-2}$$

Konfidensintervall för β_1 ges av

$$b_1\pm t_{n-2,\frac{\alpha}{2}}s_{b_1},$$

där

$$s_{b_1} = \sqrt{\frac{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n-2}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}}.$$

Prediktionsintervall

$$\underbrace{b_0 + b_1 x_{n+1}}_{\hat{y}_{n+1}} \pm t_{n-2,\frac{\alpha}{2}} \sqrt{s_e^2 \left(1 + \frac{1}{n} + \frac{(x_{n+1}-\overline{x})^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\right)}.$$

Konfidensintervall förväntat y-värde för ett nytt x-värde

$$\underbrace{b_0 + b_1 x_{n+1}}_{\hat{y}_{n+1}} \pm t_{n-2,\frac{\alpha}{2}} \sqrt{s_e^2 \left(\frac{1}{n} + \frac{(x_{n+1}-\overline{x})^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\right)}.$$

Multipel regression

n st observationer och p förklarande variabler.

Variationsorsak	SS	df	MS	F
Regression	SSR	p	$MSR = \frac{SSR}{p}$	MSR/MSE
Residual	SSE	$n-p-1$	$MSE = \frac{SSE}{(n-p-1)}$	
Totalt	SST	$n-1$		

$$R^2_{adj} = 1 - \frac{SSE/(n-p-1)}{SST/(n-1)}$$

Normalekvationerna för fallet $\hat{y} = a + b_1 t + b_2 t^2$

$$\begin{aligned}
\sum_{i=1}^{n} y_i &= a \cdot n + b_1 \sum_{i=1}^{n} t_i + b_2 \sum_{i=1}^{n} t_i^2 \\
\sum_{i=1}^{n} y_i t_i &= a \sum_{i=1}^{n} t_i + b_1 \sum_{i=1}^{n} t_i^2 + b_2 \sum_{i=1}^{n} t_i^3 \\
\sum_{i=1}^{n} y_i t_i^2 &= a \sum_{i=1}^{n} t_i^2 + b_1 \sum_{i=1}^{n} t_i^3 + b_2 \sum_{i=1}^{n} t_i^4
\end{aligned}$$

Säsongsrensning med regression

$$a_0 = \overline{y} - b \cdot \overline{t}$$

$$T_t = a_0 + b \cdot t$$

$$\begin{aligned} S_1 &= a - a_0 + c_1 \\ S_2 &= a - a_0 + c_2 \\ S_3 &= a - a_0 + c_3 \\ S_4 &= a - a_0 \end{aligned}$$

Logistisk regression

$$P\left(Y_i = 1 \mid x_i\right) = \frac{\exp\left(\beta_0 + \beta_1 x_i\right)}{1 + \exp\left(\beta_0 + \beta_1 x_i\right)} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_i)}}$$

$$\log\left(\text{odds}\right) = \beta_0 + \beta_1 x_i.$$

$$\text{odds}\,(D) = \frac{P(D)}{P(D \text{ inträffar inte})} = \frac{P(D)}{1 - P(D)}$$

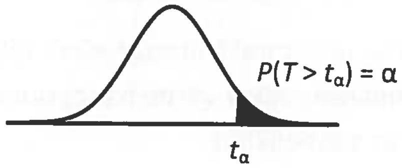
$$P(D) = \frac{\text{odds}\,(D)}{1 + \text{odds}\,(D)}$$

Konfidensintervall för oddskvoten e^{β_1}: $e^{b_1 \pm z \times s_{b_1}}$

TABELL 3. t-fördelningens kvantiler

$T \in t(v)$ där v = antal frihetsgrader.

Vilket värde har t_α om $P(T > t_\alpha) = \alpha$ där α är en given sannolikhet. Utnyttja även $P(T \leq -t_\alpha) = P(T > t_\alpha)$.

