



Försättsblad till skriftlig tentamen vid Linköpings Universitet

Datum för tentamen	2010-08-25
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	TER2
Tid	8-12
Kurskod	732G25
Provkod	TENT
Kursnamn/benämning Provnamn/benämning	Statistisk analys av samhällsdata Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	4
Jour/Kursansvarig Ange vem som besöker salen	Lotta Hallberg
Telefon under skrivtiden	
Besöker salen ca kl.	10
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Carita Lilja 1463 carita.lilja@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.
Övrigt	G=12, VG=16
Vilken typ av papper ska användas, rutigt eller linjerat	
Antal exemplar i påsen	

Tentamen i Statistisk analys av samhällsdata, 2010-08-25

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.

Jourhavande lärare: Lotta Hallberg

Redovisa och motivera kort alla dina lösningar
Data i denna tenta är hämtade från SCB

1

Du har tillgång till följande data om befolkningen i Bangladesh 1974.

Åldersgrupp	Män i tusental	Kvinnor i tusental	Födslar i tusental
0-4	7457	7192	--
5-9	5820	5625	--
10-14	4730	4533	--
15-19	4017	3777	756.9
20-24	3362	3101	1046.0
25-29	2891	2636	819.5
30-34	2391	2161	565.1
35-39	1983	1793	353.2
40-44	1635	1484	141.6
45-49	1343	1222	16.5
50-54	1066	974	--
55-59	858	784	--
60-64	658	601	--
65-69	487	440	--
70-74	339	304	--
75+	399	334	--

- Bestäm allmänna födelsetalet (Crude birth rate) CBR 1p
- Bestäm den generella fruktsamhetskvoten GBR 1p
- Bestäm den åldersspecifika fruktsamheten för gruppen 20-29år 1p
- Bestäm den summerade fruktsamheten (Total fertility rate) TFR. 2p
- Hur många barn förväntas en kvinna föda i Bangladesh 1974? 1p
- Förklara hur man går tillväga om man istället vill följa en kohort. 1p

2

Med jämna mellanrum mättes antal smittade av en viss sjukdom vid 11 tillfällen: Resultat 1, 1, 2, 3, 4, 6, 8, 13, 21, 27, 45. Som prediktionsmodell för dessa data är tillväxtmodellen (exponentiella modellen)

$y = \beta_0 (\beta_1)^t \delta$ en bra modell. $y =$ antal smittade och $t = 1, 2, \dots, 11$. δ är

slumpkomponenten. Skatta och tolka β_1 .

3p

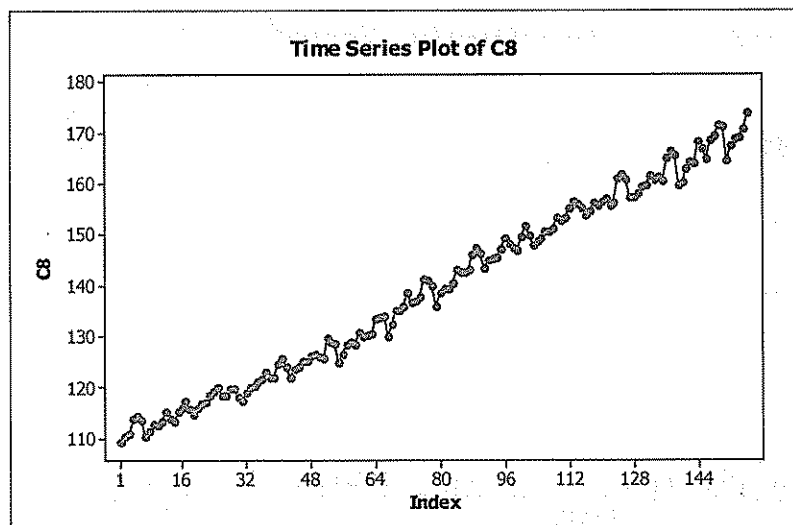
3

Följande data är hämtade från SCB hemsida.

Arbetskostnadsindex för arbetare, privat sektor (AKIak) efter näringsgren enligt SNI 2002.

