

TENTAMEN I STATISTIK B, 2014-12-05

Skrivtid: kl: 8-12
Tillåtna hjälpmedel: Ett A4-blad med egna handskrivna anteckningar samt räknedosa
Jourhavande lärare: Lotta Hallberg
Betygsgränser: Tentamen omfattar totalt 20p. 12 poäng och uppåt ger betyget G, 16 poäng och uppåt ger betyget VG.

Redovisa och motivera kort alla dina lösningar samt tolka dina svar

1

Tabellen nedan visar för åren 2009-2013 hushållens konsumtion av livsmedel (löpande priser i 1000 miljoner kronor) i Sverige och medelbefolkningen (i 1000-tal). Även Konsumentprisindex redovisas:

År	2009	2010	2011	2012	2013
<i>Konsumtion livsmedel</i>	166	169	175	179	186
<i>Folkmängd</i>	9341	9416	9483	9556	9645
<i>KPI</i>	299	302	311	314	314

- Beskriv utvecklingen av konsumtion av livsmedel i form av en indexserie med 2009 som basår.
1p
- Räkna om livsmedelskonsumtionen till fast penningvärde i priser för år 2013 med hjälp av KPI.
1,5p
- Beskriv med en indexserie hur livsmedelskonsumtionen per capita har utvecklats år 2009 till år 2013. Basår 2009.
1.5p

2

Anta att en ny bil ska inhandlas till hushållet. Åtta annonser studeras därför i dagspressen. Den enda variabel som antas påverka priset på bil är bilens ålder.

Skatta regressionskoefficienten β_1 med minsta kvadratmetoden (least squares estimate) i den enkla linjära regressionsmodellen med endast ålder som förklarande variabel. Pröva sedan hypoteserna:

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

Du får givet summorna: $\sum pris = 265$, $\sum ålder = 94$, $\sum pris * ålder = 2840$, $\sum pris^2 = 10669$, $\sum ålder^2 = 1150$. Priset är i 1000tals kronor. Signifikansnivå 5%.

Du får använda direkt att $SSE=243,28$.

3p

3

Från SCB:s hemsida har följande data hämtats.

Undervisande och forskande personal inom högskolan omräknat till heltidspersoner vid Handelshögskolan i Stockholm, Samhällsvetenskap, tjänstekategori, År 2001 – 2011.

Variabeln Tjänstekategori har värdena *Professorer, lektorer och adjunkter*. Denna variabel är omräknad till två Dummyvariabler; Professorer och lektorer.

Dummyvariabeln Professorer är 1 om Tjänstekategori = Professorer och 0 annars

Dummyvariabeln lektorer är 1 om Tjänstekategori = lektorer och 0 annars.

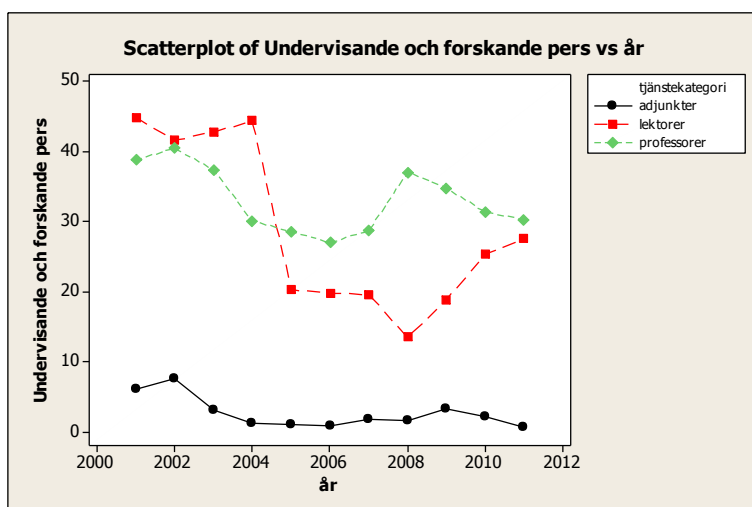
Dessutom har interaktionstermer skapats mellan dummyvariablerna och år; år*professor och år*lektor.

Först visas en graf över antalet heltidspersoner uppdelat på Tjänstekategori.

Därefter är två modeller anpassade med tillhörande residualplottar.

Uppgifter:

- Modell 2. Tolka regressionskoefficienten för år. Tolka även regressionskoefficienten för Professorer. 2p
- Vilken av modellerna är bäst om man studerar justerat R^2 ? 1p
- Pröva om med **ett** test om år*professor och år*lektor kan tas bort från modell 1. Använd signifikansnivå 5%. 2p
- En alumn som läst ett par kurser i statistik hävdar att man ska pröva att ta bort Professorer och år*professorer från modell 1 eftersom dessa inte är signifikanta. Förklara varför detta inte är en så god idé. 1p
- Kan man hävda att det saknas autokorrelation hos residualerna för de båda modellerna? 1p
- Modell 1: Beräkna ett 95% prediktionsintervall för antalet heltidsprofessorer år 2012 vid handelshögskolan i Stockholm, Samhällsvetenskap. Du får använda 'SE fit' direkt från utskrift utan att visa beräkning. Resten av beräkningarna ska redovisas. 2p



MODELL 1

Regression Analysis: Undervisande och forskande år; professorer; ...

