

# Försättsblad till skriftlig tentamen vid Linköpings universitet



Datum för tentamen	2017-08-17
Sal (1)	TER3(31)
Tid	14-18
Kurskod	732G05
Provkod	TENB
Kursnamn/benämning Provnamn/benämning	Regressions- och tidsserieanalys Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	4
Jour/Kursansvarig Ange vem som besöker salen	Annika Tillander
Telefon under skrivtiden	073-2129393
Besöker salen ca klockan	ca kl. 15
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Anna Garabska Eklund ankn. 2362 anna.grabska eklund@liu.se
Tillåtna hjälpmedel	Miniräknare Kursbok: Bowerman, O'Connel, Koehler: Forecasting, Time series, and Regression (alla upplagor tillåtna - får innehålla markeringar, understrykningar och flärpar, men inte anteckningar) A4-blad: med vad som helst skrivet på fram- och baksida (för hand/skrivet på dator/kopierat etc. - inga begränsningar).
Övrigt	
Antal exemplar i påsen	

Annika Tillander  
IDA/Statistik

## Tentamen Statistik, 732G05 2017-08-17

---

**Skrivtid:** 14.00-18.00

**Tillåtna hjälpmedel:**

**Miniräknare**

**Kursbok** Bowerman, O'Connel, Koehler: Forecasting, Time series, and Regression (alla upplagor tillåtna - får innehålla markeringar, understrykningar och flärpar, men inte anteckningar)

**A4-blad** med vad som helst skrivet på fram- och baksida (för hand/skrivet på dator/kopierat etc. - inga begränsningar).

**Betyg:** För godkänt betyg krävs 25 av 42 poäng. För väl godkänt betyg krävs 34 av 42 poäng.

Redovisa utförligt dina lösningar och tolka resultat.

Lösningförslag till denna tentamen läggs upp på kursens LISAM-sida den 18/8.

---

**Lycka till!**

### Uppgift 1 (7p)

Tabell 1. Visar priser och kvantitet för tre olika maskinkomponenter över en treårsperiod.

År	Komponent 1		Komponent 2		Komponent 3	
	$p$	$q$	$p$	$q$	$p$	$q$
2014	30	300	400	50	150	100
2015	35	200	600	100	180	150
2016	37	250	650	100	250	50

a) 4p

Beräkna ett index för de tre komponenterna i tabell 1 med 2014 som basår och Laspeyrevikter.

b) 1p

Hur stor har prisutvecklingen varit från 2014 till de två efterkommande åren (2015 och 2016)?

c) 2p

Vad är skillnaden mellan vikterna för Laspeyre och Paasche och när är Laspeyre att föredra framför Paasche?

## Uppgift 2 (6p)

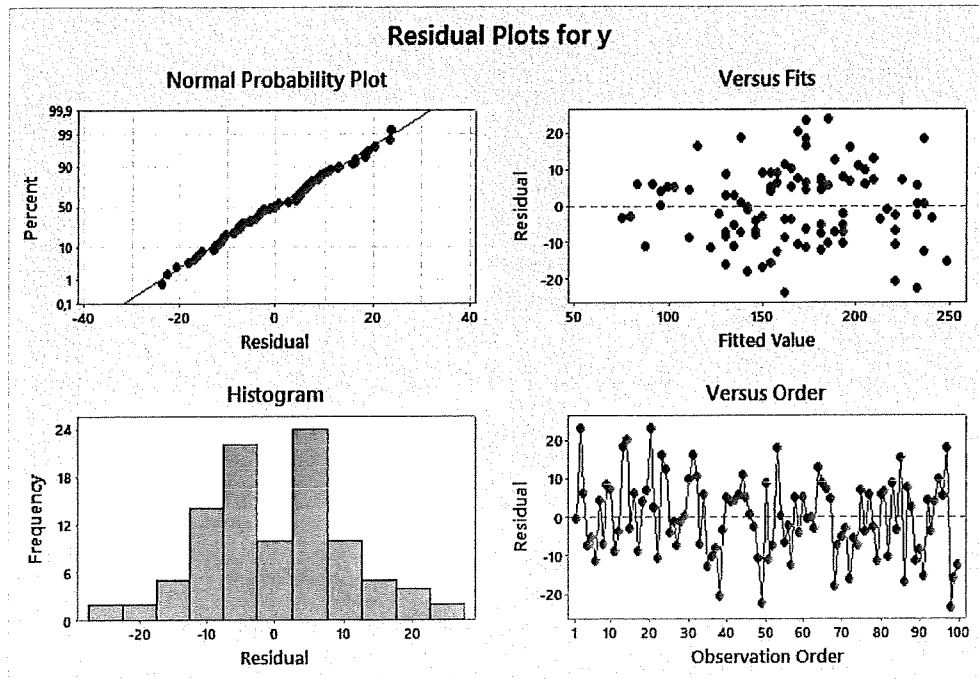
a)

