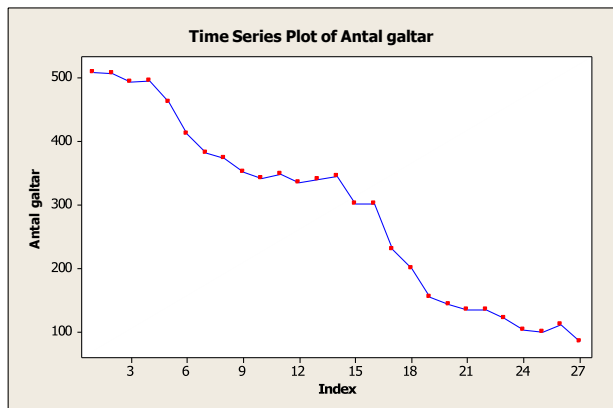


ÖVNINGAR TILL LEKTION 4 och 5

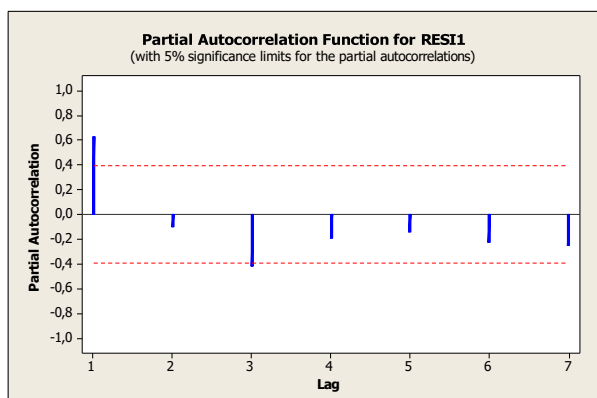
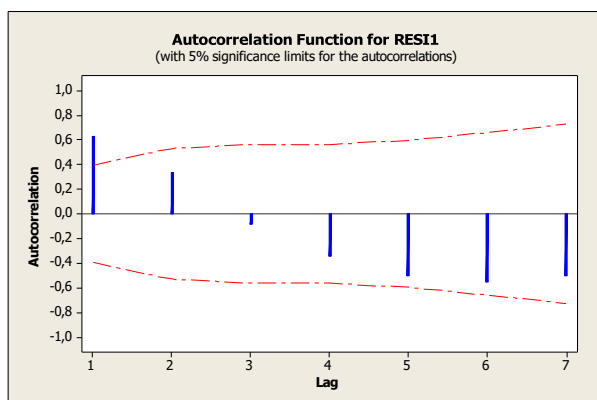
1

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 utjämningskonstanterna $\alpha = 0,4$ och $\gamma = 0,05$. Serien har totalt 27 värden och $l_{26} = 85,16$ och $b_{26} = -17,07$. De sista observationerna är $y_{25} = 98, y_{26} = 111, y_{27} = 84$.

- Beräkna prognoser för 2008 = tidpunkt 28 och 2009 då du står vid 2007 = tidpunkt 27.
- Nedan visas SAC och SPAC för residualerna från den dubbla exponentiella utjämningen. Vilken ARMA modell föreslår dessa?



- 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.
- d) Pröva med hjälp av utskriften nedan om residualerna är autokorrelerade.