The regression equation is

$$\begin{aligned} \text{Undervisande och forskande pers} = & 847 - 0,421 \text{ år} + 524 \text{ professorer} \\ & + 4390 \text{ lektorer} - 0,246 \text{ år*professor} \\ & - 2,18 \text{ år*lektor} \end{aligned}$$

Predictor	Coef	SE Coef	T	P
Constant	847	1110	0,76	0,452
år	-0,4208	0,5531	-0,76	0,453
professorer	524	1569	0,33	0,741
lektorer	4390	1569	2,80	0,009
år*professor	-0,2461	0,7822	-0,31	0,755
år*lektor	-2,1755	0,7822	-2,78	0,010

S = 5,80122 R-Sq = 88,2% R-Sq(adj) = 86,1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	6823,6	1364,7	40,55	0,000
Residual Error	27	908,7	33,7		
Total	32	7732,3			

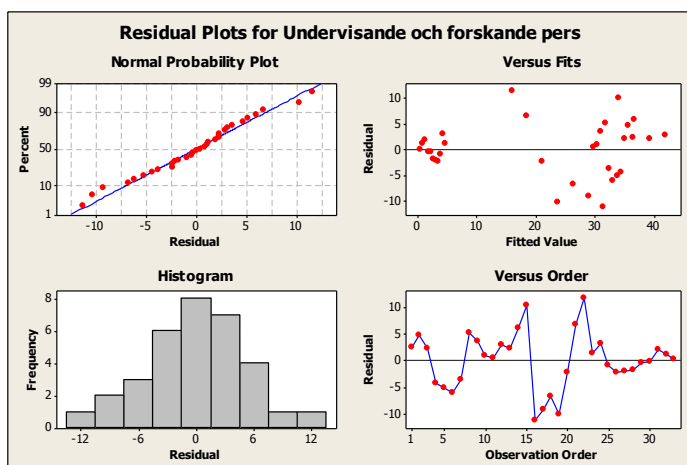
Durbin-Watson statistic = 1,08559

Predicted Values for New Observations

New Obs	Fit	SE Fit	95% CI	95% PI
1	29,17	3,75	(21,47; 36,87)	(15,00; 43,35)

Values of Predictors for New Observation

New Obs	år	professorer	lektorer	år*professor	år*lektor
1	2012	1,00	0,000000	2012	0,000000



MODELL 2

Regression Analysis: Undervisande och forskande år; professorer; lektorer

The regression equation is

$$\text{Undervisande och forskande pers} = 2466 - 1,23 \text{ år} + 30,5 \text{ professorer} + 26,3 \text{ lektorer}$$

Predictor	Coef	SE Coef	T	P
Constant	2466,0	716,5	3,44	0,002
år	-1,2280	0,3572	-3,44	0,002
professorer	30,499	2,767	11,02	0,000
lektorer	26,314	2,767	9,51	0,000

S = 6,48845 R-Sq = 84,2% R-Sq(adj) = 82,6%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	6511,4	2170,5	51,56	0,000
Residual Error	29	1220,9	42,1		
Total	32	7732,3			

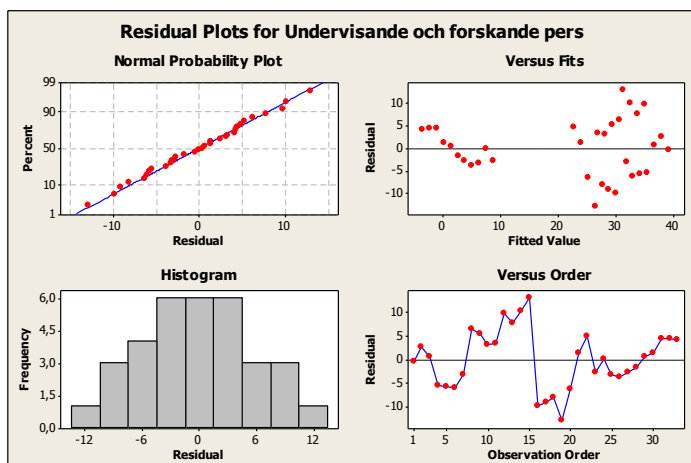
Durbin-Watson statistic = 0,796340

Predicted Values for New Observations

New Obs	Fit	SE Fit	95% CI	95% PI
1	25,80	2,90	(19,87; 31,74)	(11,27; 40,34)

