STOCKHOLMS UNIVERSITET
Statistiska institutionen
Regressionsanalys och undersökningsmetodik, höstterminen 2016

Tentamen i Regressions- och tidsserieanalys den 5 januari 2017

Datum	2017-01-05
Tid:	14.00-19.00
Ansvarig Lärare:	Jörgen Säve-Söderbergh
Antal frågor:	5
Maxpoäng:	50
Hjälpmedel:	1) Språklexikon
	2) Kalkylator utan lagrade formler eller lagrad text
Tentamensgenomgång	Måndag 16 januari kl. 15.00 i B413.

Anvisningar

Redovisa dina lösningar i en form som gör det lätt att följa tankegången. Motivera alla väsentliga steg i lösningen. Ange alla antaganden och förutsättningar som du utnyttjar. Skriv endast på en sida av arket. Börja varje ny uppgift på nytt ark.

Lycka Till!

1. En butiksägare önskade studera sambandet mellan hur mycket plats en viss vara får ta i butiken och hur mycket varan säljs. En undersökning genomfördes under nio veckor, där man först hade en liten plats för varan, en kvadratdecimeter, ända till nio kvadratdecimeter den sista veckan. Nedan återges resultatet där x står för antalet kvadratdecimeter som varan exponeras på och y för antalet försålda enheter.

x	y
1	3
2	6
3	5
4	3
5	4
6	5
7	9
8	4
9	14

Som en första åtgärd i sin undersökning önskade butiksägaren att anpassa följande modell $Y = \beta_0 + \beta_1 X + \varepsilon$ till datamaterialet.

a) Beräkna minsta-kvadrat-skattningen av β_1. (1 p)

b) Beräkna minsta-kvadrat-skattningen av β_0. (1 p)

c) Beräkna ett 95%-igt konfidensintervall för β_1. (2 p)

d) Testa nollhypotesen $H_0 : \beta_1 = 0$ mot alternativet $H_1 : \beta_1 \neq 0$ på 5% signifikansnivå. (2 p)

e) Beräkna en prognos för Y då $x = 10$. (1 p)

f) Beräkna ett 95%-igt prediktionsintervall då $x = 10$. (3 p)

2. En börsanalytiker skattade följande samband för ett livsmedelsföretag

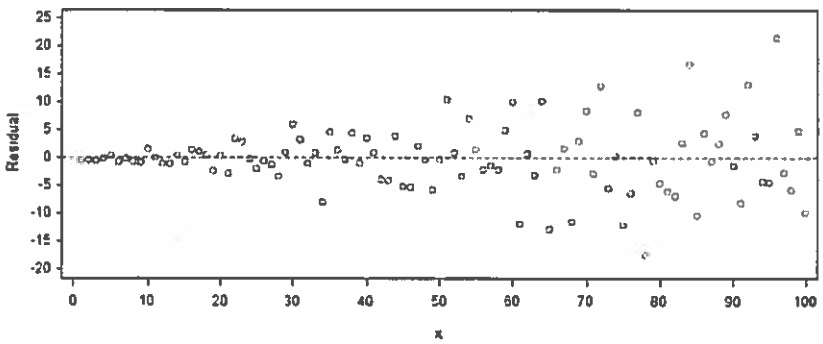
Utgifter för forskning = 7.83+0.026 Intäkter från försäljning+15.61 Rykten,

där utgifterna för forskning mäts i miljoner kronor, liksom intäkterna från försäljning. Rykten är en dummyvariabel som antar värdet 1 om många rykten florerar angående vad konkurrenterna gör och 0 vid stiltje på ryktessidan. Analytikern har en subtil metod för att skilja på *många* och *stiltje* som vi inte behöver bekymra oss om för närvarande.

a) Beräkna en prognos för utgifterna för forskning då intäkterna från försäljningen är tio miljoner kronor och ryktena kring vad konkurrenterna gör är många. (5 p)

b) Beräkna en prognos för utgifterna för forskning då intäkterna från försäljningen är tio miljoner kronor och det inte verkar hända något på konkurrenternas anläggningar rörande forskning. (5 p)

3. En forskare i psykologi anpassade en enkel lineär regressionsmodell till två variabler x och y. Följande graf visar residualerna från denna resgressionsanalys plottade mot x-värdena. Psykologen visade grafen för



dig och undrade om allting stod rätt till. Gör det det? (10 p)

4. En forskare i Göteborg har funnit att korrelationen mellan systoliskt blodtryck och bukomfång var $r = 0.298$ i ett stickprov om $n = 147$ individer. Testa på 5% signifikansnivå hypotest

$$H_0 : \rho = 0 \text{ mot } H_1 : \rho \neq 0 \quad (10p)$$

5. En psykiatriker hade behandlat patienter som hade genomlidit psykoser. Psykiatrikern önskade veta vilken patologisk avvikelse (Y) patienterna uppvisade sex månader efter behandling och hoppades att den kunde predikteras med rimlig precision utifrån från dels kunskapen om patienternas uppmätta tvångsmässiga tankar (X_1) (*Variabeln heter* ***think*** *i SAS-utskriften*) och dels deras fientliga inställning till omgivningen (X_2) (**suspicio** i utskriften). Samtliga variabler är uppmätta på olika

psykologiska skalor som sätter siffror på de psykologiska tillstånden. Data insamlades för $n = 53$ patienter.

