

TENTAMEN I STATISTISK ANALYS AV SAMHÄLLSDATA, 2017-06-09

Skrivtid: kl: 8-12
Hjälpmedel: Räknedosa. Ett handskrivet A4 blad med egna anteckningar (båda sidor).
Jourhavande lärare: Lotta Hallberg
Betyg: För godkänt krävs minst 12 av 20 poäng och för väl godkänt krävs minst 16 av 20 poäng.

Redovisa och motivera kort alla dina lösningar

Tolka (om möjligt) alla dina resultat!

Följande data har hämtats från SCBs hemsida.

Genomsnittlig månadslön för månadsavlönade, statlig sektor (KLS) polisväsendet heltidsanställning
Månad 2010M01 - 2017M02

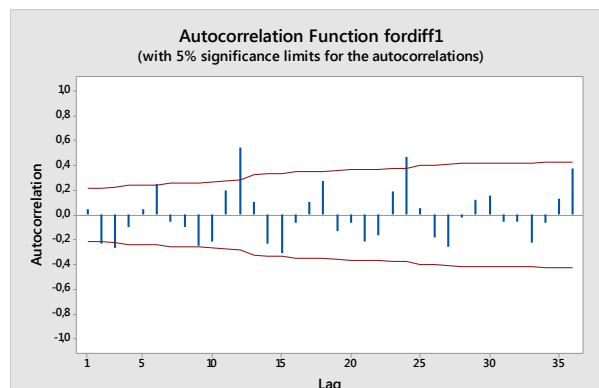
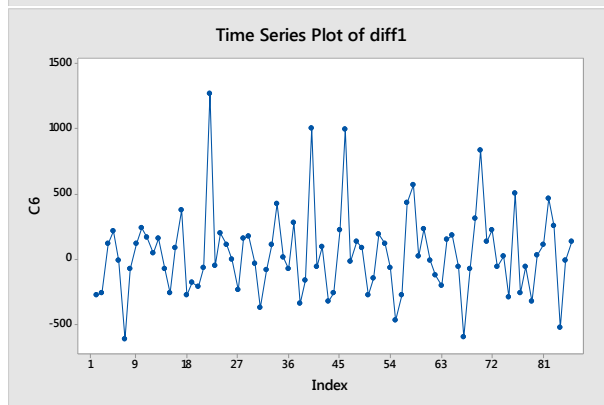
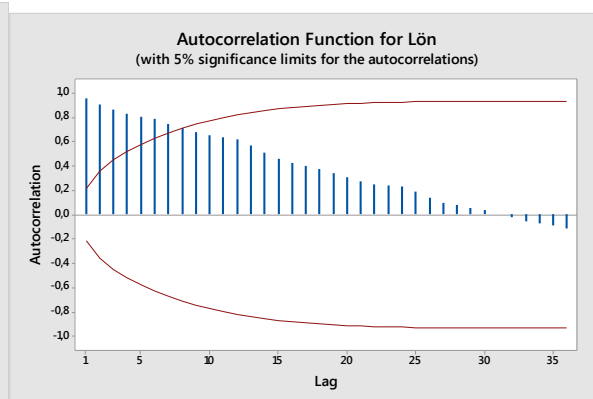
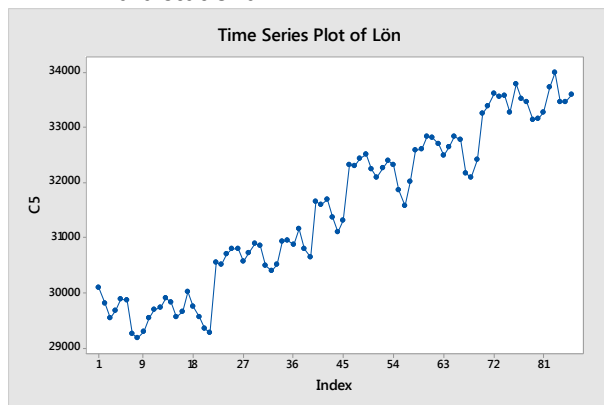
Dessa data kommer att användas i hela tentamen.