Månad 1996M01-2008M12. Ger tidpunkt 1 till 156

Handel, reparation av fordon, hotell och restauranger



4 modeller är anpassade i Minitab och tidsserien ligger i C8.
Uppgifterna kommer efter alla utskrifter

Modell 1

Time Series Decomposition for C8

Multiplicative Model

Data C8
Length 156
NMissing 0

Fitted Trend Equation

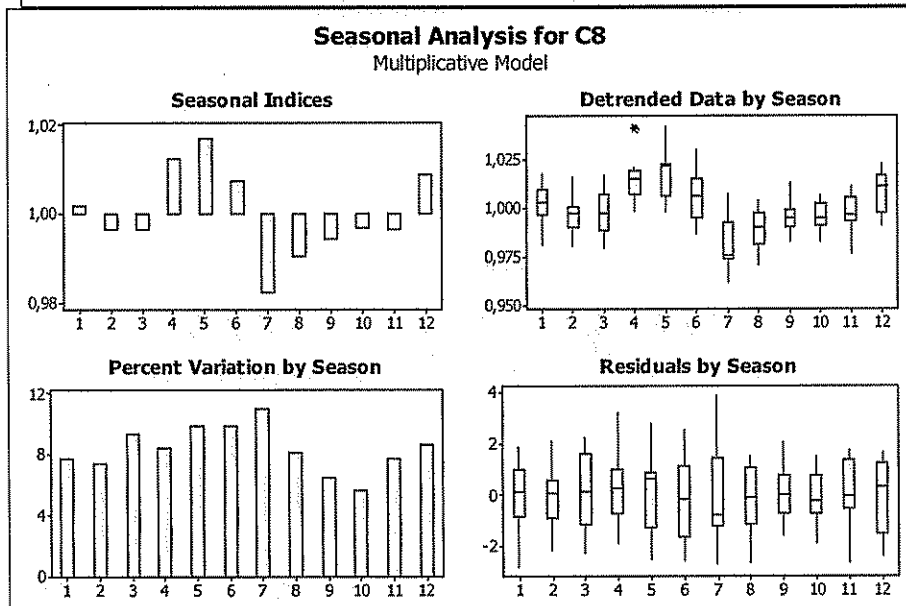
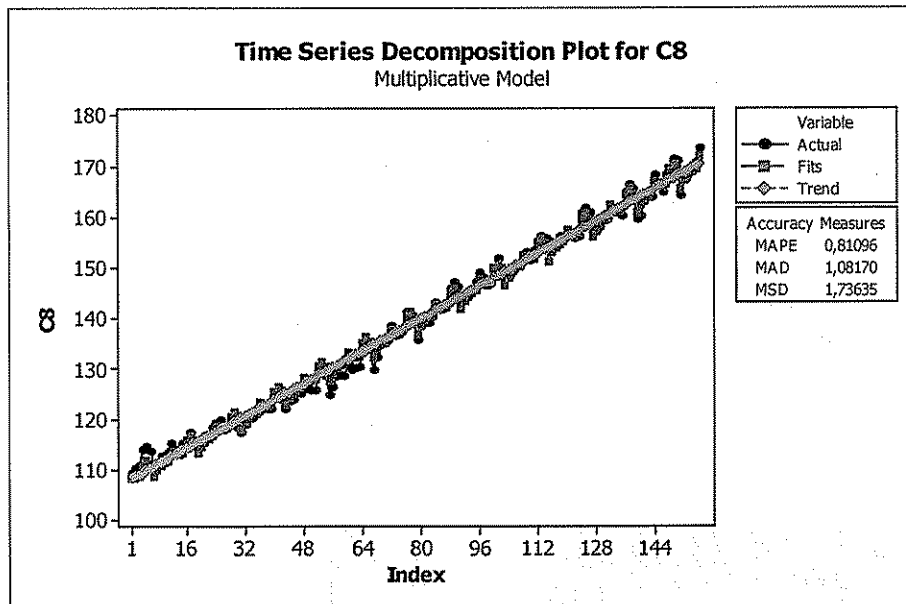
$$Y_t = 107,764 + 0,402 * t$$