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%

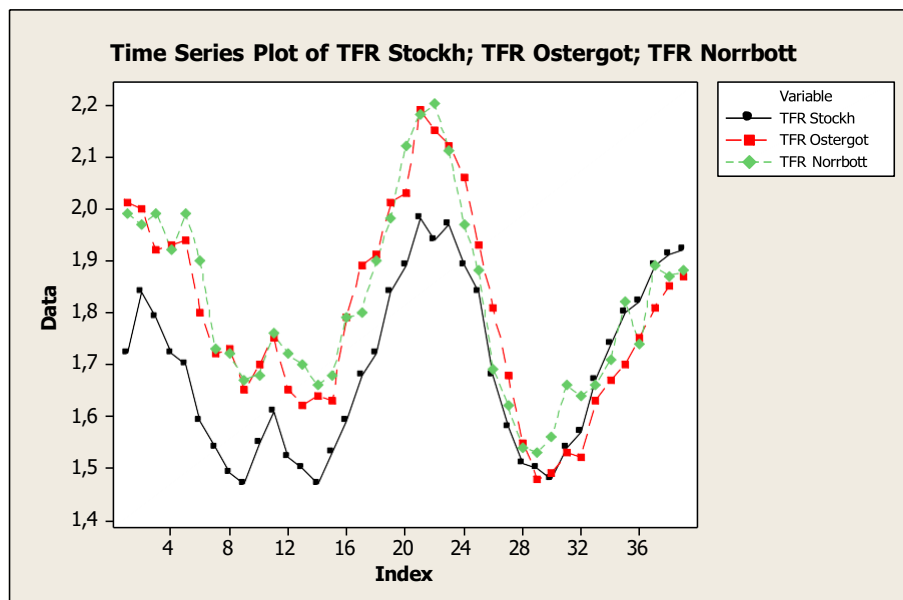
Analysis of Variance

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

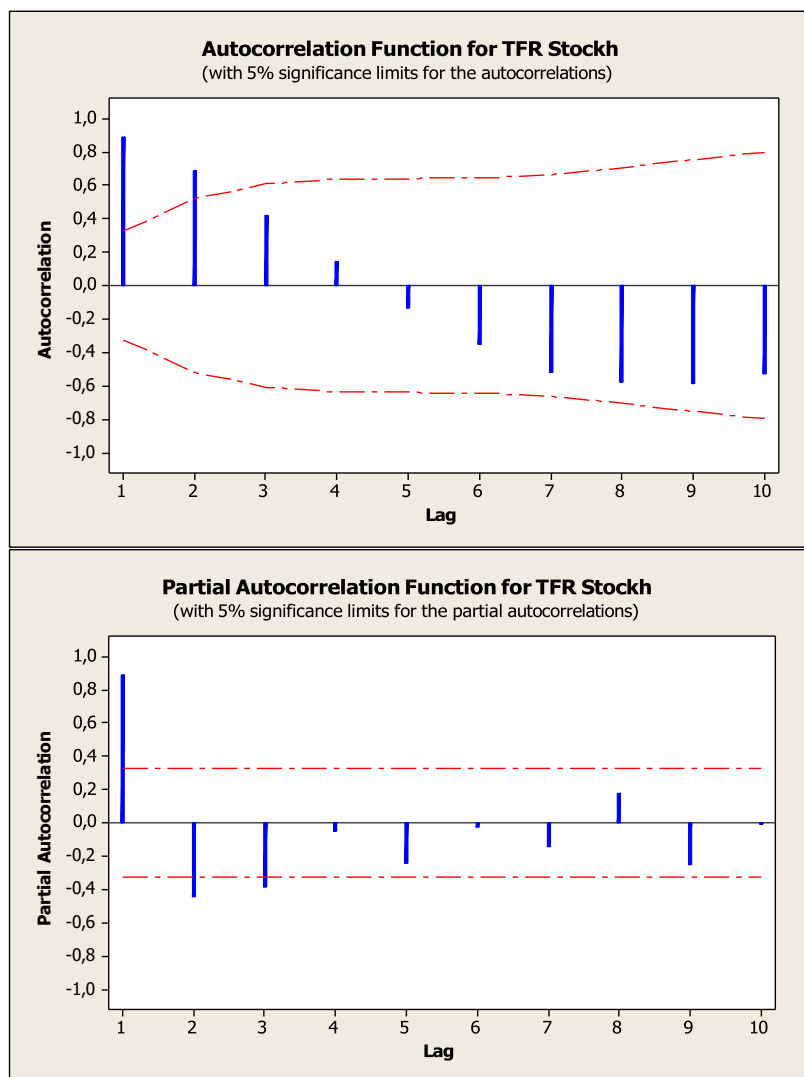
Durbin-Watson statistic = 0,543939

2

I grafen nedan ses TFR= total fertility rate för länen Stockholm, Östergötland och Norrbotten åren 1970 till 2008. Värden på TFR 2007 är 1,84 och för 2008 1,86



- a) Anta nu att vi vill anpassa en ARIMA modell till TFR för Stockholms län. SAC och SPAC visas nedan tillsammans med 6 modeller. Ta hjälp av utskrifterna nedan och välj den bästa modellen. Förklara noga hur du går tillväga.
- b) Gör en prognos för TFR för 2009 och 2010 med hjälp av modell 1.
- c) Uttryck modell tre med hjälp av bakåtskiftoperatorer.



Modell 1

ARIMA Model: TFR Stockh

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	1,4604	0,1361	10,73	0,000
AR	2	-0,6015	0,1395	-4,31	0,000
Constant		0,238751	0,009373	25,47	0,000
Mean		1,69201	0,06643		

Number of observations: 39

Residuals: SS = 0,121049 (backforecasts excluded)
MS = 0,003362 DF = 36

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

Lag	12	24	36	48
Chi-Square	52,6	68,3	75,1	*
DF	9	21	33	*
P-Value	0,000	0,000	0,000	*

Modell 2

ARIMA Model: TFR Stockh

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
MA	1	-1,4312	0,1071	-13,36	0,000
MA	2	-0,8051	0,1011	-7,96	0,000
Constant		1,69354	0,03478	48,70	0,000
Mean		1,69354	0,03478		

Number of observations: 39

Residuals: SS = 0,190671 (backforecasts excluded)
MS = 0,005296 DF = 36

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

Lag	12	24	36	48
Chi-Square	57,5	79,8	98,9	*
DF	9	21	33	*
P-Value	0,000	0,000	0,000	*

Modell 3

ARIMA Model: TFR Stockh

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	1,8239	0,0847	21,52	0,000
AR	2	-0,9506	0,0810	-11,73	0,000
MA	1	0,7377	0,1869	3,95	0,000
MA	2	-0,1429	0,1832	-0,78	0,441
Constant		0,212843	0,003504	60,74	0,000
Mean		1,68042	0,02767		

Number of observations: 39

Residuals: SS = 0,0963436 (backforecasts excluded)
MS = 0,0028336 DF = 34

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

Lag	12	24	36	48
Chi-Square	17,0	25,0	33,7	*
DF	7	19	31	*
P-Value	0,017	0,162	0,338	*