En preliminär analys av datamaterialet gjordes och resultatet återfinns i bilaga 1.

a) Beräkna determinationskoefficienten R^2 då psykiatrikern har anpassat modellen $Y = \beta_0 + \beta_1 X_1 + \varepsilon$. (3 p)

b) Testa om variabeln X_2 bidrar till att prediktera Y i en modell som redan innehåller variabeln X_1. Använd 5% signifikansnivå. (5 p)

c) Genomför testet i b) genom att använda t-fördelningen istället. (2 p)

STOCKHOLMS UNIVERSITET
Statistiska institutionen
Undersökningmetodik och regressionsanalys, ht 2016
Jörgen Säve-Söderbergh

Bilaga till uppgift 5 på tentamen i Regressionsanalys den 5 januari 2017

Regression av Y på X1

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1535.85697	1535.85697	6.39	0.0146
Error	51	12255	240.30025		
Corrected Total	52	13791			

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-25.04084	19.00283	-1.32	0.1935
think	1	15.95111	6.30947	2.53	0.0146

Regression av Y på X1 och X2

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	2753.87136	1376.93568	6.24	0.0038
Error	50	11037	220.74597		
Corrected Total	52	13791			

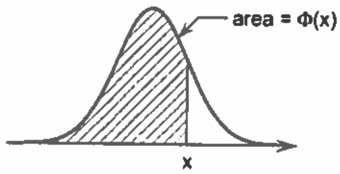
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-0.63535	20.96833	-0.03	0.9759
think	1	23.45144	6.83851	3.43	0.0012
suspicio	1	-7.07261	3.01092	-2.35	0.0228

Tabeller

Tabell 1. Standardiserad normalfördelning

$\Phi(x) = P(X \leq x)$ där $X \in N(0, 1)$

För negativa värden, utnyttja att $\Phi(x) = 1 - \Phi(-x)$

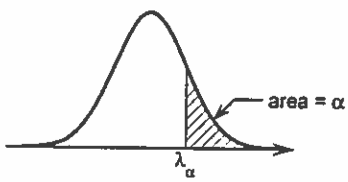


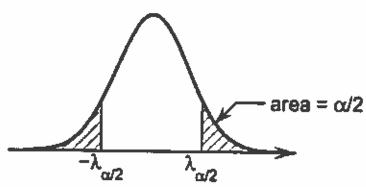
x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.99865									
3.1	.99903									
3.2	.99931									
3.3	.99952									
3.4	.99966									
3.5	.99977									
3.6	.99984									
3.7	.99989									
3.8	.99993									
3.9	.99995									
4.0	.99997									

Tabell 2. Normalfördelningens kvantiler

$P(X > \lambda_\alpha) = \alpha$ där $X \in N(0, 1)$

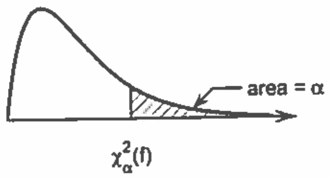
α	λ_α	α	λ_α
0.1	1.2816	0.001	3.0902
0.05	1.6449	0.0005	3.2905
0.025	1.9600	0.0001	3.7190
0.01	2.3263	0.00005	3.8906
0.005	2.5758	0.00001	4.2649





Tabell 4. χ^2**-fördelningen**

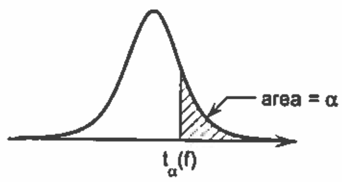
$P(X > \chi^2_\alpha(f)) = \alpha$ där $X \in \chi^2(f)$



f \ α	0.9995	0.999	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005	0.001	0.0005
1	0.00	0.00	0.00	0.00	0.00	0.00	3.84	5.02	6.63	7.88	10.83	12.12
2	0.00	0.00	0.01	0.02	0.05	0.10	5.99	7.38	9.21	10.60	13.82	15.20
3	0.02	0.02	0.07	0.11	0.22	0.35	7.81	9.35	11.34	12.84	16.27	17.73
4	0.06	0.09	0.21	0.30	0.48	0.71	9.49	11.14	13.28	14.86	18.47	20.00
5	0.16	0.21	0.41	0.55	0.83	1.15	11.07	12.83	15.09	16.75	20.52	22.11
6	0.30	0.38	0.68	0.87	1.24	1.64	12.59	14.45	16.81	18.55	22.46	24.10
7	0.48	0.60	0.99	1.24	1.69	2.17	14.07	16.01	18.48	20.28	24.32	26.02
8	0.71	0.86	1.34	1.65	2.18	2.73	15.51	17.53	20.09	21.95	26.12	27.87
9	0.97	1.15	1.73	2.09	2.70	3.33	16.92	19.02	21.67	23.59	27.88	29.67
10	1.26	1.48	2.16	2.56	3.25	3.94	18.31	20.48	23.21	25.19	29.59	31.42
11	1.59	1.83	2.60	3.05	3.82	4.57	19.68	21.92	24.72	26.76	31.26	33.14
12	1.93	2.21	3.07	3.57	4.40	5.23	21.03	23.34	26.22	28.30	32.91	34.82
13	2.31	2.62	3.57	4.11	5.01	5.89	22.36	24.74	27.69	29.82	34.53	36.48
14	2.70	3.04	4.07	4.66	5.63	6.57	23.68	26.12	29.14	31.32	36.12	38.11
15	3.11	3.48	4.60	5.23	6.26	7.26	25.00	27.49	30.58	32.80	37.70	39.72
16	3.54	3.94	5.14	5.81	6.91	7.96	26.30	28.85	32.00	34.27	39.25	41.31
17	3.98	4.42	5.70	6.41	7.56	8.67	27.59	30.19	33.41	35.72	40.79	42.88
18	4.44	4.90	6.26	7.01	8.23	9.39	28.87	31.53	34.81	37.16	42.31	44.43
19	4.91	5.41	6.84	7.63	8.91	10.12	30.14	32.85	36.19	38.58	43.82	45.97
20	5.40	5.92	7.43	8.26	9.59	10.85	31.41	34.17	37.57	40.00	45.31	47.50
21	5.90	6.45	8.03	8.90	10.28	11.59	32.67	35.48	38.93	41.40	46.80	49.01
22	6.40	6.98	8.64	9.54	10.98	12.34	33.92	36.78	40.29	42.80	48.27	50.51
23	6.92	7.53	9.26	10.20	11.69	13.09	35.17	38.08	41.64	44.18	49.73	52.00
24	7.45	8.08	9.89	10.86	12.40	13.85	36.42	39.36	42.98	45.56	51.18	53.48
25	7.99	8.65	10.52	11.52	13.12	14.61	37.65	40.65	44.31	46.93	52.62	54.95
26	8.54	9.22	11.16	12.20	13.84	15.38	38.89	41.92	45.64	48.29	54.05	56.41
27	9.09	9.80	11.81	12.88	14.57	16.15	40.11	43.19	46.96	49.64	55.48	57.86
28	9.66	10.39	12.46	13.56	15.31	16.93	41.34	44.46	48.28	50.99	56.89	59.30
29	10.23	10.99	13.12	14.26	16.05	17.71	42.56	45.72	49.59	52.34	58.30	60.73
30	10.80	11.59	13.79	14.95	16.79	18.49	43.77	46.98	50.89	53.67	59.70	62.16
40	16.91	17.92	20.71	22.16	24.43	26.51	55.76	59.34	63.69	66.77	73.40	76.09
50	23.46	24.67	27.99	29.71	32.36	34.76	67.50	71.42	76.15	79.49	86.66	89.56
60	30.34	31.74	35.53	37.48	40.48	43.19	79.08	83.30	88.38	91.95	99.61	102.69
70	37.47	39.04	43.28	45.44	48.76	51.74	90.53	95.02	100.43	104.21	112.32	115.58
80	44.79	46.52	51.17	53.54	57.15	60.39	101.88	106.63	112.33	116.32	124.84	128.26
90	52.28	54.16	59.20	61.75	65.65	69.13	113.15	118.14	124.12	128.30	137.21	140.78
100	59.90	61.92	67.33	70.06	74.22	77.93	124.34	129.56	135.81	140.17	149.45	153.17