Seasonal Indices

Period	Index
1	1,00187
2	0,99643
3	0,99648
4	1,01247
5	1,01710
6	1,00748
7	0,98201
8	0,99012
9	0,99436
10	0,99661
11	0,99634
12	1,00873

Accuracy Measures

MAPE 0,81096
MAD 1,08170
MSD 1,73635



Modell 2

Trend Analysis for C8

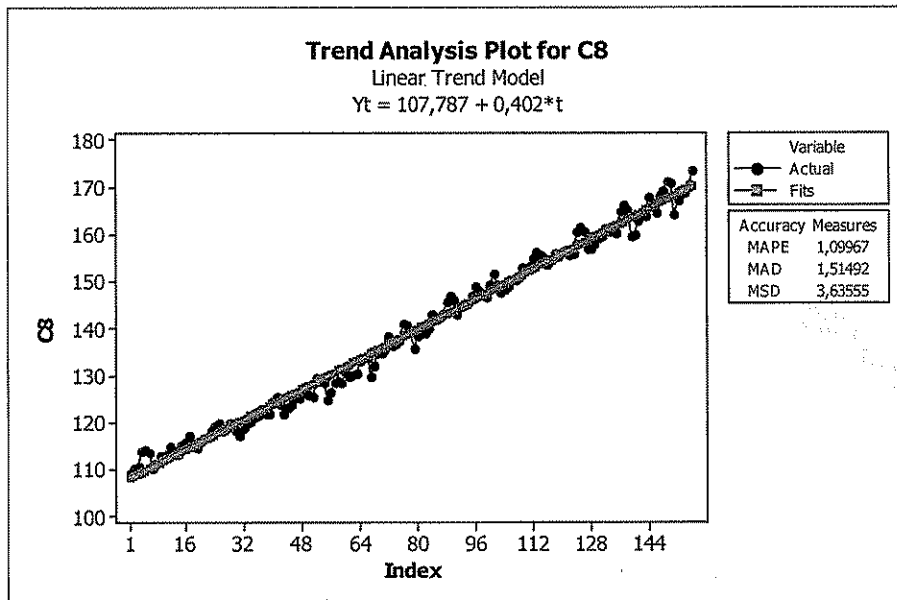
Data C8
 Length 156
 NMissing 0

Fitted Trend Equation

$$Y_t = 107,787 + 0,402 * t$$

Accuracy Measures

MAPE 1,09967
 MAD 1,51492
 MSD 3,63555



Modell 3 Winters' Method for C8

Multiplicative Method

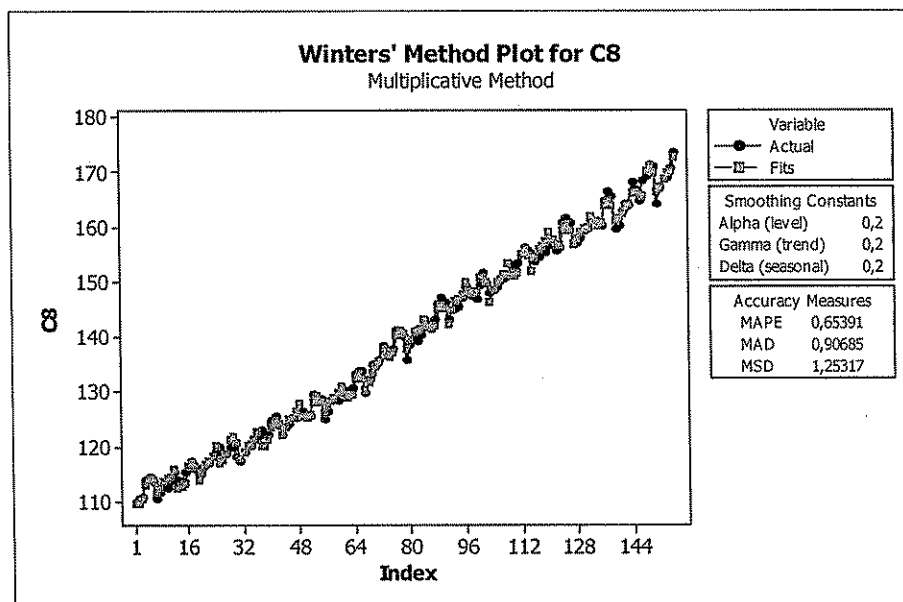
Data C8
Length 156