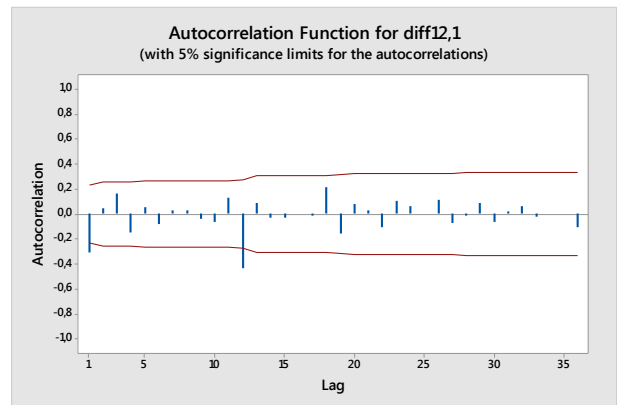
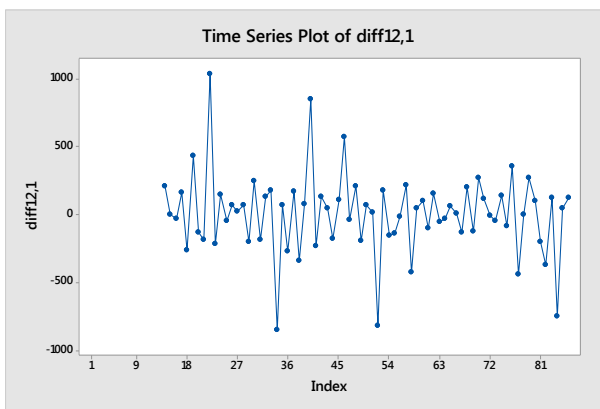
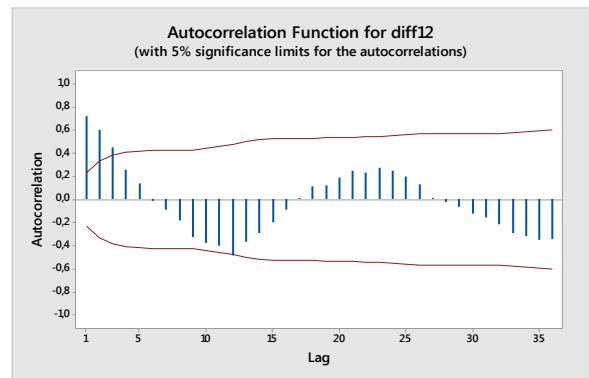
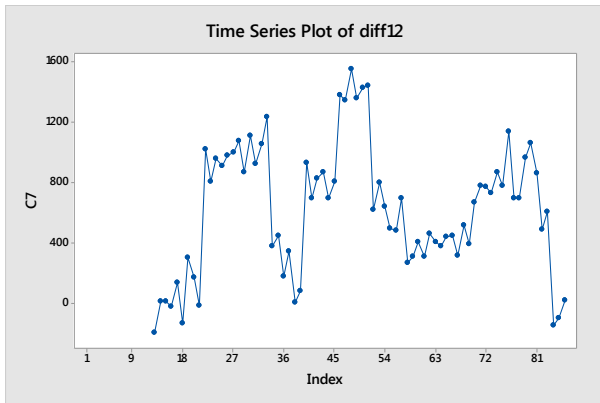
Denna genomsnittslön kommer att kallas antingen lön eller löner i uppgifterna nedan. Diff1 är lön differentierad för trend, diff12 är lön differentierad för säsong och diff12,1 är lön differentierad för båda.

1

En SARIMA-modell ska anpassas. För att kunna göra det så måste först tidsserien göras stationär.

- a) Det är tre kriterier som ska vara uppfyllda för att man ska kunna kalla en tidserie stationär. Vilka är dessa tre kriterier? 1,5p
- b) Nedan visas några grafer. Ta hjälp av dessa och tala om vilken av tidsserierna som du anser vara stationär. 1,5p





2

Följande modell har anpassats till Lön

ARIMA Model: Lön

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
SAR 12	0,6908	0,0891	7,75	0,000
Constant	7,00	27,85	0,25	0,802

Differencing: 1 regular difference

Number of observations: Original series 86, after differencing 85

Residuals: SS = 5453106 (backforecasts excluded)
MS = 65700 DF = 83

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	18,0	34,0	37,5	56,2
DF	10	22	34	46
P-Value	0,055	0,049	0,312	0,143

- a) Ange p,d,q och P,D,Q för den säsong ARIMA som är anpassad ovan. 1p

De sista 13 värdena i tidsserien lön är givna nedan

År,månad	Lön
2016M02	33563
2016M03	33270
2016M04	33779
2016M05	33521
2016M06	33463
2016M07	33134
2016M08	33161
2016M09	33268
2016M10	33731
2016M11	33986
2016M12	33461
2017M01	33449
2017M02	33587