Tabell 3. ***t*****-fördelningen**

$P(X > t_\alpha(f)) = \alpha$ där $X \in t(f)$



f	α 0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.08	6.31	12.71	31.82	63.66	318.31	636.62
2	1.89	2.92	4.30	6.96	9.92	22.33	31.60
3	1.64	2.35	3.18	4.54	5.84	10.21	12.92
4	1.53	2.13	2.78	3.75	4.60	7.17	8.61
5	1.48	2.02	2.57	3.36	4.03	5.89	6.87
6	1.44	1.94	2.45	3.14	3.71	5.21	5.96
7	1.41	1.89	2.36	3.00	3.50	4.79	5.41
8	1.40	1.86	2.31	2.90	3.36	4.50	5.04
9	1.38	1.83	2.26	2.82	3.25	4.30	4.78
10	1.37	1.81	2.23	2.76	3.17	4.14	4.59
11	1.36	1.80	2.20	2.72	3.11	4.02	4.44
12	1.36	1.78	2.18	2.68	3.05	3.93	4.32
13	1.35	1.77	2.16	2.65	3.01	3.85	4.22
14	1.35	1.76	2.14	2.62	2.98	3.79	4.14
15	1.34	1.75	2.13	2.60	2.95	3.73	4.07
16	1.34	1.75	2.12	2.58	2.92	3.69	4.01
17	1.33	1.74	2.11	2.57	2.90	3.65	3.97
18	1.33	1.73	2.10	2.55	2.88	3.61	3.92
19	1.33	1.73	2.09	2.54	2.86	3.58	3.88
20	1.33	1.72	2.09	2.53	2.85	3.55	3.85
21	1.32	1.72	2.08	2.52	2.83	3.53	3.82
22	1.32	1.72	2.07	2.51	2.82	3.50	3.79
23	1.32	1.71	2.07	2.50	2.81	3.48	3.77
24	1.32	1.71	2.06	2.49	2.80	3.47	3.75
25	1.32	1.71	2.06	2.49	2.79	3.45	3.73
26	1.31	1.71	2.06	2.48	2.78	3.43	3.71
27	1.31	1.70	2.05	2.47	2.77	3.42	3.69
28	1.31	1.70	2.05	2.47	2.76	3.41	3.67
29	1.31	1.70	2.05	2.46	2.76	3.40	3.66
30	1.31	1.70	2.04	2.46	2.75	3.39	3.65
40	1.30	1.68	2.02	2.42	2.70	3.31	3.55
60	1.30	1.67	2.00	2.39	2.66	3.23	3.46
120	1.29	1.66	1.98	2.36	2.62	3.16	3.37
∞	1.28	1.64	1.96	2.33	2.58	3.09	3.29

Tabell 2. F-fördelningens kvantiler

$X \in F(\nu_1,\nu_2)$ där ν_1,ν_2 = antal frihetsgrader i täljaren respektive nämnaren. Vilket värde har f_α om $P(X > f_\alpha) = \alpha$ där α är en given sannolikhet.