Modell 4

ARIMA Model: TFR Stockh

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	-0,0983	0,1193	-0,82	0,415
AR	2	0,9589	0,0965	9,93	0,000
MA	1	-1,0187	0,0307	-33,21	0,000

Constant	0,24560	0,02221	11,06	0,000
Mean	1,7620	0,1594		

Number of observations: 39

Residuals: SS = 0,169205 (backforecasts excluded)
MS = 0,004834 DF = 35

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

Lag	12	24	36	48
Chi-Square	62,7	85,3	98,2	*
DF	8	20	32	*
P-Value	0,000	0,000	0,000	*

Modell 5

ARIMA Model: TFR Stockh

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	0,9874	1,0917	0,90	0,372
MA	1	0,1911	1,0675	0,18	0,859
MA	2	0,7696	0,8651	0,89	0,380
Constant		0,021833	0,008850	2,47	0,019
Mean		1,7349	0,7032		

Number of observations: 39

Residuals: SS = 0,460975 (backforecasts excluded)
MS = 0,013171 DF = 35

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

Lag	12	24	36	48
Chi-Square	92,2	118,6	140,2	*
DF	8	20	32	*
P-Value	0,000	0,000	0,000	*

Modell 6

ARIMA Model: TFR Stockh

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	0,8928	0,0901	9,91	0,000
MA	1	-0,3128	0,1703	-1,84	0,075
Constant		0,18499	0,01404	13,18	0,000
Mean		1,7255	0,1309		

Number of observations: 39

Residuals: SS = 0,150590 (backforecasts excluded)
MS = 0,004183 DF = 36

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

Lag	12	24	36	48
Chi-Square	44,6	61,8	71,1	*
DF	9	21	33	*
P-Value	0,000	0,000	0,000	*

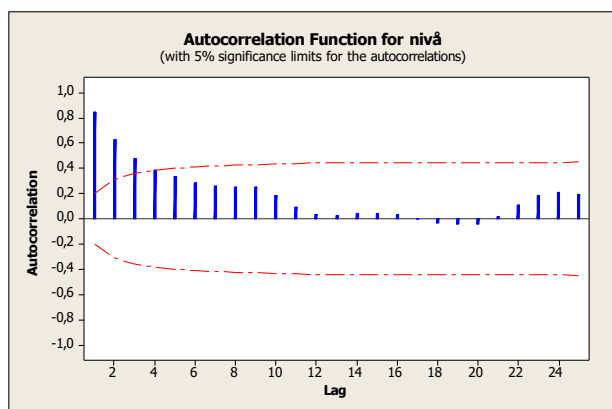
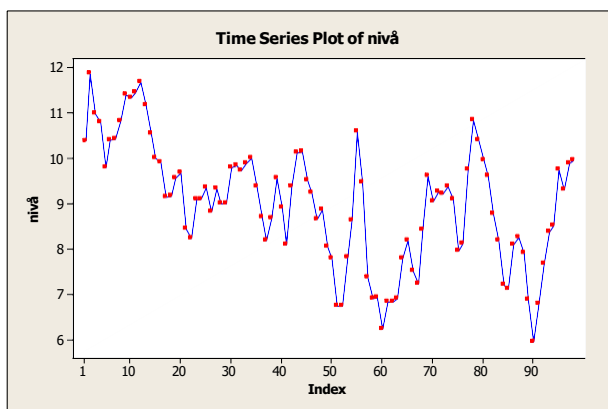
3

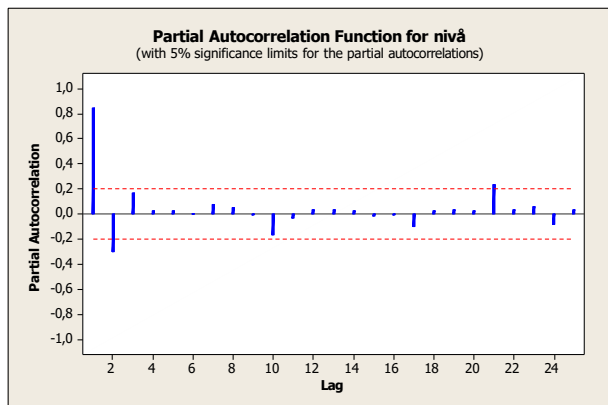
Vattennivån (mätt i fot) i en viss sjö har mätts under 98 år. Se tidsserieplotten för nivå nedan.

- a) Man kan ana att en viss sänkning av nivån har skett. Skatta den linjära trendfunktionen. Du kan direkt använda de uträknade summorna nedan där Y = nivå och x = år.

$$\sum Y = 881,4 \quad \sum Y^2 = 8096,6 \quad \sum x = 4851 \quad \sum x^2 = 318549 \quad \sum xY = 41694,8$$

- b) Om lutningen på linjen i a-uppgiften är -0,025, Hur många fot har sjön sjunkit under de 98 åren enligt modellen?
- c) Nedan visas SAC och SPAC för tidserien vattennivå. Vilken ARMA modell verkar vara rimlig att pröva enligt SAC och SPAC?
- d) Fyra ARMA modeller är anpassade nedan. Välj den bästa av dessa. Förklara hur du går tillväga.
- e) Skriv upp hur den valda modellen i d-uppgiften ser ut. Skriv den även med bakåtskiftoperatorer.