Smoothing Constants

Alpha (level) 0,2
Gamma (trend) 0,2
Delta (seasonal) 0,2

Accuracy Measures

MAPE 0,65391
MAD 0,90685
MSD 1,25317



Modell 4 Double Exponential Smoothing for C8

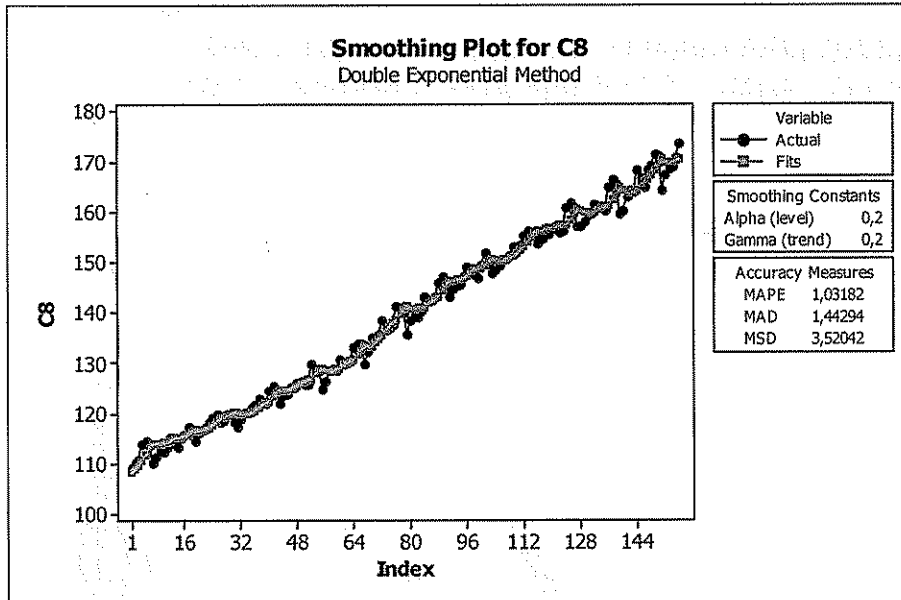
Data C8
Length 156

Smoothing Constants

Alpha (level) 0,2
Gamma (trend) 0,2

Accuracy Measures

MAPE 1,03182
MAD 1,44294
MSD 3,52042



Här är data för det sista året:

2008M01	166,7
2008M02	164,7
2008M03	168,4
2008M04	169,2
2008M05	171,4
2008M06	171,1
2008M07	164,3
2008M08	167,4
2008M09	168,7
2008M10	168,9
2008M11	170,7
2008M12	173,7