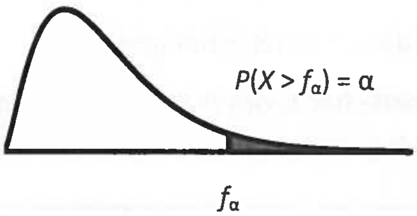


v	α = 0,1	0,05	0,025	0,010	0,005	0,0025	0,0010	0,0005
1	3,078	6,314	12,706	31,821	63,657	127,321	318,309	636,619
2	1,886	2,920	4,303	6,965	9,925	14,089	22,327	31,599
3	1,638	2,353	3,182	4,541	5,841	7,453	10,215	12,924
4	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	1,476	2,015	2,571	3,365	4,032	4,773	5,893	6,869
6	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792
23	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,768
24	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745
25	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725
26	1,315	1,706	2,056	2,479	2,779	3,067	3,435	3,707
27	1,314	1,703	2,052	2,473	2,771	3,057	3,421	3,690
28	1,313	1,701	2,048	2,467	2,763	3,047	3,408	3,674
29	1,311	1,699	2,045	2,462	2,756	3,038	3,396	3,659
30	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646
35	1,306	1,690	2,030	2,438	2,724	2,996	3,340	3,591
40	1,303	1,684	2,021	2,423	2,704	2,971	3,307	3,551
45	1,301	1,679	2,014	2,412	2,690	2,952	3,281	3,520
50	1,299	1,676	2,009	2,403	2,678	2,937	3,261	3,496
55	1,297	1,673	2,004	2,396	2,668	2,925	3,245	3,476
60	1,296	1,671	2,000	2,390	2,660	2,915	3,232	3,460
65	1,295	1,669	1,997	2,385	2,654	2,906	3,220	3,447
70	1,294	1,667	1,994	2,381	2,648	2,899	3,211	3,435
75	1,293	1,665	1,992	2,377	2,643	2,892	3,202	3,425

Forts. nästa sida

TABELL 5. *F*-fördelningens kvantiler

$X \in F(v_1, v_2)$ där v_1, v_2 = antal frihetsgrader i täljaren respektive nämnaren. Vilket värde har f_α om $P(X > f_\alpha) = \alpha$ där α är en given sannolikhet.



α = 0,05

	v_1 = 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
v_2 = 1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,0	243,9	244,7	245,4	245,9
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40	19,41	19,42	19,42	19,43
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,73	8,71	8,70
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91	5,89	5,87	5,86
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68	4,66	4,64	4,62
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,98	3,96	3,94
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57	3,55	3,53	3,51
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28	3,26	3,24	3,22
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07	3,05	3,03	3,01
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,89	2,86	2,85
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79	2,76	2,74	2,72
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69	2,66	2,64	2,62
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60	2,58	2,55	2,53
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57	2,53	2,51	2,48	2,46
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,45	2,42	2,40
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46	2,42	2,40	2,37	2,35
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38	2,35	2,33	2,31
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,31	2,29	2,27
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31	2,28	2,26	2,23
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28	2,25	2,22	2,20
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20	2,16	2,14	2,11	2,09
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,13	2,09	2,06	2,04	2,01
35	4,12	3,27	2,87	2,64	2,49	2,37	2,29	2,22	2,16	2,11	2,07	2,04	2,01	1,99	1,96
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04	2,00	1,97	1,95	1,92
45	4,06	3,20	2,81	2,58	2,42	2,31	2,22	2,15	2,10	2,05	2,01	1,97	1,94	1,92	1,89
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,99	1,95	1,92	1,89	1,87
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92	1,89	1,86	1,84
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02	1,97	1,93	1,89	1,86	1,84	1,81
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,91	1,88	1,84	1,82	1,79
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,89	1,85	1,82	1,79	1,77
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,79	1,75	1,72	1,69	1,67

Forts. nästa sida

9

Stockholms universitet

Statistiska institutionen

Rättningsblad

Datum: 10/6-2019

Sal: Ugglevikssalen

Tenta: Regressions- och tidsserieanalys

Kurs: Regressionsanalys och undersökningsmetodik

ANONYMKOD: 0009-KWE

☒ Jag godkänner att min tenta får läggas ut anonymt på hemsidan som studentsvar.

OBS! SKRIV ÄVEN PÅ BAKSIDAN AV SKRIVBLADEN

Markera besvarade uppgifter med kryss

1	2	3	4	5	6	7	8	9	Antal inl. blad
x	x	x	x	x					5
Lär.ant. 28p	16p	20p	20p						

POÄNG	BETYG	Lärarens sign.
84p	B	RC

SU, STATISTIK

Skrivsal: UG Anonymkod: 0009-KWB Blad nr: 1

1) 28p

1 a)

$$\bar{y} = \frac{3+4+5+6+10}{5} = \frac{28}{5} = 5,6$$

$$\bar{x} = \frac{40+60+70+80+100}{5} = \frac{350}{5} = 70$$

$$\hat{y} = a + bx$$

$$b = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{\sum x_i y_i - \frac{1}{5}(\sum x \sum y)}{\sum x^2 - \frac{1}{5}(\sum x)^2}$$

$$\sum x_i y_i = (3\cdot 40)+(4\cdot 60)+(5\cdot 70)+(6\cdot 80)+(10\cdot 100) = 2190$$