Modell 1

ARIMA Model: nivå

Estimates at each iteration

Relative change in each estimate less than 0,0010

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	1,1018	0,0980	11,24	0,000
AR	2	-0,2944	0,0982	-3,00	0,003
Constant		1,73932	0,06796	25,59	0,000
Mean		9,0335	0,3529		

Number of observations: 98

Residuals: SS = 42,9367 (backforecasts excluded)

MS = 0,4520 DF = 95

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

Lag	12	24	36	48
Chi-Square	7,7	16,5	22,2	28,3
DF	9	21	33	45
P-Value	0,565	0,741	0,923	0,976

Modell 2

ARIMA Model: nivå

Estimates at each iteration

Relative change in each estimate less than 0,0010

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
MA	1	-1,0891	0,0862	-12,64	0,000
MA	2	-0,5604	0,0866	-6,47	0,000
Constant		9,0026	0,1929	46,67	0,000
Mean		9,0026	0,1929		

Number of observations: 98

Residuals: SS = 49,6183 (backforecasts excluded)
MS = 0,5223 DF = 95

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

Lag	12	24	36	48
Chi-Square	32,9	40,8	51,3	72,6
DF	9	21	33	45
P-Value	0,000	0,006	0,022	0,006

Modell 3

ARIMA Model: nivå

Estimates at each iteration

Relative change in each estimate less than 0,0010

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	0,1808	0,7602	0,24	0,813
AR	2	0,4015	0,6055	0,66	0,509
MA	1	-0,9711	0,7464	-1,30	0,196
MA	2	-0,3120	0,2573	-1,21	0,228
Constant		3,7773	0,1553	24,33	0,000
Mean		9,0426	0,3717		

Number of observations: 98

Residuals: SS = 42,0701 (backforecasts excluded)
MS = 0,4524 DF = 93

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

Lag	12	24	36	48
Chi-Square	4,7	14,5	21,2	28,5
DF	7	19	31	43
P-Value	0,702	0,752	0,905	0,956

Modell 4

ARIMA Model: nivå

Estimates at each iteration

Relative change in each estimate less than 0,0010

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	0,7618	0,0761	10,01	0,000
MA	1	-0,3651	0,1091	-3,35	0,001
Constant		2,15445	0,09232	23,34	0,000
Mean		9,0462	0,3876		

Number of observations: 98

Residuals: SS = 42,4992 (backforecasts excluded)

MS = 0,4474 DF = 95

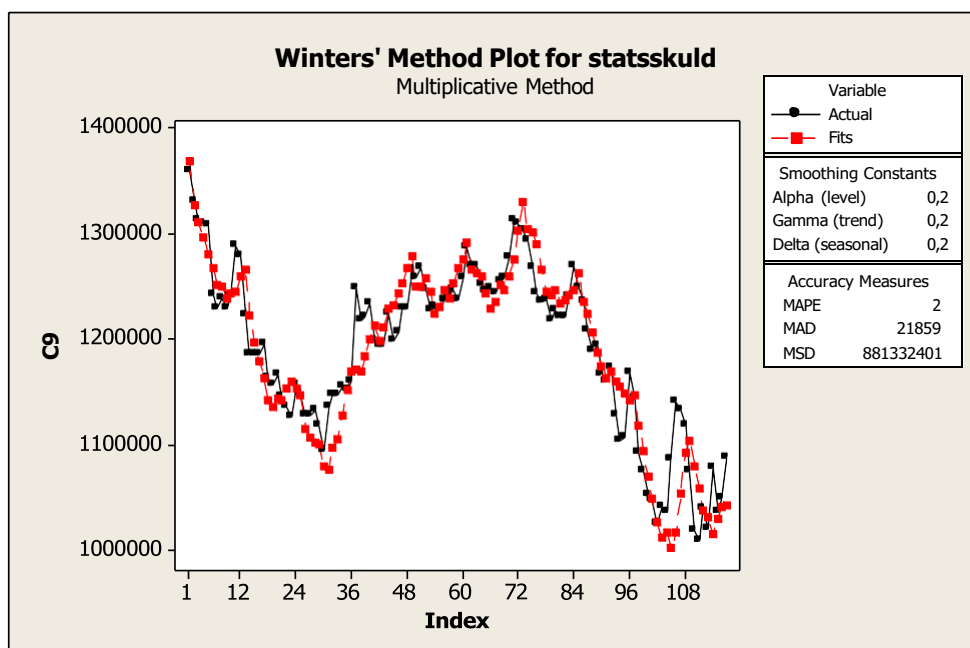
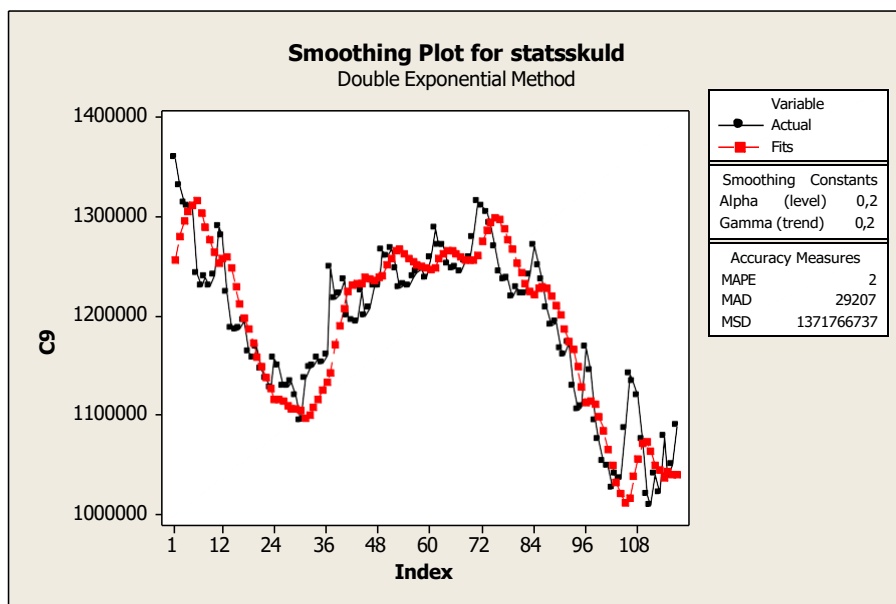
Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	5,6	16,0	22,7	29,3
DF	9	21	33	45
P-Value	0,782	0,772	0,912	0,966