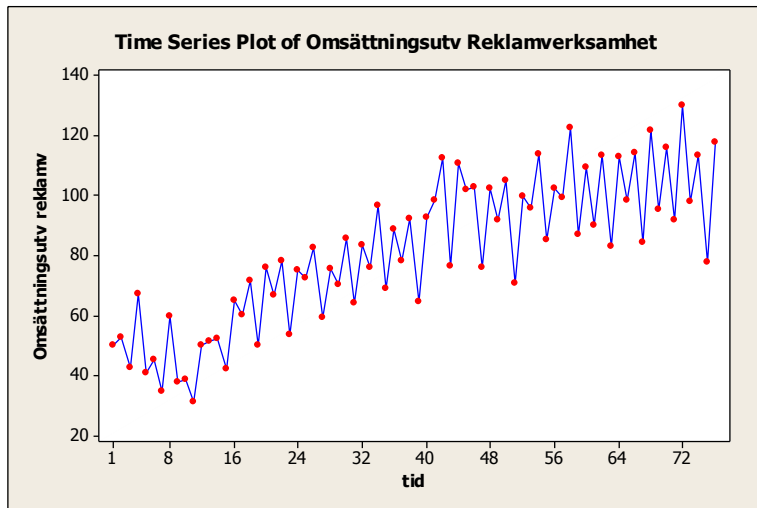
Values of Predictors for New Observations

New Obs	år	professorer	lektorer
1	2012	1,00	0,000000



4

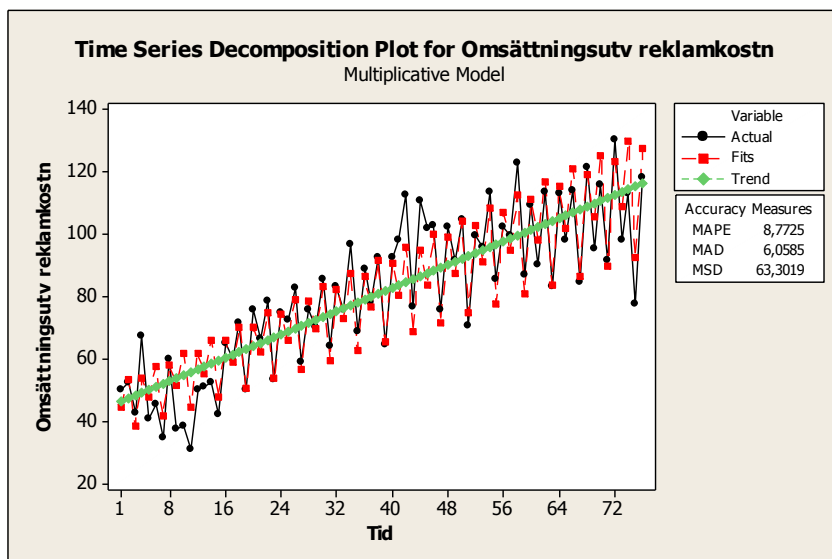
Nu ska omsättningsutvecklingen av reklamverksamhet i riket analyseras. Det är en indexserie med 2005 som basår. Kvartalsdata 1990K1 - 2008K4.



Den klassisk komponentuppdelnings metoden har använts nedan för skattning av komponenter.

Både en multiplikativ modell och en additiv modell har anpassats nedan.

- a) Vilken av anpassningarna anser du vara bäst? Motivera. 1p
- b) Tolka säsongskomponenten för kvartal 2 i båda modellerna. 1p
- c) Beräkna prognoser för nästföljande kvartal med hjälp av båda modellerna. Vilken av prognoserna verkar mest rimlig? 2p



Multiplicative Model

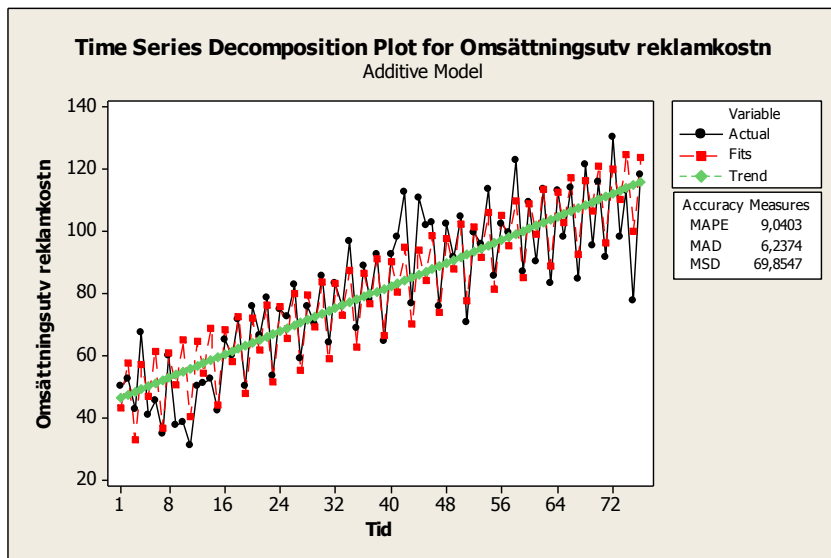
Data Omsättningsutv reklamkostn
 Length 76
 NMissing 0

Fitted Trend Equation

$$Y_t = 45,29 + 0,932 * t$$

Seasonal Indices

Period	Index
1	0,96205
2	1,13468
3	0,80477
4	1,09851



Additive Model

Data Omsättningsutv
 reklamkostn
 Length 76
 NMissing 0

Fitted Trend Equation

$$Y_t = 45,32 + 0,931 * t$$

Seasonal Indices

Period	Index
1	-3,1016
2	10,4422
3	-15,2578
4	7,9172