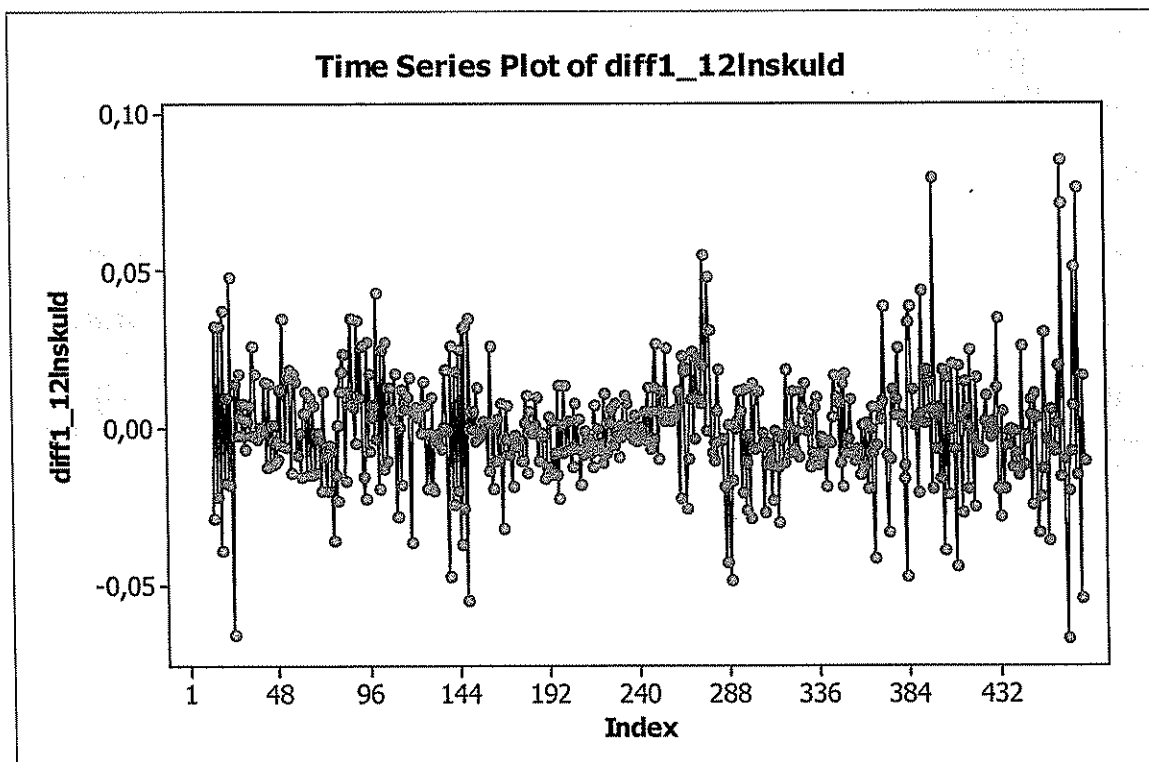
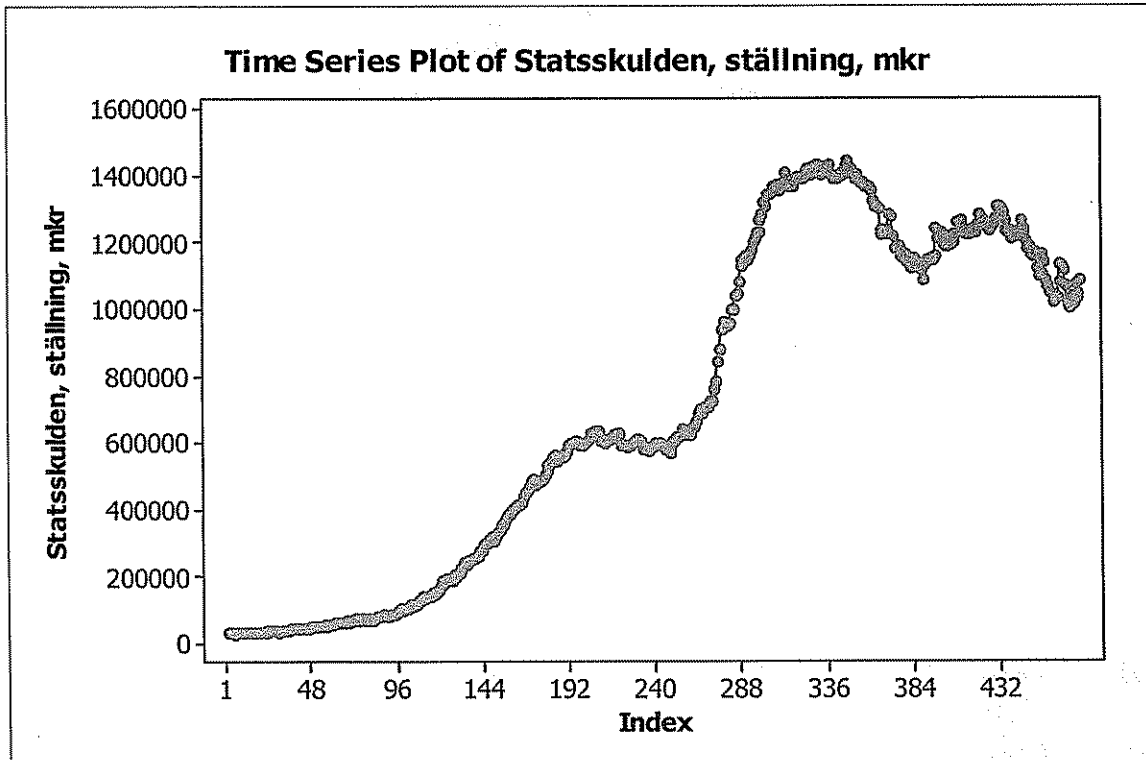
- Vilka är för- resp nackdelar med modell av typ 1 och 2 gentemot modell 3 och 4? 2p
- Vilken av de fyra modellerna väljer du? Motivera väl varför du vill/vill inte ha med säsongskomponent/effekt. 1p
- Det finns ytterligare modeller som inkluderar säsongseffekt som skulle kunna anpassas här. Beskriv minst en av dessa modeller. 1p
- Beräkna två prognoser för december 2008 med hjälp av modell 1 och modell 2. Jämför med de sanna värdena. Vilken prognos anser du vara den mest trovärdiga? 2p