$$\frac{(\sum x)\cdot(\sum y)}{5} = \frac{9800}{5} = 1960$$

$$\sum x^2 - \frac{1}{5}(\sum x)^2 = 26500 - \frac{1}{5}(122500) = 2000$$

$$b = \frac{2190-1960}{2000} = 0,115$$ R

a) 10p

$$a = \bar{y} - b\cdot\bar{x} = 5,6 - 70\cdot 0,115 = (-2,45)$$ R

Svar: $\hat{y} = -2,45 + 0,115x$ R

b)

b) 5p

$H_0: \beta = 0$

$H_1: \beta \neq 0$

signifikant nivå 5%

vänd →

besiutsregel om $f_{obs} > f_{krit}$ förkasta H_0

$f_{krit} = f_{(1,3)}(0{,}05) = 10{,}13$

R

$f_{obs} = \frac{MSR}{MSE}$

$MSR = \frac{SSR}{p}$ $p=1$
$n=5$

$MSE = \frac{SSE}{(n-p-1)}$

$SST = \sum(Y_i - \bar{Y})^2 = \sum Y_i^2 - \frac{1}{5}\left(\sum Y_i\right)^2 = 186 - \frac{1}{5} 784 = 29{,}2$

$SSE = SST - b^2(x_i - \bar{x})^2 = 29{,}2 - 0{,}115^2 \cdot 2000 = 2{,}75$

$SSR = SST - SSE = 29{,}2 - 2{,}75 = 26{,}45$

$SSE = 2{,}75$ $MSE = 0{,}9166667$

$SSR = 26{,}45$ $MSR = 26{,}45$

R

$F_{obs} = \frac{26{,}45}{0{,}916667} = 28{,}8545455 > 10{,}13$

Slutsats man ~~ska~~ förkasta H_0

c) $H_0: \beta = 0$

$H_1: \beta < 0$

Sigignifikansnivå 5%

0) 3p

SU, STATISTIK

Skrivsal: UG **Anonymkod:** 0009-KWE **Blad nr:** 2

c) beslutsregel om $t_{obs} < t_{krit}$ förkasta H_0

$t_{krit} = (-t_{(3)}(0{,}05)) = -2{,}353$ R

$t_{obs} = \frac{b - b_0}{S_b}$ $b_0 = 0$, $b = 0{,}115$

$S_b = \sqrt{\frac{MSE}{\Sigma(x_i - \bar{x})^2}} = \sqrt{\frac{0{,}9166667}{2000}} = 4{,}583333 \cdot 10^{-4}$

STORT FEL (SE LÖSNINGAR

$t_{obs} = \frac{0{,}115}{S_b} = 250{,}9090909$ −2p

F

$250{,}9090909 < -2{,}353$

förkasta ej H_0

D) $\sqrt{MSE \cdot \left(0{,}2 + \frac{(x-70)^2}{2000}\right)} \cdot 3{,}182 =$ $y - 5{,}785 = 1{,}8$

$y = 7{,}585$

$MSE = 0{,}916667$

$1{,}8 = 3{,}182 \cdot \sqrt{MSE \cdot \left(0{,}2 + \frac{(x-70)^2}{2000}\right)} =$

d) 10p

$0{,}5656819 = \sqrt{MSE \left(0{,}2 + \frac{(x-70)^2}{2000}\right)}$

$0{,}565681962^2 = MSE\left(0{,}2 + \frac{(x-70)^2}{2000}\right)$

R

$$= 0{,}3199960822 = MSE\left(0{,}2 + \frac{(x-70)^2}{2000}\right)$$

$$= \frac{0{,}3199960822}{MSE} - 0{,}2 = \frac{(x-70)^2}{2000}$$

$$= (x-70)^2 = \left(\frac{0{,}3199960822}{MSE} - 0{,}2\right) \cdot 2000 =$$

$$(x-70)^2 = 298{,}1732702$$

$$(x-70) = 17{,}26769441$$

$$x = 87{,}26769441$$

R

Svar: I konfident intervallet har man använt 87,26769441 som X_{n+1}.