2p

Den teoretiska modellen för multipel linjär regression  $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$ , förklara vad de olika beteckningarna ( $y, \beta_i, x_i, \varepsilon$ ) står för.

b)

4p



Figur 1. "Residual plots" från Minitab.

Givet figur 1:

1. Vad är egentligen "residual" i "Residual Plots for y" och vad motsvarar det i den teoretiska modellen?
2. Vad visar "Normal Probability Plot" och varför finns den med i "Residual Plots for y"?
3. Vad är egentligen "Fitted value" som finns med i "Versus fit"?
4. Vad visar "Versus order" och varför finns den med i "Residual Plots for y"?

### Uppgift 3 (17p)

Följande modell skattar inkomst för män i USA (data från 1978)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 D_1 + \beta_4 D_2 + \beta_5 D_3 + \beta_6 D_4 + \beta_7 D_3 * D_4 + \varepsilon$$

där  $y$  = inkomst i dollar,  $x_1$  = ålder (Age),  $x_2$  = utbildningslängd i år (Edu),  $D_1 = 1$  för afroamerikan (Black) annars 0,  $D_2 = 1$  för latinamerikan (Hispanic) annars 0,  $D_3 = 1$  för gift (Married) annars 0 och  $D_4 = 1$  ej fullgjort grundskola (Dropout) annars 0. Nedan visas delar från Minitabutskriften baserat på 15992 personer.

#### Analysis of Variance

Source	DF	SS	MS	F-Value
Regression	7	1,39762E+11	19965994226	<b>B</b>
Error	15984	1,34856E+12	84369057	
Total	<b>A</b>	1,48832E+12		

#### Model Summary

S	R-sq	R-sq(adj)
<b>C</b>	<b>D</b>	9,35%

#### Coefficients

Term	Coef	SE Coef	T-Value	P-Value
Constant	7380	591	12.49	0.000
Age	<b>E</b>	7.50	7.89	0.000
Edu	280.2	<b>F</b>	7.43	0.000
Black				
1	-2095	282	<b>G</b>	0.000
Hispanic				
1	<b>H</b>	288	-3.03	0.002
Married				
1	3948	205	19.22	0.000
Dropout				
1	-2807	324	-8.68	0.000
Married*Dropout				
1 1	1882	349	5.38	0.000

a)

4p

Fyll i de saknade värdena som är ersatta med bokstäver i fetstil från **A-H**. Redovisa dina beräkningar.

b)

5p

Testa, på 5% signifikansnivå, om modellen som helhet är signifikant.

c) 6p

Gör prediktion för inkomst med 95% konfidensintervall för gruppen afroamerikanska män med ålder 35 år, 12 års utbildningslängd och fullgjort grundskola. Detta dels för gifta (distance value = 0.0010) och dels för ogifta (distance value = 0.0012).

d) 2p

Regressionskoefficienten för gift är positiv medan regressionskoefficienten för ej fullgjort grundskola är negativ och sedan är interaktionstermen mellan dem positiv. Hur ska detta tolkas?

### Uppgift 4 (12p)

Nedanstående data visar antal sålda enheter (Q) och relativa priset (P) för 15 observationer av en viss produkt. Nedersta raden visar summan för de olika kolumnerna.