4

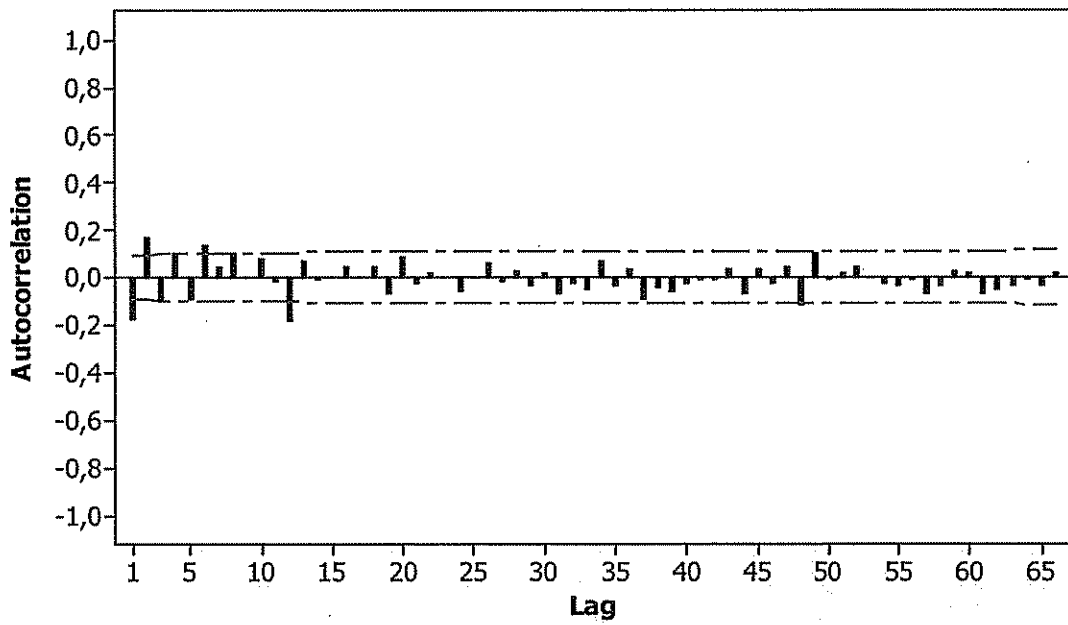
I grafen nedan visas en tidsserie över statsskulden i Sverige, januari 1970 till september 2009.

För att kunna anpassa en ARIMA modell så har serien först logaritmerats, sedan har den differentierats för trend och säsong. Denna serie heter nedan `diff1_12lnskuld`.

- Förklara varför man har logaritmerat och differentierat på detta sätt. 2p
- Föreslå en vettig SARIMA modell med hjälp av SAC och SPAC nedan. Det räcker att du ger värden på $(p,d,q)(P,D,Q)$. Förklara varför du väljer just dessa värden. 2p



Autocorrelation Function for diff1_12Inskuld
(with 5% significance limits for the autocorrelations)



Partial Autocorrelation Function for diff1_12Inskuld
(with 5% significance limits for the partial autocorrelations)