- b) Beräkna en prognos för genomsnittslönen för en person anställd inom polisväsendet för mars och april 2017 med hjälp av modellen ovan. 3p
- c) Även en modell med klassisk komponentuppdelning har anpassats. Se utskrift nedan. Beräkna med denna modell prognoser för genomsnittslönen för en person anställd inom polisväsendet för mars och april 2017. 1p
- d) Tolka säsongskomponenten för december. 1p

Time Series Decomposition for Lön

Additive Model

Data Lön
Length 86
NMissing 0

Fitted Trend Equation
 $Y_t = 29119,8 + 54,98 \times t$

Seasonal Indices
Period Index
1 321,755
2 162,484
3 -96,432
4 51,672
5 167,443
6 43,464
7 -373,891
8 -612,057
9 -431,682
10 204,901
11 204,276
12 358,068

Accuracy Measures
MAPE 0,7
MAD 210,8
MSD 73304,7

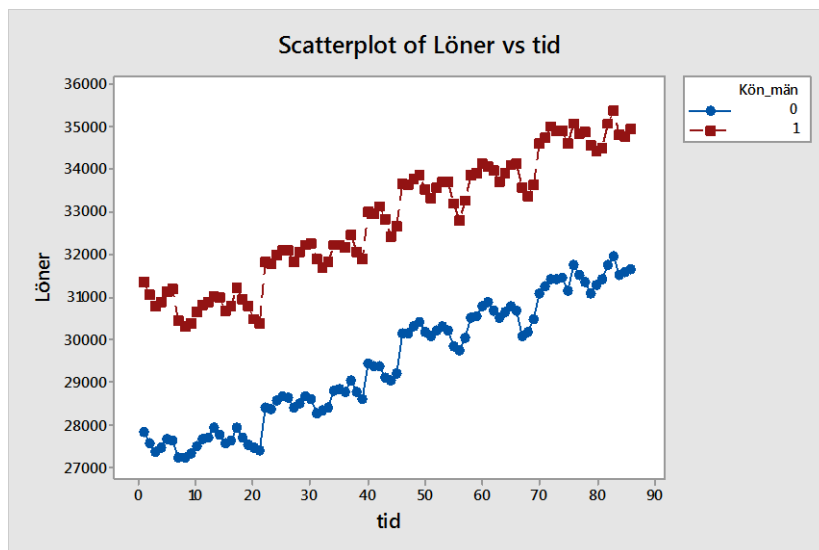
3

Använd nu observationerna på lön som gavs i uppgift 2

- Räkna om Lön till fastbasindex från månad 2016M10 till månad 2017M02 men där basmånad är 2016M08, dvs index för 2016M08 ska vara 100. 2p
- Om man vill jämföra hur den genomsnittliga månadslönen för polisväsendet ligger i förhållande till den allmänna prisutvecklingen i samhället. Vilken indexserie ska man då jämföra med? Och hur gör man en sådan jämförelse? Vad kallas en sådan jämförande indexserie? 2p

4

Nu heter variabel Lön=Löner. Den är nu uppdelad på kön. De röda prickarna är lön för män. Se graf nedan.



Följande modell har anpassats.

Regression Analysis: Löner versus tid; Kön_män

Analysis of Variance

Source	DF	Adj SS	Adj MS
Regression	2	826456922	413228461
Error	169	24301569	143796
Total	171	850758492	

Model Summary

S	R-sq	R-sq (adj)	R-sq (pred)
379,205	97,14%	97,11%	97,04%

Coefficients

Term	Coef	SE Coef
Constant	26962,5	65,1
tid	56,84	1,16
Kön_män	3355,2	57,8

Regression Equation

$$\text{Löner} = 26962,5 + 56,84 \text{ tid} + 3355,2 \text{ Kön}_m\text{än}$$

Till uppgifterna nedan ska du använda signifikansnivå 5 %

- Sätt upp korrekta hypoteser och pröva om vi har signifikant regression. 2p
- Testa om det är signifikant skillnad i lön mellan män och kvinnor. Hur stor är den skattade skillnaden? 3p
- Beräkna ett 95% prediktionsintervall för predikterad lön för män mars 2017 med hjälp av modellen ovan. Du får använda direkt att medelfelet för prediktionen är 65,1. 2p