



Försättsblad till skriftlig tentamen vid Linköpings Universitet

Datum för tentamen	2013-08-16
Sal (1) Om tentan går i flera salar ska du bifoga ett försättsblad till varje sal och <u>ringa in</u> vilken sal som avses	G34
Tid	8-12
Kurskod	732G25
Provkod	TEN1
Kursnamn/benämning Provnamn/benämning	Statistisk analys av samhällsdata Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	5
Jour/Kursansvarig Ange vem som besöker salen	Tommy Schyman
Telefon under skrivtiden	
Besöker salen ca kl.	10
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	annelie.almquist@liu.se
Tillåtna hjälpmedel	Räknedosa. Bowerman, B.J., O'Connell, R, Koehler, A.: <i>Forecasting, Time Series and Regression</i> , alla upplagor tillåtna, som inte får innehålla anteckningar men får ha markeringar och flärpar. Liten anteckning på flärp tillåten.
Övrigt	G=12, VG=16
Vilken typ av papper ska användas, rutigt eller linjerat	
Antal exemplar i påsen	

TENTAMEN I STATISTISK ANALYS AV SAMHÄLLSDATA, 2013-08-16

Skrivtid: kl: 8-12

Hjälpmedel: Räknedosa. Bowerman, B.J., O'Connell, R, Koehler, A.: *Forecasting, Time Series and Regression*, alla upplagor tillåtna, som inte får innehålla anteckningar men får ha markeringar och flärpar. Liten anteckning på flärp tillåten.

Jourhavande lärare: Tommy Schyman

Redovisa och motivera kort alla dina lösningar

Tolka (om möjligt) alla dina resultat!

1

Tabellen nedan är hämtad från SCBs hemsida. (Något modifierad)

392 Priset på mjölk och ägg 2000–2005 i löpande priser och i 2005 års priser

Priset på 1 liter mjölk			Kilopriset på ägg		
År	Löpande priser	2005 års priser	År	Löpande priser	2005 års priser
2000	6,40	6,88	2000	18,80	20,22
2001	6,60	6,93	2001	18,40	19,32
2002	6,90	7,09	2002	19,90	20,45
2003	7,30	7,36	2003	21,80	21,98
2004	7,20	7,23	2004	23,10	23,21
2005	7,20	7,20	2005	21,50	21,50

Beräkna implicitprisindex dels för mjölk och dels för ägg med år 2000 som basår. Hur stor prisförändring har det varit mellan år 2000 och år 2005 för de båda varorna? 2p

2

I tabellen nedan visas viss data över befolkningen i Linköping år 2005.

Åldersintervall	Totala antalet kvinnor i intervallet	Totala antalet levande födda barn i intervallet
15-19	4217	7
20-24	5799	75
25-29	5052	214
30-34	4552	336
35-39	4462	157
40-44	4438	22

Inga barn föddes av kvinnor äldre än 44 år. Totalt fanns det 69005 män och 68270 kvinnor.

- Beräkna det allmänna födelsetalet. 1p
- Beräkna den generella fruktsamhetsknoten 1p
- Beräkna den summerade fruktsamhetsknoten och tolka den (TFR). 2p
- Beräkna bruttoreproduktionen. 1p

3

Betrakta följande tidsserie med kvartalsdata från tre år.

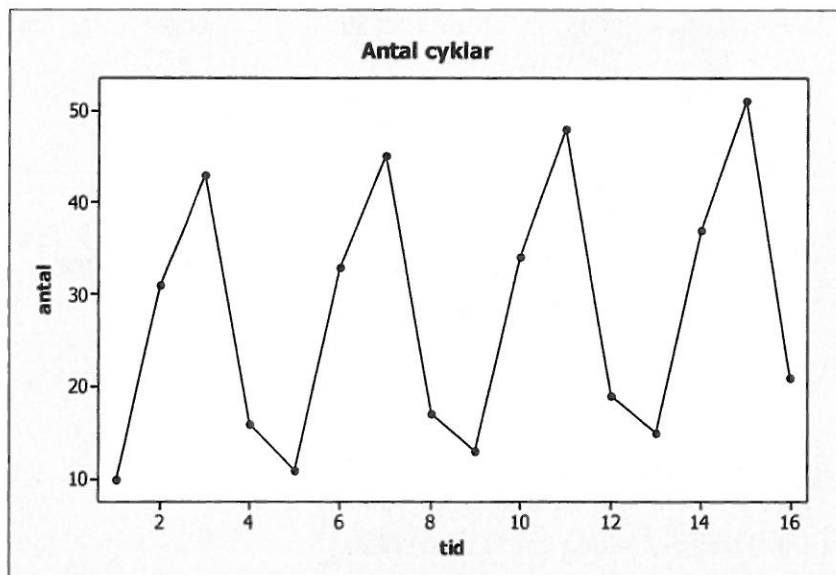
30, 26, 24, 50, 36, 28, 32, 52, 38, 28, 30, 54

Skatta de fyra säsongeffekterna med additiv metod (klassisk komponentuppdelning). Tolka också de fyra säsongkomponenterna. 3p

4

I figuren nedan visas försäljningen av antal cyklar i en viss affär. Det är kvartalsdata över fyra år. Två modeller är anpassade. $Y =$ antal cyklar.