$\alpha = 0{,}05$

	ν_1 = 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ν_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

Tabell 2 forts. *F*-fördelningens kvantiler

$\alpha = 0{,}05$

	$\nu_1 =$ 16	17	18	19	20	25	30	35	40	50	60	70	80	100	∞
$\nu_2 = 1$	246,5	246,9	247,3	247,7	248,0	249,3	250,1	250,7	251,1	251,8	252,2	252,5	252,7	253,0	254,3
2	19,43	19,44	19,44	19,44	19,45	19,46	19,46	19,47	19,47	19,48	19.48	19,48	19,48	19,49	19,50
3	8,69	8,68	8,67	8,67	8,66	8,63	8,62	8,60	8,59	8,58	8,57	8,57	8,56	8.55	8,53
4	5,84	5,83	5,82	5,81	5,80	5,77	5,75	5,73	5,72	5,70	5,69	5,68	5,67	5,66	5,63
5	4,60	4,59	4.58	4,57	4,56	4,52	4,50	4,48	4,46	4,44	4.43	4,42	4,41	4,41	4,37
6	3,92	3,91	3,90	3,88	3,87	3,83	3,81	3,79	3,77	3,75	3,74	3,73	3,72	3,71	3,67
7	3,49	3,48	3,47	3,46	3.44	3,40	3,38	3,36	3,34	3,32	3.30	3.29	3,29	3,27	3,23
8	3,20	3,19	3.17	3,16	3,15	3,11	3,08	3,06	3,04	3,02	3,01	2,99	2,99	2,97	2,93
9	2,99	2,97	2,96	2,95	2,94	2,89	2,86	2,84	2,83	2,80	2,79	2,78	2,77	2,76	2,71
10	2,83	2,81	2,80	2,79	2,77	2,73	2,70	2,68	2,66	2,64	2,62	2,61	2,60	2,59	2,54
11	2,70	2,69	2.67	2,66	2,65	2,60	2,57	2,55	2,53	2,51	2.49	2,48	2,47	2,46	2,40
12	2,60	2,58	2,57	2,56	2,54	2,50	2,47	2,44	2,43	2,40	2,38	2,37	2,36	2,35	2,30
13	2,51	2,50	2,48	2,47	2,46	2,41	2,38	2,36	2,34	2,31	2.30	2,28	2,27	2,26	2,21
14	2,44	2,43	2,41	2,40	2,39	2,34	2,31	2,28	2,27	2,24	2,22	2,21	2,20	2,19	2,13
15	2,38	2,37	2,35	2,34	2,33	2,28	2,25	2,22	2,20	2,18	2,16	2.15	2,14	2,12	2,07
16	2,33	2,32	2.30	2,29	2,28	2,23	2,19	2,17	2,15	2,12	2,11	2,09	2,08	2,07	2,01
17	2,29	2,27	2,26	2,24	2,23	2,18	2,15	2,12	2,10	2,08	2.06	2,05	2,03	2,02	1,96
18	2,25	2,23	2,22	2,20	2,19	2,14	2,11	2,08	2,06	2,04	2,02	2,00	1,99	1,98	1,92
19	2,21	2,20	2,18	2,17	2,16	2,11	2,07	2,05	2,03	2,00	1,98	1,97	1,96	1,94	1,88
20	2,18	2,17	2,15	2,14	2,12	2,07	2,04	2,01	1,99	1,97	1,95	1,93	1,92	1,91	1,84
25	2,07	2,05	2,04	2,02	2,01	1,96	1,92	1,89	1,87	1,84	1,82	1,81	1,80	1,78	1,71
30	1,99	1,98	1,96	1,95	1,93	1,88	1,84	1,81	1,79	1,76	1.74	1,72	1,71	1,70	1,62
35	1,94	1,92	1,91	1,89	1,88	1,82	1,79	1,76	1,74	1,70	1,68	1,66	1,65	1,63	1,56
40	1,90	1,89	1,87	1,85	1,84	1,78	1,74	1,72	1,69	1,66	1,64	1,62	1,61	1,59	1,51
45	1,87	1,86	1,84	1,82	1,81	1,75	1,71	1,68	1,66	1,63	1,60	1,59	1,57	1,55	1,47
50	1,85	1,83	1,81	1,80	1,78	1,73	1,69	1,66	1,63	1,60	1,58	1,56	1,54	1,52	1,44
60	1,82	1,80	1,78	1,76	1,75	1,69	1,65	1,62	1,59	1,56	1,53	1,52	1,50	1,48	1,39
70	1,79	1,77	1,75	1,74	1,72	1,66	1,62	1,59	1,57	1,53	1,50	1,49	1,47	1,45	1,35
80	1,77	1,75	1,73	1.72	1,70	1,64	1,60	1,57	1,54	1,51	1,48	1,46	1,45	1,43	1,32
100	1,75	1,73	1,71	1.69	1,68	1,62	1,57	1,54	1,52	1,48	1,45	1,43	1,41	1,39	1,28
∞	1,64	1,62	1,60	1,59	1,57	1,51	1,46	1,42	1,39	1,35	1,32	1,29	1,27	1,24	1,00

Formelsamling regressionsanalys

Enkel linjär regression

$$\hat{\beta}_1 = \frac{SSXY}{SSX} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{n\sum_{i=1}^{n}(x_i y_i) - \sum_{i=1}^{n}(x_i)\sum_{i=1}^{n}(y_i)}{n\sum_{i=1}^{n}(x_i^2) - \left(\sum_{i=1}^{n}x_i\right)^2} = \frac{\sum_{i=1}^{n}(x_i y_i) - n\bar{x}\bar{y}}{\sum_{i=1}^{n}(x_i^2) - n(\bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$$

SSY=SSR+SSE

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

$$r = \frac{SSXY}{\sqrt{SSX \times SSY}} = \frac{S_X}{S_Y}\hat{\beta}_1$$

$$S_{Y|X}^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

$$S_{\hat{\beta}_1} = \frac{S_{Y|X}}{S_X\sqrt{n-1}} \qquad S_{\hat{\beta}_0} = S_{Y|X}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{(n-1)S_X^2}}$$

$$\hat{Y}_{X_0} \pm t_{n-2,1-\alpha/2}S_{Y|X}\sqrt{1 + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{(n-1)S_X^2}} \qquad \hat{Y}_{X_0} \pm t_{n-2,1-\alpha/2}S_{Y|X}\sqrt{\frac{1}{n} + \frac{(x_0-\bar{x})^2}{(n-1)S_X^2}}$$

$$b_1 \pm t_{n-2,1-\alpha/2}S_{b_1}$$

Multipel regression

$$F = \frac{MSR}{MSE} = \frac{(SSY - SSE)/k}{SSE/(n-k-1)}$$

$$b_j \pm t_{n-k-1,1-\alpha/2}S_{b_j}$$

$$R_{Y|X_1,\ldots,X_k} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}}_i)^2}}$$

$$R_{Y|X_1,\ldots,X_k}^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2 - \sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = \frac{SSY - SSE}{SSY}$$

