

Försättsblad till skriftlig tentamen vid Linköpings universitet



Datum för tentamen	2018-06-08
Sal (2)	G33(1) <u>TER1(53)</u>
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	013-28 12 14
Besöker salen ca klockan	kl 10
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Anna Grabska Eklund 013-28 23 62 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 Regressions- och tidsserieanalys, 732G05/732G41
2018-06-08

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 da-
tor/kopierat etc. - inga begränsningar).

Betyg: För godkänt betyg krävs 24 av 40 poäng. För väl godkänt betyg krävs 32 av
40 poäng.

Redovisa utförligt dina lösningar och tolka resultat.

Lösningförslag till denna tentamen läggs upp på kursens LISAM-sida den 12/6.

Lycka till!

Uppgift 1 (6p)

Tabell 1. Försäljning (inkl. moms) av drycker inom handeln i löpande priser i miljoner kronor efter varugrupp och år

År	Drycker		KPI
	Alkoholfria	Alkoholhaltiga	
2013	20690	34707	314
2014	21653	36078	313
2015	22778	37415	313
2016	23614	38618	316

a) 2p

Beräkna om respektive pris för alkoholfria och alkoholhaltiga drycker till fastpris med 2013 som basår.

b) 2p

Beräkna ett enkelt fastbasindex för respektive dryck med 2013 som basår.

c) 2p

Beräkna ett relativprisindex med 2013 som basår för alkoholhaltiga jämfört mot alkoholfria drycker och tolka resultatet.

Uppgift 2 (6p)

a) 2p

Vilka är modellantagandena för linjär regression?

b) 4p

Visa hur modellantagandena kontrolleras grafiskt, dels hur det ser ut när antagandena är uppfyllda och dels när de inte är uppfyllda.

Uppgift 3 (15p)

Följande modell skattar Pineo-Porter yrkesprestige poäng, data kommer från Kanada 1971 där lägst poäng gällde för vaktmästare (17.3) och högst poäng för läkare (87.2).

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4D_1 + \beta_5D_2 + \beta_6D_1 \times x_2 + \beta_7D_2 \times x_2 + \varepsilon$$
 där y = Pineo-Porter yrkesprestige poäng, x_1 = median utbildningslängd i år, x_2 = medianinkomst i Kanadensiska dollar, x_3 = genomsnittlig andel kvinnor inom yrket i %, $D_1 = 1$ för yrkeskategori akademiker, $D_2 = 1$ för yrkeskategori arbetare och referensgrupp för yrkeskategori är kontor (t.ex. yrken inom administration och kundtjänst).
 Nedan visas delar från Minitabutskriften baserat på 101 yrken.

Regression Analysis: Prestige versus Utbildningslängd; Inkomst; Andel kvinnor; Akademiker; Arbetare

Analysis of Variance						
Source	DF	Adj SS	Adj MS	F-Value	P-Value	
Regression	7	25829	C	D	0,000	
Error	A	B	43,15			
Total	100	29842				

Model Summary			
S	R-sq	R-sq(adj)	
E	86,55%	85,54%	

Coefficients					
Term	Coef	SE Coef	T-Value	P-Value	
Constant	-9,05	7,10	-1,28	0,205	
Utbildningslängd	F	0,629	4,41	0,000	
Inkomst	0,003201	0,000698	4,59	0,000	
Andel kvinnor	0,0783	0,0290	2,70	0,008	
Akademiker	26,82	G	5,51	0,000	
Arbetare	0,82	4,78	H	0,863	
Inkomst*Akademiker	-0,002348	0,000700	-3,35	0,001	
Inkomst*Arbetare	0,000446	0,000806	0,55	0,582	

a)

4p

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

b)

6p

Gör prediktion för Pineo-Porter yrkesprestige poäng med 95% prediktionsintervall för ett yrke med 10 års median utbildningslängd, medianinkomst 10 000 Kanadensiska dollar och genomsnittlig andel kvinnor 10%. Detta för respektive yrkeskategori: Kontor (distance value = 0.323959), Akademiker (distance value = 0.1719672) och Arbetare (distance value = 0.1618114).

c)

5p

En reducerad modell utan interaktionstermer ger följande resultat:

Regression Analysis: Prestige versus Utbildningslängd; Inkomst; Andel kvinnor; Akademiker; Arbetare

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	24639	4927,70	89,96	0,000
Error	95	5204	54,77		
Total	100	29842			

Model Summary

S	R-sq	R-sq(adj)
7,40098	82,56%	81,65%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value
Constant	-6,72	7,17	-0,94	0,351
Utbildningslängd	3,777	0,672	5,62	0,000
Inkomst	0,001110	0,000268	4,15	0,000
Andel kvinnor	0,0111	0,0293	0,38	0,705
Akademiker	9,36	2,89	3,24	0,002
Arbetare	4,98	2,60	1,92	0,058

Testa på 5% signifikansnivå hur reducering av interaktionstermerna påverkar modellen.

Uppgift 4 (13p)

Följande kvartalsvisa omsättning per hundra tusen kronor rapporterade butiken SäljDet för de tre senaste åren:

År	2015				2016				2017			
Kvartal	1	2	3	4	1	2	3	4	1	2	3	4
Omsättning	3.6	4.4	4.5	10.6	3.8	4.6	4.7	11.2	4.2	4.9	4.8	11.8

Följanden kvadratsummor erhöles:

$$\sum_{i=1}^n \text{Tid}_i = 78, \sum_{i=1}^n \text{Oms}_i = 73.1, \sum_{i=1}^n \text{Tid}_i^2 = 650, \sum_{i=1}^n \text{Oms}_i^2 = 551.99 \text{ och}$$

$$\sum_{i=1}^n \text{Oms}_i \times \text{Tid}_i = 516.6$$

a)

5p

Skatta en enkel linjär regressionsmodell för omsättning över tid.

a)

2p

Beräkna ett lämpligt centrerat glidande medelvärde.

b)

3p

Gör modellantaganden och beräkna standardiserade säsongskomponenter.

c)

1p

Säsongrensa omsättningen.

d)

2p

Rita upp originaldata för omsättningen, regressionslinjen och säsongrensade omsättningen i samma graf.