- Prediktera antal sålda cyklar för första kvartalet år fem med båda modellerna. 2p
- Tolka för båda modellerna den skattade säsongeffekten för kvartal 3. 2p
- Kan residualerna från modell 1 anses vara autokorrelerade? Kan modell 1 anses vara en bra modell? Förklara och motivera väl. 1p



Modell 1

Regression Analysis: y versus tid; kv_1; kv_2; kv_3

The regression equation is

$$y = 13,3 + 0,500 \text{ tid} - 4,50 \text{ kv}_1 + 16,5 \text{ kv}_2 + 29,0 \text{ kv}_3$$

Predictor	Coef	SE Coef	T	P
Constant	13,2500	0,5056	26,20	0,000
tid	0,50000	0,03769	13,27	0,000
kv_1	-4,5000	0,4900	-9,18	0,000
kv_2	16,5000	0,4827	34,19	0,000
kv_3	29,0000	0,4782	60,64	0,000

S = 0,674200 R-Sq = 99,8% R-Sq(adj) = 99,8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	2990,00	747,50	1644,50	0,000
Residual Error	11	5,00	0,45		
Total	15	2995,00			

Durbin-Watson statistic = 2,2

Modell 2 Time Series Decomposition for y

Additive Model

Data y
Length 16
NMissing 0

Fitted Trend Equation

$$Y_t = 23,531 + 0,496 * t$$

Seasonal Indices

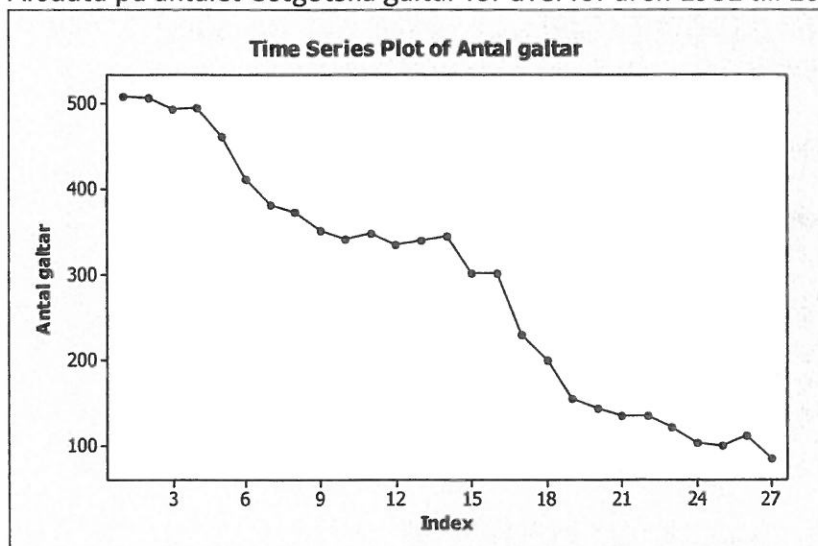
Period	Index
1	-14,8438
2	6,4063
3	18,4063
4	-9,9688