Korrelation

$$Z = \frac{\frac{1}{2}ln\left[\frac{(1+r)}{(1-r)}\right] - \frac{1}{2}ln\left[\frac{(1+\rho_0)}{(1-\rho_0)}\right]}{1/\sqrt{n-3}} \sim N(0,1)$$

Statistiska institutionen

Stockholms universitet

Rättningsblad

Datum: 5/1-17

Sal: Brunnsvikssalen

Tenta: Regressions- och tidsserieanalys

Kurs: Regressionsanalys och undersökningsmetodik

ANONYMKOD: REG-0034

☒ 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					11
Lär.ant. ~~[illegible]~~	10	10	4	9					

7

HJ

POÄNG	BETYG	Lärarens sign.
40	B	

SU, STATISTIK

Skrivsal: Brunnsvikssalen **Anonymkod:** REG-0034 **Blad nr:** 1

1 a) $\hat{\beta}_1 = \frac{\sum(x_i y_i) - n\bar{x}\bar{y}}{\sum(x_i^2) - n(\bar{x})^2} = \frac{313 - (9 \cdot 5 \cdot 5{,}89)}{285 - (9 \cdot (5^2))} = 0{,}8$

$n = 9$

X	y	$x_i y_i$	x_i^2
1	3	3	1
2	6	12	4
3	5	15	9
4	3	12	16
5	4	20	25
6	5	30	36
7	9	63	49
8	4	32	64
9	14	126	81
$\sum$ 45	53	313	285

$\bar{x} = 5$ $\bar{y} = 5{,}89$

Svar: $\hat{\beta}_1 = 0{,}8$ R

b) $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

$\hat{\beta}_0 = 5{,}89 - (0{,}8 \cdot 5) = 1{,}89$

Svar: $\hat{\beta}_0 = 1{,}89$ R

SU, STATISTIK

Skrivsal: Brannviksssalen Anonymkod: REG-0054 Blad nr: 2

1c) 95%-igt konfidensintervall för β_1 fås genom formeln $\hat{\beta}_1 \pm t_{n-2,1-\alpha/2} s_{\beta_1}$

$$s_{\beta_1} = \frac{s_{y|x}}{s_x \cdot \sqrt{n-1}}$$

$$s_{y|x} = \sqrt{s^2_{y|x}} \qquad s^2_{y|x} = \frac{1}{n-2} \sum_{i=1}^{n} (y_i - \hat{y})^2$$

Formeln för $\hat{y}$ har vi från 1a)

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1$ $\hat{y} = 1{,}89 + 0{,}8 \cdot x$

x	$\beta_0 + \beta_1 x_1 = \hat{y}$	$(y_i - \hat{y})^2$
1	$1{,}89 + (0{,}8 \cdot 1) = 2{,}69$	0,0016
2	$1{,}89 + (0{,}8 \cdot 2) = 3{,}49$	6,3001
3	$1{,}89 + (0{,}8 \cdot 3) = 4{,}29$	0,4041
4	$1{,}89 + (0{,}8 \cdot 4) = 5{,}09$	4,3681
5	$1{,}89 + (0{,}8 \cdot 5) = 5{,}89$	3,5721
6	$1{,}89 + (0{,}8 \cdot 5) = 6{,}69$	2,8561
7	$1{,}89 + (0{,}8 \cdot 5) = 7{,}49$	2,2801
8	$1{,}89 + (0{,}8 \cdot 5) = 8{,}29$	18,4041
9	$1{,}89 + (0{,}8 \cdot 5) = 9{,}09$	24,1081

$\sum$ 62,4889 R

Fortsättning →

SU, STATISTIK

Skrivsal: Brunnsviksalen Anonymkod: REG-0034 Blad nr: 3

1 c)

$$S^2_{y|x} = \frac{1}{n-1}\sum(y_i-\hat{y})^2 = \frac{1}{9-2}\cdot 62,4889 = 8,927$$ R

$$S_{y|x} = \sqrt{8,927} = 2,9878$$

$$S_x = \sqrt{S_x^2} = \frac{1}{n-1}\sum(x_i-\bar{x})^2 = \frac{1}{9-1}\cdot 60 = 7,5$$

$$\sqrt{7,5} = 2,7386$$

x	$(x_i-\bar{x})^2$
1	16
2	9
3	4
4	1
5	0
6	1
7	4
8	9
9	16

$\sum 60$

Det vi vill räkna ut är $s\beta_1$, formeln är $\frac{s_{y|x}}{s_x\cdot\sqrt{n-1}}$ där vi har

$s_{y|x} = 2,9878$ och $s_x = 2,7386$.

$$s\beta_1 = \frac{2,9878}{2,7386\cdot\sqrt{9-1}} = 0,3857$$ R

Formeln för 95%-igt ki är

$\beta_1 \pm t_{n-2,\,1-\alpha/2}\, s\beta_1$ vilket blir

$0,8 \pm t(7)_{0,025}\ s\beta_1 = 2,36\cdot 0,3857 = 0,9103$

Svar:

$0,8 \pm 0,9103$

95%-igt ki: $(-0,1103\ ;\ 1,7103)$ R

SU, STATISTIK

Skrivsal: Brunnsvikssalen Anonymkod: REG-0034 Blad nr: 4

1d) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

$$t_{obs} = \frac{\beta_1 - \beta_1^{(0)}}{S_{\beta_1}} = \frac{0,8}{0,3857} = 2,07412$$ R