4

Nedan visas två typer av utjämningar som använts på en tidsserie som visar Sveriges statskuld månadsvis från januari 2000 till september 2009.

Ge en kort beskrivning av de bägge metoderna och för vilka tidsserier de är lämpliga att använda. Vilken utjämningsmetod verkar rimligast här?



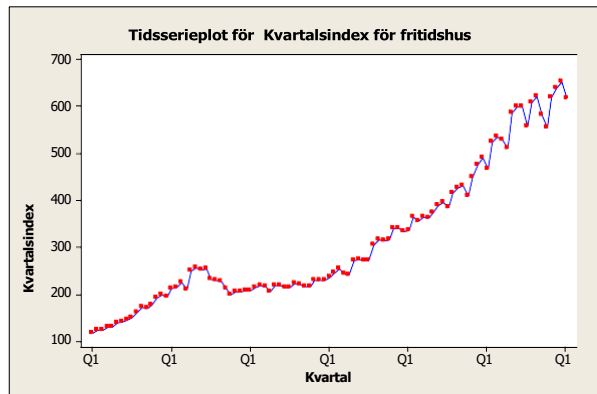
5

I grafen nedan visas

Fastighetsprisindex för fritidshus. Kvartal 1986K1-2010K1

(Q1 som syns i vissa grafer nedan betyder kvartal 1)

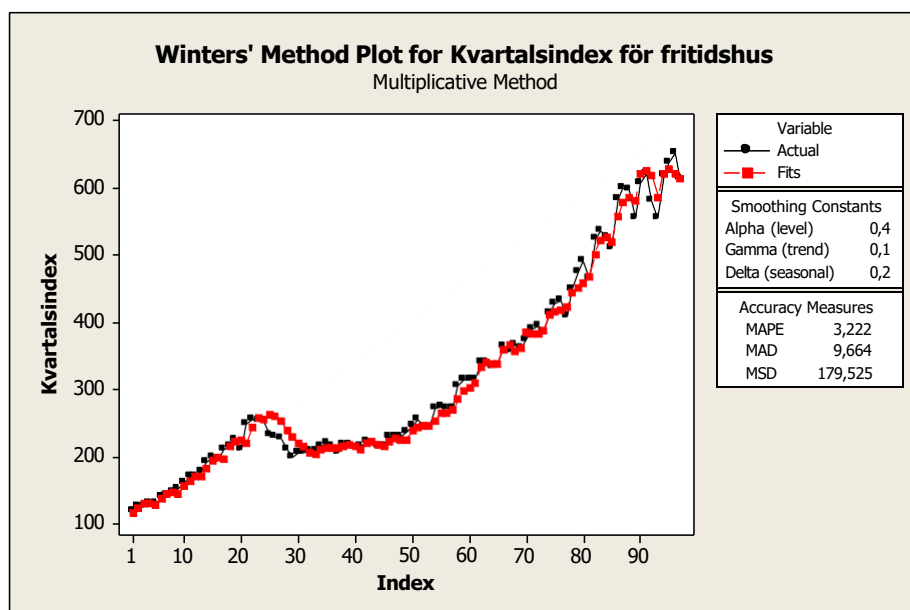
Serien kallas också bara **index**



Nedan har två typer av modeller anpassats och från modell 1 och 3 har prognoser beräknats.

- a) Studera modell 1. Visa hur modellen ser ut. Förklara hur parametrarna skattas samt hur man går tillväga för att hitta bäst anpassning. Är du nöjd med modellen. Motivera.

