

TENTAMEN I STATISTISK ANALYS AV SAMHÄLLSDATA, 2017-10-09

- Skrivtid:** kl: 14-18
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

1

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.

- a) Beräkna det allmänna födelsetalet. 0.5p
- b) Beräkna den generella fruktsamhetskvoten 0.5p
- c) Beräkna den summerade fruktsamhetskvoten och tolka den. 2p
- d) Beräkna bruttoreproduktionen. 1p

2

Nedan visas konsumtion i löpande och fast pris för varugruppen livsmedel, tobak och drycker:

År	Värde i löpande pris	Värde i fast pris	KPI
1990	144124	150666	207,6
1991	151420	151420	227,2
1992	146762	153781	232,6
1993	152194	154032	243,6

- a) Bilda ett implicitprisindex för de fyra åren och förklara vad detta index beskriver. Basår 1990. 2p
- b) Hur kan man från en given indexserie (som exempelvis den i a-uppgiften) bilda ett relativprisindex med hjälp av KPI. Vad beskriver en sådan serie? 1p
- c) Hur stor förändring (i %) har det varit i den allmänna prisutvecklingen i Sverige mellan år 1991 och 1993? 1p

3

Anta att kvaliteten på en produkt ska mätas med ett visst mått, Score2 men att detta mått är mycket dyrt att ta fram. Det finns då ett annat mått som är betydligt billigare Score1. Score2 kan uppskattas med Score1 via en enkel linjär regressionsanalys.

Regression Analysis: Score2 versus Score1

The regression equation is
 $\text{Score2} = 1.12 + 0.218 \text{ Score1}$

Predictor	Coef	SE Coef
Constant	1.1177	0.1093
Score1	0.21767	0.01740

S = 0.127419 R-Sq = 95.7% R-Sq(adj) = 95.1%

Analysis of Variance

Source	DF	SS	MS
Regression	1	2.5419	2.5419
Residual Error	7	0.1136	0.0162
Total	8	2.6556	

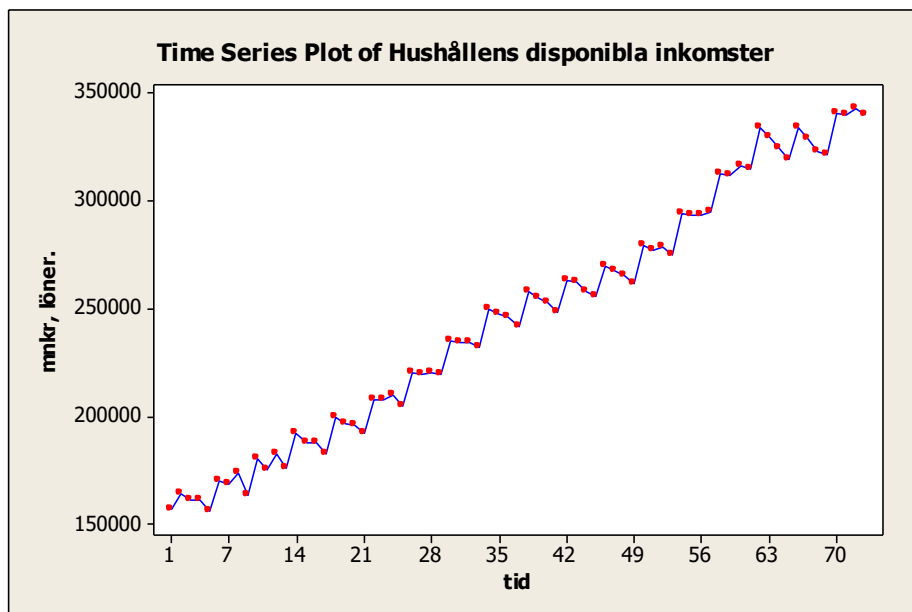
- Pröva på 5% signifikansnivå om Score1 är en signifikant förklarande variabel för Score2. 1p
- Hur mycket av variationen i Score2 förklaras av Score1. 1p
- Prediktera storleken på Score2 då Score1 är 5. 1p

4

Utjämningsserien 116 108 95 125 133 110 150 108 med ett 3 punkters centrerat glidande medelvärde. 2p

5

Nedan visas en tidsserie över Hushållens disponibla inkomster (ENS95), mnkr, löner. Kvartalsdata 1993K1 till 2011K1



Nedan har tre SARIMA-modeller anpassats.

- Jämför de tre modellerna noggrant. Vilken anser du vara den bästa modellen? Motivera. 2p
- Låt y_t vara den disponibla inkomsten. Låt vidare z_t vara diff4_1, dvs z_t är resultatet av att y_t är differentierad först med en tidsförskjutning och sedan med 4 tidsförskjutningar (lag). Uttryck z_t i y_t . 2p
- Modell 2 är den enklaste av de tre SARIMA-modellerna därför ska vi titta närmare på den. Sist i tentan visas vissa data från slutet på tidsserien. Även prognoser som Minitab har beräknat. Visa hur prognoserna är beräknade. 3p

MODELL 1

ARIMA Model: Hushållens disponibla inkomster

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
SAR 4	0,2287	0,1662	1,38	0,173
SMA 4	0,8707	0,1084	8,03	0,000
Constant	81,09	48,85	1,66	0,102

Differencing: 1 regular, 1 seasonal of order 4

Number of observations: Original series 73, after differencing 68

Residuals: SS = 449299813 (backforecasts excluded)
MS = 6912305 DF = 65

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

Lag	12	24	36	48
Chi-Square	14,4	22,3	38,8	44,2
DF	9	21	33	45
P-Value	0,107	0,382	0,224	0,504

MODELL 2

ARIMA Model: Hushållens disponibla inkomster

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
SMA 4	0,8050	0,0736	10,94	0,000
Constant	91,30	74,71	1,22	0,226

Differencing: 1 regular, 1 seasonal of order 4

Number of observations: Original series 73, after differencing 68

Residuals: SS = 458964898 (backforecasts excluded)
MS = 6954014 DF = 66

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

Lag	12	24	36	48
Chi-Square	17,0	27,6	44,8	50,6
DF	10	22	34	46
P-Value	0,075	0,188	0,102	0,296

MODELL 3

ARIMA Model: Hushållens disponibla inkomster

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
MA 1	-0,0119	0,1241	-0,10	0,924
SMA 4	0,8022	0,0743	10,80	0,000
Constant	92,01	76,32	1,21	0,232

Differencing: 1 regular, 1 seasonal of order 4

Number of observations: Original series 73, after differencing 68

Residuals: SS = 458838843 (backforecasts excluded)
MS = 7059059 DF = 65

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

Lag	12	24	36	48
Chi-Square	16,8	27,3	44,7	50,4
DF	9	21	33	45
P-Value	0,051	0,161	0,085	0,267

Data och prognoser från modell 2

Forecasts from period 73

Period	Forecast	95% Limits		Actual
		Lower	Upper	
74	357528	352358	362698	
75	355662	348351	362973	

diff1_4	RES11	inkomst	tid
-9378	-6061,13	324719	64
-4181	-3101,63	319592	65
-4935	-3355,14	333811	66
-171	-3116,58	329046	67
-630	-5600,40	323315	68
3485	896,94	321673	69
5036	2243,86	340928	70
3664	1063,90	339827	71
8770	4170,47	342866	72
-1236	-605,28	339988	73