$t_{krit}: t_{n-k-1,\ \alpha/2} = t(7)_{0,025}$ vilket enligt tabell 3 har värdet 2,36

Vi förkastar Ho om $t_{obs} > t_{krit}$

$2,07412 < 2,36$

Svar: Den är ej signifikant på 5% nivå. R

1e) Vårt $\hat{Y}$ var $\hat{Y} = 1,89 + 0,8X$

Nu ska vi predicera Y när $X = 10$

$Y = 1,89 + 0,8 \cdot 10 = 9,89$ R

Svar: Prognosen för y när $X = 10$ är 9,89

SU, STATISTIK

Skrivsal: Brunnsvikssalen **Anonymkod:** REG-0034 **Blad nr:** 5

2

Y = Utgifter för forskning

X_1 = Intäkter från försäljning

X_2 = Rykten

Formeln blir då: $Y = 7{,}83 + 0{,}026X_1 + 15{,}61X_2$

a) Då intäkterna är 10 miljoner kronor antar variabeln X_1 värdet 10

"Många rykten" innebär att dummyvariabeln X_2 antar värdet 1

$Y = 7{,}83 + 0{,}026 \cdot 10 + 15{,}61 \cdot 1$

$Y = 23{,}7$

Svar: Vid denna situation är utgifterna för forskning 23,7 miljoner kronor R

b) X_1 har samma värde på 10

X^2 antar värdet 0.

$Y = 7{,}83 + 0{,}026 \cdot 10 + 15{,}61 \cdot 0$

$Y = 8{,}09$

Svar: Utgifterna för forskning blir vid detta tillfälle 8,09 miljoner kronor R

SU, STATISTIK

Skrivsal: Brunnvikssalen **Anonymkod:** REG-0034 **Blad nr:** 6

3) För att kunna dra några slutsatser kring datan behöver jag göra en deskriptiv analys och kontrollera att fem antaganden uppfylls:

Existens: Är y en stokastisk variabel av x? exakt linje!
Linjäritet: Går det att se något linjärt samband?
Oberoende: Är variablerna oberoende?
Normalitet: Är datan normalfördelad?
Homoskedasticitet

Att existens uppfylls är något vi antar, det går att se ett linjärt samband mellan x och y. Om variablerna är oberoende är något vi skulle behöva testa med hjälp av en korrelationsmatris där en hög korrelation skulle peka på ett beroende och en låg korrelation peka på oberoende. Inte heller normalitet går att bevisa/motbevisa med det material vi har utan en mer anpassad och detaljerad analys skulle behöva göras

SU, STATISTIK

Skrivsal: Brännvinssalen **Anonymkod:** REG-0034 **Blad nr:** 7

(3) fortsättning!

Det antagande som inte uppfylls och som går att avläsa av grafen är antagandet om homoskedasticitet. Ju högre X-värde desto större blir variansen för residualerna vilket är väldigt tydligt när man tittar på grafen. Detta gör att datan blir väldigt svår att analysera i och med så pass stigande varianser. Datan är rent utsagt heteroskedastiskt, residualerna har olika varianser för olika värden på X.

Bra

SU, STATISTIK

Skrivsal: Brunnviksalen Anonymkod: REG-0034 Blad nr: 8

(4) $r = 0,298$

$n = 147$

$\rho = 0$

Jag behöver använda korrelationsmatrisen för att få ut ett värde på Z.

$$Z = \frac{\frac{1}{2}\ln\left[\frac{1+r}{1-r}\right] - \frac{1}{2}\ln\left[\frac{1+\rho_0}{1-\rho_0}\right]}{1/\sqrt{n-3}}$$

Fel testvariabel

$$Z = \frac{0,5\ln\left[\frac{1+0,298}{1-0,298}\right] - 0,5\ln\left[\frac{1+0}{1-0}\right]}{1/\sqrt{n-3}}$$

$$Z = \frac{0,3073 - 0}{1/12} = 3,6876$$

Nu har jag ett observerat Z-värde på 3,6876.
Det kritiska värdet fås genom $Z_{0,05} = 1,6449$.

Vi förkastar $H_0: \rho = 0$ om det observerade Z-värdet är större än det kritiska vilket stämmer då $3,6876 > 1,6449$.

Svar: Vi förkastar nollhypotesen på 5% signifikansnivå.

4p

SU, STATISTIK

Skrivsal: Brannvikssalen Anonymkod: REG-0034 Blad nr: 9

5

a)

Determinationskoefficienten R^2 visar hur stor del av modellen som kan förklaras med hjälp av X-variablerna och fås genom formeln

$$\frac{SSR}{SSY}$$

Både SSR och SSY fås i uppgiften

SSR = 1535,85697

SSY = 13791

$$R^2 = \frac{SSR}{SSY} = \frac{1535,85697}{13791} = 0,1114$$

Svar: Då psykiatrikern har anpassat modellen $Y = \beta_0 + \beta_1 X_1 + \varepsilon$ är $R^2 = 0,1114$.