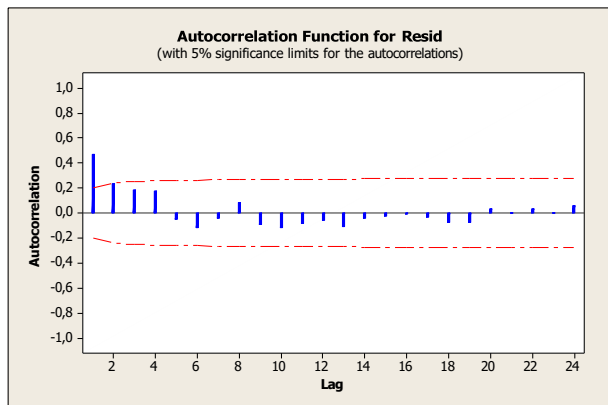
Modell 1



Forecasts

Period	Forecast	Lower	Upper
98	671,366	647,691	695,041
99	682,891	657,686	708,095
100	673,571	646,553	700,588
101	650,058	620,998	679,119

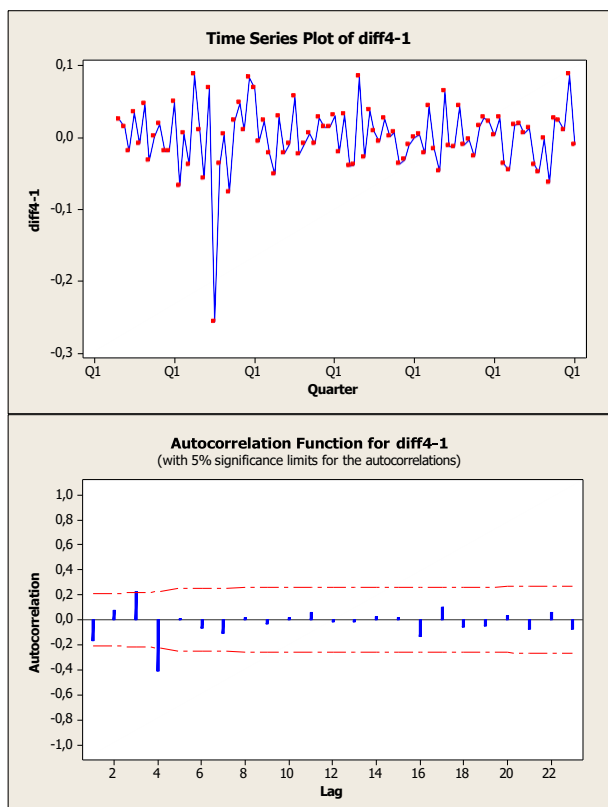
SAC för residualer från modell 1.

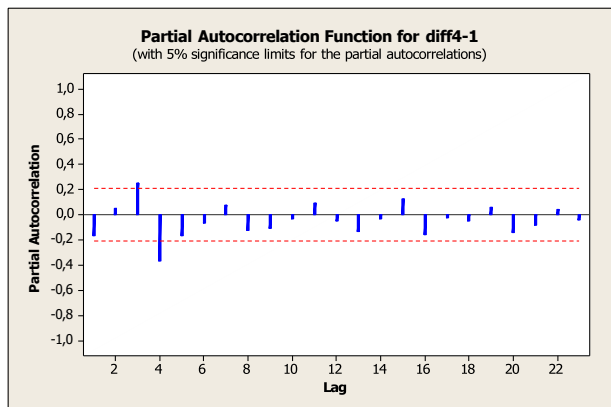


Nedan har två säsongs ARIMA modeller anpassats. Först har data logaritmerats med den naturliga logaritmen. Denna serie heter **ln index**. Sedan har **ln index** differentierats för säsong och för trend. Denna serie heter **diff4,1**

- Förklara varför serien har logaritmerats.
- Vilken av modell 2 och 3 anser du vara bäst. Motivera.
- För modell 3 har prognoser för ln index beräknats. Räkna om dessa till prognoser för index. Jämför prognosintervallen från modell 1 och modell 3. Vilka föredrar du?
- Låt $\text{diff4,1} = z_t$. Visa hur modell 3 ser ut för z_t .
- Skriv modell 2 med bakåtskiftoperatorer.

Modell 2 och 3





Modell 2

ARIMA Model: In index

Estimates at each iteration

Iteration	SSE		Parameters			
0	1,10237	0,100	0,100	0,100	0,100	0,081
1	0,28629	0,088	-0,043	0,112	0,243	-0,027
2	0,27429	-0,062	-0,034	-0,036	0,261	-0,029
3	0,26170	-0,212	-0,024	-0,183	0,282	-0,031
4	0,19901	-0,362	0,049	-0,316	0,413	-0,020
5	0,17517	-0,226	0,085	-0,166	0,481	-0,013
6	0,14776	-0,376	0,134	-0,301	0,592	-0,007
7	0,13583	-0,243	0,175	-0,151	0,679	-0,003
8	0,13004	-0,393	0,195	-0,295	0,750	-0,002
9	0,12763	-0,257	0,215	-0,145	0,795	-0,001
10	0,12626	-0,407	0,215	-0,292	0,831	-0,001
11	0,12584	-0,269	0,228	-0,142	0,849	-0,000
12	0,12560	-0,418	0,224	-0,292	0,864	-0,000
13	0,12548	-0,278	0,238	-0,142	0,873	-0,000
14	0,12543	-0,363	0,230	-0,231	0,875	-0,000
15	0,12540	-0,311	0,235	-0,174	0,878	-0,000
16	0,12540	-0,369	0,234	-0,235	0,880	-0,000
17	0,12538	-0,309	0,240	-0,171	0,883	-0,000
18	0,12537	-0,343	0,239	-0,207	0,884	-0,000
19	0,12536	-0,339	0,243	-0,202	0,886	-0,000
20	0,12536	-0,339	0,243	-0,202	0,886	-0,000