N	Q	P	Q <sup>2</sup>	P <sup>2</sup>	Q * P	log(Q)	log(P)	log(Q) <sup>2</sup>	log(P) <sup>2</sup>	log(Q) * P	Q * log(P)	log(Q) * log(P)
1	196	78	38416	6084	15288	2.3	1.9	5.3	3.6	178.8	370.9	4.3
2	110	107	12100	11449	11770	2.0	2.0	4.2	4.1	218.4	223.2	4.1
3	89	121	7921	14641	10769	1.9	2.1	3.8	4.3	235.9	185.4	4.1
4	146	92	21316	8464	13432	2.2	2.0	4.7	3.9	199.1	286.7	4.3
5	147	91	21609	8281	13377	2.2	2.0	4.7	3.8	197.2	288.0	4.2
6	108	109	11664	11881	11772	2.0	2.0	4.1	4.2	221.6	220.0	4.1
7	266	66	70756	4356	17556	2.4	1.8	5.9	3.3	160.0	484.0	4.4
8	301	62	90601	3844	18662	2.5	1.8	6.1	3.2	153.7	539.5	4.4
9	57	154	3249	23716	8778	1.8	2.2	3.1	4.8	270.4	124.7	3.8
10	79	130	6241	16900	10270	1.9	2.1	3.6	4.5	246.7	167.0	4.0
11	376	54	141376	2916	20304	2.6	1.7	6.6	3.0	139.1	651.4	4.5
12	137	95	18769	9025	13015	2.1	2.0	4.6	3.9	203.0	270.9	4.2
13	655	40	429025	1600	26200	2.8	1.6	7.9	2.6	112.6	1049.3	4.5
14	126	100	15876	10000	12600	2.1	2.0	4.4	4.0	210.0	252.0	4.2
15	112	107	12544	11449	11984	2.0	2.0	4.2	4.1	219.3	227.3	4.2

Σ 2905 1406 901463 144606 215777 32.883 29.219 73.185 57.257 2965.902 5340.352 63.445

a) 5p

Använd regressionsanalys för att skatta produktens priselasticitet, rita upp data på lämpligt vis med den skattade regressionslinjen (för den som inte vill utgå från summorna i ovanstående tabell finns tabellen också i större typsnitt på sidan 5).

b) 5p

Testa på 5% nivå om produkten är normalt priskänslig, dvs har priselasticiteten  $-1$ .

c) 2p

Använd din skattade modell för att ge en prognos av förändringen i antal sålda enheter då relativpriset ökar med c:a 3%.

N	Q	P	Q <sup>2</sup>	P <sup>2</sup>	Q * P	log(Q)	log(P)	log(Q) <sup>2</sup>	log(P) <sup>2</sup>	log(Q) * P	Q * log(P)	log(Q) * log(P)
1	196	78	38416	6084	15288	2.3	1.9	5.3	3.6	178.8	370.9	4.3
2	110	107	12100	11449	11770	2.0	2.0	4.2	4.1	218.4	223.2	4.1
3	89	121	7921	14641	10769	1.9	2.1	3.8	4.3	235.9	185.4	4.1
4	146	92	21316	8464	13432	2.2	2.0	4.7	3.9	199.1	286.7	4.3
5	147	91	21609	8281	13377	2.2	2.0	4.7	3.8	197.2	288.0	4.2
6	108	109	11664	11881	11772	2.0	2.0	4.1	4.2	221.6	220.0	4.1
7	266	66	70756	4356	17556	2.4	1.8	5.9	3.3	160.0	484.0	4.4
8	301	62	90601	3844	18662	2.5	1.8	6.1	3.2	153.7	539.5	4.4
9	57	154	3249	23716	8778	1.8	2.2	3.1	4.8	270.4	124.7	3.8
10	79	130	6241	16900	10270	1.9	2.1	3.6	4.5	246.7	167.0	4.0
11	376	54	141376	2916	20304	2.6	1.7	6.6	3.0	139.1	651.4	4.5
12	137	95	18769	9025	13015	2.1	2.0	4.6	3.9	203.0	270.9	4.2
13	655	40	429025	1600	26200	2.8	1.6	7.9	2.6	112.6	1049.3	4.5
14	126	100	15876	10000	12600	2.1	2.0	4.4	4.0	210.0	252.0	4.2
15	112	107	12544	11449	11984	2.0	2.0	4.2	4.1	219.3	227.3	4.2