SU, STATISTIK

Skrivsal: Brunnvikssalen Anonymkod: REG-0034 Blad nr: 10

(5)

b) För att ta reda på om X2 bidrar måste jag jämföra SSR mellan modellen för X1 och X2 med SSR för modellen med endast X1. För det vi vill ta reda på är om förklaringsgraden ökat

Bra

$SSR(X1, X2) - SSR(X1) = 2753{,}87136 - 1535{,}85697 = 1218{,}0144$ R

Sedan använder jag formeln $F(X1, X2) = \frac{SSR(X1,X2) - SSR(X1)}{MSE(X1,X2)}$

för att få ett observerant F-värde

$F_{obs} = \frac{1218{,}0144}{220{,}74597} = 5{,}5177$ R

enligt tabell 2

$F_{krit}: F_{k,\, n-k-1\, \alpha} = F(\cancel{2}\, 1; 50)_{0{,}05} \longrightarrow 3{,}18$ ✓

X2 är signifikant om $F_{obs} > F_{krit}$ $5{,}5177 > 3{,}18$

Svar: X2 bidrar till att prediktera Y på 5% signifikansnivå. R

SU, STATISTIK

Skrivsal: Brannvikssalen **Anonymkod:** REG-0034 **Blad nr:** 11

5

c) I ett t-test utgår jag från parameterskattningarna och använder testvariabeln

$$t_{obs} = \frac{\beta_2 - \beta_1^{(0)}}{s_{\beta_2}}$$

$\beta_2 = -7{,}07261$

$s_{\beta_2} = 3{,}01092$

$$t_{obs} = \frac{-7{,}07261 - 0}{3{,}01092} = -2{,}349 \quad R$$

enligt tabell 3

$t_{krit}: t_{1-\alpha/2,\, n-k-1} = t_{0{,}025}(50) \longrightarrow 2{,}01 \quad R$

OBS: Värdet 50 finns inte i tabellen så jag tog värdena för 40 (2,02) och 60 (2,00) och delade med 2 för 2,01. OK

Då testet är dubbelsidigt spelar minustecknet i tobs ingen roll.

Vi förkastar nollhypotesen om $t_{obs} > t_{krit}$

$2{,}349 > 2{,}01$

Svar: Vi förkastar nollhypotesen på 5% nivå. R

Statistiska institutionen

Stockholms universitet

Rättningsblad

Datum: 5/1-17

Sal: Brunnsvikssalen

Tenta: Regressions- och tidsserieanalys

Kurs: Regressionsanalys och undersökningsmetodik

ANONYMKOD: reg 0030

☒ 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					4
Lär.ant. ~~8~~ 6	10	10	10	10					

HS

POÄNG	BETYG	Lärarens sign.
46	~~B~~ A	JSS

SU, STATISTIK

Skrivsal: BrunnsV. Anonymkod: reg 0030 [redacted] Blad nr: 1

1. $n = 9$ $Y = B_0 + B_1 X + \varepsilon$

X	Y	XY	X^2	$(X-\bar{X})^2$
1	3	3	1	16
2	6	12	4	9
3	5	15	9	4
4	3	12	16	1
5	4	20	25	0
6	5	30	36	1
7	9	63	49	4
8	4	32	64	9
9	14	126	81	16
Σ 45	53	313	285	60

$\bar{X}$: $45/9 = 5$ $\bar{Y}$: $53/9 = 5{,}89$

$$\hat{B}_1 = \frac{\sum_{i=1}^{n}(X_i Y_i) - n\bar{x}\bar{y}}{\sum_{i=1}^{n}(x_i^2) - n(\bar{x})^2}$$

$\hat{B}_0 = \bar{y} - \hat{B}_1 \bar{x}$ sätter in värden i formlerna:

a) $$\hat{B}_1 = \frac{313 - 9 \cdot 5 \cdot 5{,}89}{285 - 9(5)^2} = \frac{313 - 265{,}05}{285 - 225} = \frac{47{,}95}{60} = 0{,}799 \approx 0{,}80$$

$\hat{B}_1 = 0{,}80$ R

b) $\hat{B}_0 = 5{,}89 - 0{,}80 \cdot 5 = 1{,}89$

$\hat{B}_0 = 1{,}89$ R

c) $b_1 \pm t_{n-2,\,1-\alpha/2}\, S_{b_1}$

$t_{(0{,}025;\,7)}$ ges av tabell 3: $2{,}36$ R

$$S_{\hat{b}_1} = \frac{S_{y|x}}{S_x \sqrt{n-1}}$$ där $$S^2_{y|x} = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$ och

$$S^2_x = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

→

$$s^2_{y|x} = \frac{1}{9-2} \sum (y_i - \hat{y}_i)^2 = \frac{1}{7} \cdot 63{,}35 = 9{,}05$$

$$s_{y|x} = \sqrt{9{,}05} = 3{,}01$$

$$s^2_x = \frac{1}{9-1} \sum (60) = 0{,}125 \cdot 60 = 7{,}5$$

$$s_x = \sqrt{7{,}5} = 2{,}74$$

Det som saknas för att få fram $s_{\hat\beta_1}$ nu är $\sum (y_i - \hat{y}_i)^2$.

Vår skattade regression blir $y = 1{,}89 + 0{,}8x$

x	y		$\hat{y}$	$(y-\hat{y})^2$
1	3	$\hat{y} = 1{,}89 + 0{,}8 \cdot 1 =$	2,69	$(3-2{,}69)^2 = 0{,}961$
2	6	$\hat{y} = 1{,}89 + 0{,}8 \cdot 2 =$	3,49	$(6-3{,}49)^2 = 6{,}30$
3	5	$\hat{y} = 1{,}89 + 0{,}8 \cdot 3 =$	4,29	$(5-4{,}29)^2 = 0{,}50$
4	3	-//- $\cdot 4 =$	5,09	$(3-5{,}09)^2 = 4{,}37$
5	4	$\cdot 5 =$	5,89	$(4-5{,}89)^2 = 3{,}57$
6	5	-//- $\cdot 6 =$	6,69	$(5-6{,}69)^2 = 2{,}86$
7	9	$\cdot 7 =$	7,49	$(9-7{,}49)^2 = 2{,}28$
8	4	-//- $\cdot 8 =$	8,29	$(4-8{,}29)^2 = 18{,}40$
9	14	$\cdot 9 =$	9,09	$(14-9{,}09)^2 = 24{,}11$
				$\sum$ 63,35

Ok

$$s_{\hat\beta_1} = \frac{3{,}01}{2{,}74\sqrt{9-1}} = 0{,}39$$

R

Så vårt koffidens intervall blir:

$0{,}8 \pm 2{,}36 \cdot 0{,}39 =$

$0{,}8 \pm 0{,}92$

$(-0{,}12 ; 1{,}72)$ Eftersom det täcker 0 så är det inte signifikant

R

Bra

d) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$ T-fördelning

$\alpha = 0{,}05 \Rightarrow$ dubbelsidigt $= 0{,}025_{(f=7)} = 2{,}36$

→

SU, STATISTIK

Skrivsal: Brunnsv Anonymkod: reg 0030 Blad nr: 2

1d forts) Beslutsregel: förkasta H0 om $t_{obs} > \text{kritiskt } (2{,}36)$

$\rightarrow$ tidigare uträknat

$$t_{obs} = \frac{\hat{\beta}_1}{S\hat{\beta}_1} = \frac{0{,}8}{0{,}39} = 2{,}05 \text{ R}$$

$2{,}05 < 2{,}36$ så vi kan ej förkasta H0. R

β_1 är således ej signifikant på 5% nivån.

e) $y = 1{,}89 + 0{,}8x$

$x = 10$

$1{,}89 + 0{,}8 \cdot 10 = \underline{11{,}89}$ ✓

f) $$\hat{y}_{x0} \pm t_{n-2,\,1-\alpha/2}\, S_{y/x} \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{(n-1)S_x^2}}$$

$$11{,}89 \pm \underset{2{,}36}{t(0{,}025;7)} \cdot 3{,}01 \sqrt{1 + \frac{1}{9} + \frac{(10-5)^2}{(9-1)\,7{,}5}} =$$

$$11{,}89 \pm 2{,}36 \cdot 3{,}01 \sqrt{1 + 0{,}11 + \frac{25}{60}} =$$

$$11{,}89 \pm 5{,}37\sqrt{1{,}526} =$$

$$11{,}89 \pm 5{,}37 \cdot 1{,}235 =$$

$$11{,}89 \pm 6{,}63$$

$(5{,}26;\ 18{,}52)$ ✓

$\rightarrow$

2. utgifter = 7,93 + 0,026 intäkter + 15,61 rykten

a) intäkter = 10 m
rykte = 1

utgifter = 7,83 + 0,026·10 + 15,61·1 = 23,7

utgifterna skattas till <u>23,7</u> miljoner kronor vid många rykten och försäljning på 10 m

b) utgifter = 7,83 + 0,026·10 + 15,61·0 = 8,09

utgifterna skattas till <u>8,09</u> mkr vid stilje och försäljning på 10 m.

Vi kan se att företaget beräknas gå med förlust i tider av många rykten och gå med vinst i tider av få rykten.

R

3) Nej, Grafen ser ut att bryta mot antagandet homoskedacitet! R Detta innebär att variansen hos y inte är konstant i hela spektrumet. SSE och MSE riskerar då bli högre och våra skattningar för parametrarna blir sämre. Den ser även ut att bryta mot normalitetantagandet. nej, i E prövas Detta betyder att för varje x kombo så antas y vara normalfördelad och om detta bryts så blir inte skattningarna så bra. Det är även möjligt att antagandet om oberoende är på väg brytas. Det innebär att observationerna av y ska vara oberoende varand om detta bryts får vi hög SSE och MSE och sämre skattningar

10

SU, STATISTIK

Skrivsal: Brunnsvik Anonymkod: rey 0030 Blad nr: 3

4. $r = 0{,}298$

$n = 147$

$H_0: \rho = 0$

$H_1: \rho \neq 0$

$r^2 = 0{,}089$

Testas enligt: $\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$

$\frac{0{,}298\sqrt{147-2}}{\sqrt{1-0{,}089}} =$

$\frac{0{,}298\sqrt{145}}{\sqrt{0{,}911}} = \frac{3{,}589}{0{,}9544} =$

$3{,}76$ R

T-kritiskt = 1,97
(0,025, 145)

Beslutsregel: Förkasta H_0 om obs > kritiskt

$3{,}76 > 1{,}97$ R

H_0 förkastas därmed på 5% signifikansnivå.

R

→

5) $n=53$

a) $Y = \beta_0 + \beta_1 X_1 + \varepsilon$

$$R^2 = \frac{SSR}{SSY}$$ Från tabell 1 n.Har vi värden

$SSY = 13791$

$SSR = 1535{,}85697$

$$\frac{1535{,}85697}{13791} = 0{,}11$$ R

$R^2 = \underline{0{,}11}$, dvs modellen förklarar 11 % av variationen i Y.

b) $H_0 : \beta_2 = 0$

$H_1 : \beta_2 \neq 0$

Test : $$\frac{SSR_{x_1 x_2} - SSR_{x_1}}{MSE_{x_1 x_2}} = \frac{2753{,}87136 - 1535{,}85697 = 1218{,}01439}{220{,}74597} = 5{,}52$$ R

Beslutsregel : Förkasta H_0 om obs > kritiskt

$F_{(2,50)} = 3{,}18$. $5{,}52 > 3{,}18$ och vi förkastar H_0.

β_2 bidrar därmed på 5 % signifikansnivå. R

c) T-test : $$\frac{\hat{\beta}_2}{S_{\hat{\beta}_2}} = \frac{-7{,}07261}{3{,}01092} = -2{,}35$$ R

$H_0 : \beta_2 = 0$

$H_1 : \beta_2 \neq 0$

Beslutsregel : Förkasta H_0 om T_{obs} > kritiskt $_{0{,}025} = 2{,}01$

→

SU, STATISTIK

Skrivsal: Brunnsv Anonymkod: reg 0030 Blad nr: 4

5 c forts)

$-2{,}35 < -2{,}01$ dvs vi kan förkasta

H_0. B_2 bidrar på 5% signifikansnivå.

1L