Relative change in each estimate less than 0,0010

Final Estimates of Parameters

Type		Coef	SE Coef
AR	1	-0,3391	0,7272
SAR	4	0,2429	0,1517
MA	1	-0,2025	0,7617
SMA	4	0,8857	0,0914
Constant		-0,0001336	0,0006898

Differencing: 1 regular, 1 seasonal of order 4

Number of observations: Original series 97, after differencing 92

Residuals: SS = 0,122825 (backforecasts excluded)
MS = 0,001412 DF = 87

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

Lag	12	24	36	48
Chi-Square	6,0	10,2	21,9	33,0
DF	7	19	31	43
P-Value	0,541	0,947	0,885	0,865

Modell 3

ARIMA Model: In index

Estimates at each iteration

Iteration	SSE	Parameters			
0	1,10237	0,100	0,100	0,090	
1	0,26502	-0,046	0,246	-0,027	
2	0,18848	0,028	0,396	-0,014	
3	0,15119	0,101	0,546	-0,006	
4	0,13435	0,169	0,696	-0,002	
5	0,12828	0,224	0,846	-0,000	
6	0,12799	0,216	0,868	-0,000	
7	0,12796	0,220	0,875	-0,000	
8	0,12795	0,224	0,879	-0,000	
9	0,12795	0,228	0,882	-0,000	
10	0,12795	0,230	0,884	-0,000	
11	0,12795	0,232	0,886	-0,000	

Unable to reduce sum of squares any further

Final Estimates of Parameters

Type	Coef	SE Coef
SAR 4	0,2317	0,1403
SMA 4	0,8855	0,0867
Constant	-0,0001128	0,0005739

Differencing: 1 regular, 1 seasonal of order 4

Number of observations: Original series 97, after differencing 92

Residuals: SS = 0,125022 (backforecasts excluded)
MS = 0,001405 DF = 89

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

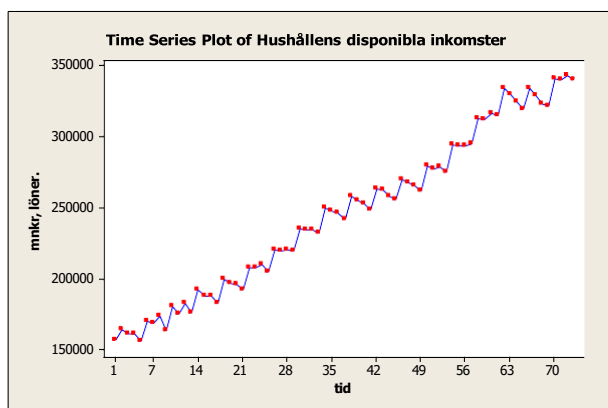
Lag	12	24	36	48
Chi-Square	7,2	11,3	23,8	36,0
DF	9	21	33	45
P-Value	0,620	0,956	0,880	0,830

Forecasts from period 97

Period	Forecast	95% Limits		Actual
		Lower	Upper	
98	6,51030	6,43683	6,58378	
99	6,53415	6,43024	6,63806	
100	6,53354	6,40628	6,66081	
101	6,49809	6,35114	6,64504	

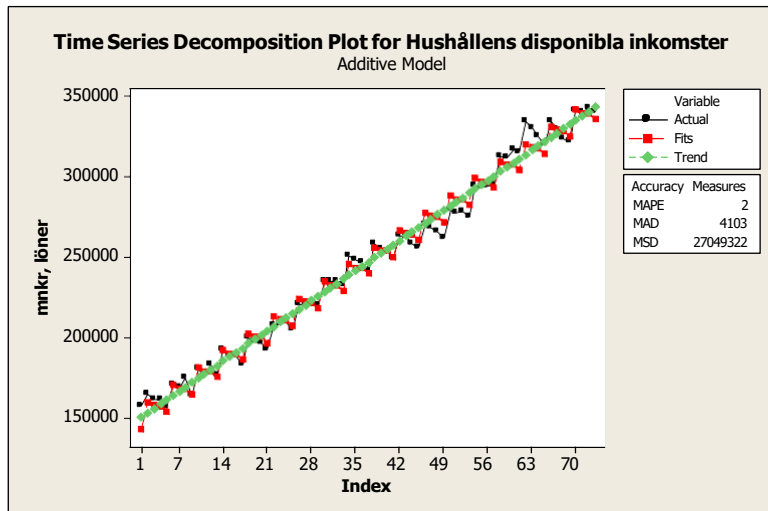
6

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



- Tolka säsongskomponenten för kvartal 4 i modell 1 och 2.
- Hur mycket kommer den disponibla inkomsten öka per år enligt trendlinjen. Välj modell 1 eller 2 vid denna tolkning.
- Gör prognos för kvartal 2 och 3 år 2011 dels med modell 1 och dels med modell 2.
- Tolka Durbin-Watson statistikan i modell 2.
- Eftersom både modell 1 och 2 uppvisade autokorrelerade residualer så har 3 SARIMA-modeller anpassats (modell 3 till modell 5). Jämför dessa tre modeller noggrant. Vilken anser du vara den bästa modellen. Motivera.
- Låt y_t vara den disponibla inkomsten. Låt vidare z_t vara $\text{diff}4,1$, 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 .
- Modell 4 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.