R

SU, STATISTIK

Skrivsal: UG   Anonymkod: 0009-KWE   Blad nr: 3

2) 16p

2) a)

x	t	t^2	y_{log}	$y_{log} \cdot t$
2014	÷2	4	0,301029	0,602059913
2015	1	1	0,4772	0,477121254 7
2016	0	0	0,6989	0
2017	1	1	0,84509	0,84509
2018	2	4	1	2
Σ	0	10	3,32221	2,72019303

$n=5$

$\hat{y} = a \cdot b^t$

$a' = \dfrac{\Sigma \log y}{n} = 0,664438589$

$a = 10^{a'}$

$b = 10^{b'}$

$b' = \dfrac{\Sigma \log y \cdot t}{t^2} = $ ____

$a = 4,617892927$

$b = 1,87075076$

FEL (SE LÖSNINGAR) −2p

$\hat{y}$ 4,617892927 · 1,870750 76

b) år 2019 $t=3$ kom försäljningen vara

$4,617892927 \cdot 1,820750 76^3 = 30,2337244$

svar: 30,2337244 HTkr

−2p

SU, STATISTIK

Skrivsal: UG **Anonymkod:** 0009 KWS **Blad nr:** 4

3 a)

Ftest $H_0: \beta_0 = \beta_1 = \beta_2 = 0$ 3) 20p

H_1: minst en $\beta_0 \neq 0$

Signifikant nivå: $\alpha = 5\%$

beslutsregel om $f_{obs} > f_{krit}$ förkasta H_0

$f_{(2,17)} = 3,59$ R

givet SSR = 1397,64   MSR = 698,82

SSE = 1171,16   MSE = 68,89176471

SST = 2568,80

$\frac{MSR}{MSE} = 10,14373783 > 3,59$ R R

R

Slutsats förkasta H_0 R a) 10p

t test - b_1 $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

Signifikans nivå 5%

beslutsregel om $t_{obs} > |t_{krit}|$ förkasta H_0

$t_{krit} = t_{(18)}(0,05) = 1,734$

$t_{obs} = \frac{0,97}{0,26} = 3,730769231 > 1,734$

förkasta H_0

A

t-test x_2

$H_0\ \beta_2 = 0$

$H_1\ \beta_2 \neq 0$

Signifikans nivå 5%

beslutsregel om $|t_{obs}| > |t_{krit}|$ förkasta H_0

$t_{krit} = |1,734|$

$t_{obs}\ \left|\frac{-13,71}{3,87}\right| = 3,542635659 > 1,734$

förkasta H_0

Slutsats testen visar att modelln ger sigintfikanta resultat och borde ändras.

B $Y = inkomst$ i TKR

$x_1 = ålder$

$x_2 = 1$ - kvinna, 0 = man

$Y = 14,64 + 0,97x_1 + 13,71x_2$

Vid samma ålder så tjänar en man enligt modeln 13710 kr mer än en kvinna vid samma ålder R

b) 1p

SU, STATISTIK

Skrivsal: UG Anonymkod: 0009-KWE Blad nr: 5

4) 20p

$T_t = a_0 + b \cdot t + C_1 \cdot D_1 + C_2 \cdot D_2 + C_3 \cdot D_n$

a) $D_1 = 1$ om kvartal 1 annars 0

$D_2 = 1$ om kvartal 2 annars 0

$D_3 = 1$ om kvartal 3 annars 0

$a_0 = \bar{Y} - b \cdot \bar{t}$

$\bar{t} = \frac{1+2 \ldots +16}{16} = 8,5$

$\bar{Y} = 4,75000$

$C_1 = 3,05625$

$C_2 = -3,46250$

$C_3 = -2,981255$

$b = 0,26875$

$a = 3,31250$

$a_0 = 2,465625$

$T = 3,31250 - 2,465625 + 3,05625 D_1 - 3,46250 D_2 - 2,98125 D_3$

$S_1 = 3,05625 - 2,465626 + 3,05625 = 3,646875$

$S_2 = 3,05625 - 2,465626 - 3,4625 = -2,871875$

$S_3 = 3,05625 - 2,465625 - 2,98125 = -2,390625$

$S_4 = 3,05625 - 2,465625$ $0,590625$

R — SE LÖSNINGAR

a) 10p

b) man kan se att det finns en tydlig trend att det används betydligt mindre

Δ

natur gas under kvartal 2 och kvartal 3 än blad det gör under 1 och 4 mest används under kvartal 1.

SE LÖSNINGAR

b) 0p

c) kvartal 2 2019 = t=18

kvartal 3 2019 = t=19

kvartal 2 = $a_0 + b \cdot 18 + c_2 = 2{,}465625 + 18 \cdot 0{,}26875 - 3{,}46250$

$= 3{,}840625$

R

kvartal 3 = $2{,}465625 + 19 \cdot 0{,}26875 - 2{,}98125 = 4{,}5915625$

R

c) 10p

Svar: kvartal 2 = 3,840625 m³ naturgas

kvartal 3 : 4,5915625 m³ natur gas.