Accuracy Measures

MAPE 2,04255
MAD 0,46967
MSD 0,36983

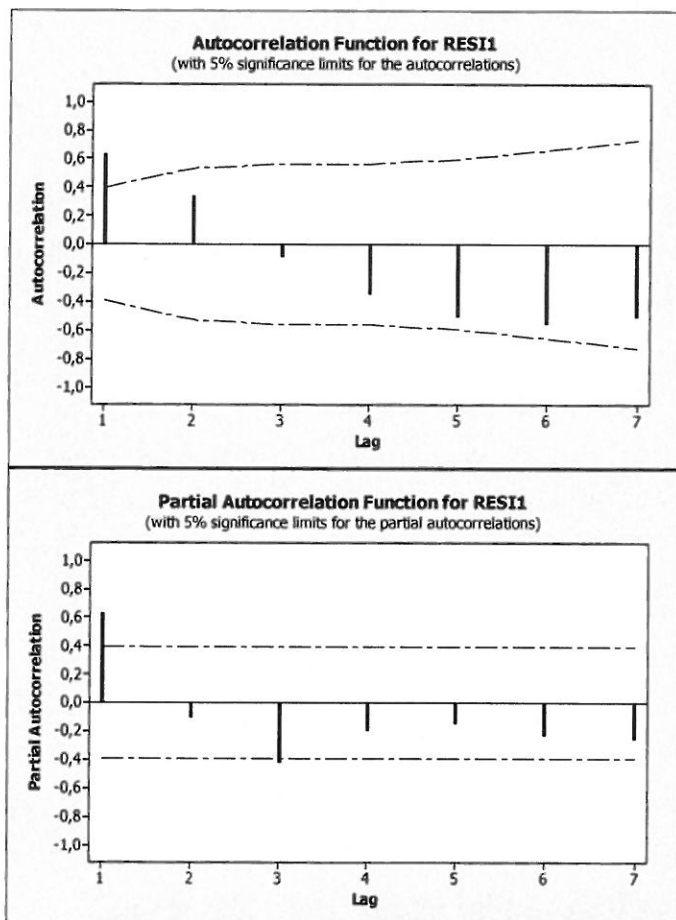
5

Årsdata på antalet Östgötska galtar för avel för åren 1981 till 2007 visas i följande graf.



För att prognostisera antalet galtar = y_t år 2008 och 2009 har Dubbel exponentiell utjämning utförts i Minitab med utjämningkonstanterna $\alpha = 0,4$ och $\gamma = 0,05$. Serien har totalt 27 värden och $l_{26} = 85,16$ och $b_{26} = -17,07$. De sista observationerna är $y_{25} = 98, y_{26} = 111, y_{27} = 84$.

- a) Beräkna prognoser för 2008 = (tidpunkt 28) och 2009 då du står vid 2007 = (tidpunkt 27). 2p
- b) Nedan visas SAC och SPAC för residualerna från den dubbla exponentiella utjämnningen. Vilken ARMA-modell föreslår dessa för residualerna? 1p



- c) Nedan har en enkel linjär regressionsmodell anpassats till antal avelsgaltar. Vid vilket år har antalet avelsgaltar nått noll i Östergötland enligt modellen? Tidsvariabeln är kodad 1 till 27. 1p
- d) Pröva med hjälp av utskriften nedan om residualerna är autokorrelerade. 1p

Regression Analysis: Antal galtar versus tid

The regression equation is
Antal galtar = 535 - 17,6 tid

Predictor	Coef	SE Coef	T	P
Constant	535,25	10,96	48,85	0,000
tid	-17,6392	0,6840	-25,79	0,000

S = 27,6813 R-Sq = 96,4% R-Sq(adj) = 96,2%

Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	1	509649	509649	665,12	0,000
Residual Error	25	19156	766		
Total	26	528806			

Durbin-Watson statistic = 0,543939

Formelsamling Index

Sammansatta fastbasindex:

$$I_t = i_{1,t} \cdot w_1 + i_{2,t} \cdot w_2 + \dots + i_{n,t} \cdot w_n$$

där n är antalet ingående varor/tjänster, $i_{1,t}, \dots, i_{n,t}$ är enkla prisindex för ingående varor, alla med basår t_0 och w_1, \dots, w_n väljs enligt ett viktsystem:

$$\text{Laspeyre: } w_i = \frac{p_{i,t_0} \cdot q_{i,t_0}}{\sum_j p_{j,t_0} \cdot q_{j,t_0}}$$

$$\text{Paasche: } w_i = \frac{p_{i,t_0} \cdot q_{i,t}}{\sum_j p_{j,t_0} \cdot q_{j,t}}$$

Kedjeprisindex:

$$I_t = \frac{L_{0,1}}{100} \cdot \frac{L_{1,2}}{100} \cdot \dots \cdot \frac{L_{t-1,t}}{100} \cdot 100$$

där

$$L_{t-1,t} = \sum_{i=1}^n \frac{p_{i,t}}{p_{i,t-1}} \cdot 100 \cdot w_{i,t-1,t}$$

är årslänken från år $t-1$ till t för n ingående varor/tjänster. $w_{i,t-1,t}$ väljs enligt ett viktsystem:

$$\text{Laspeyre: } w_{i,t-1,t}^L = \frac{\text{Försäljningsvärdet för vara } i \text{ år } t-1}{\text{Totala försäljningsvärdet år } t-1}$$

$$\text{Paasche: } w_{i,t-1,t}^P = \frac{\text{Försäljningsvärdet för vara } i \text{ år } t \text{ i priser för år } t-1}{\text{Totala försäljningsvärdet år } t \text{ i priser för år } t-1}$$

Med representantvaror byts "Försäljningsvärdet för vara i " mot "Försäljningsvärdet för varugrupp i " i vikterna.

Implicitprisindex:

$$I_t = \frac{\text{Försäljningsvärdet av varan/tjänsten/gruppen år } t \text{ i löpande priser}}{\text{Försäljningsvärdet av varan/tjänsten/gruppen år } t \text{ i basårets priser}} \cdot 100$$

Relativprisindex:

$$I_t^R = \frac{I_t^v}{I_t^0} \cdot 100$$

där I_t^v = Prisindex för aktuell vara/tjänst/grupp och I_t^0 = Prisindex för den större jämförelsegruppen, t ex KPI.