Modell 1



Time Series Decomposition for Hushållens disponibla inkomster

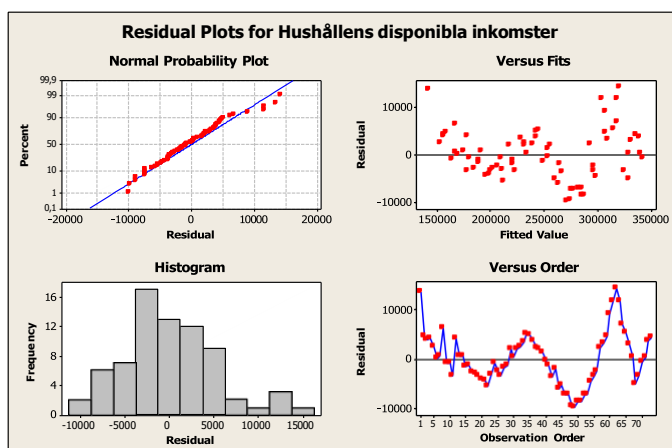
Additive Model

Data Hushållens disponibla inkomster
Length 73
NMissing 0

Fitted Trend Equation
 $Y_t = 147508 + 2679 \cdot t$

Seasonal Indices

Period	Index
1	-7283,02
2	6593,86
3	1824,67
4	-1135,52



Modell 2

Regression Analysis: Hushållens dispo versus kv2; kv3; kv4; tid

The regression equation is

$$\text{Hushållens disponibla inkomster} = 140745 + 13090 \text{ kv2} + 8259 \text{ kv3} + 5702 \text{ kv4} + 2678 \text{ tid}$$

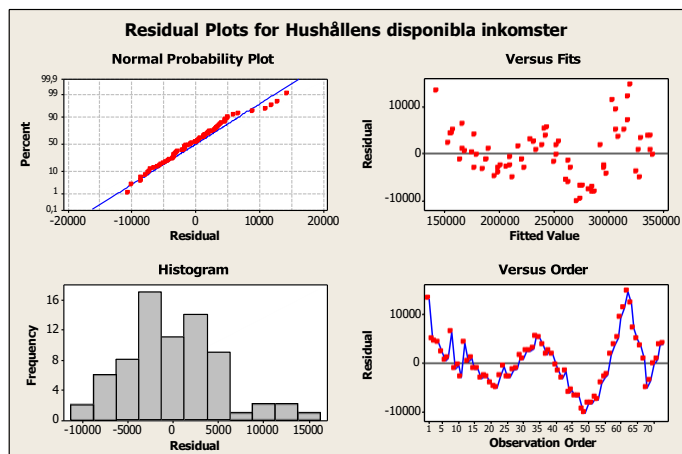
Predictor	Coef	SE Coef	T	P
Constant	140745	1657	84,96	0,000
kv2	13090	1769	7,40	0,000
kv3	8259	1769	4,67	0,000
kv4	5702	1769	3,22	0,002
tid	2678,35	29,88	89,62	0,000

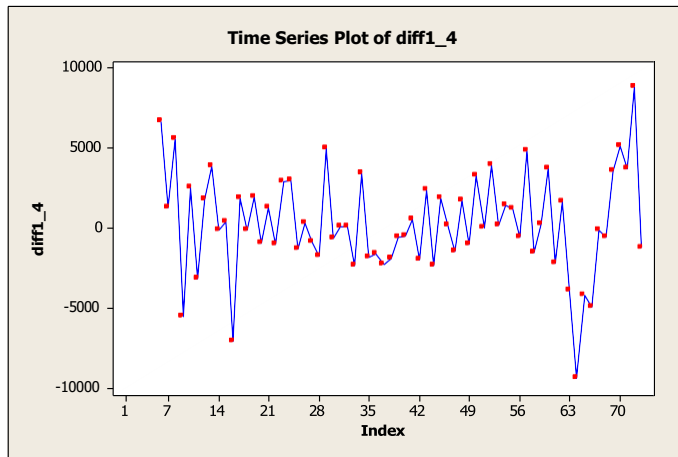
S = 5377,16 R-Sq = 99,2% R-Sq(adj) = 99,1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	2,33447E+11	58361829379	2018,47	0,000
Residual Error	68	1966141149	28913840		
Total	72	2,35413E+11			

Durbin-Watson statistic = 0,258596





Modell 3

ARIMA Model: Hushållens disponibla inkomster

Unable to reduce sum of squares any further

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 4

ARIMA Model: Hushållens disponibla inkomster

Unable to reduce sum of squares any further

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 5

ARIMA Model: Hushållens disponibla inkomster

Unable to reduce sum of squares any further

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 4

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