

Tentamen Statistik, 732G05 2016-09-24

Skrivtid: 08.00-12.00

Tillåtna hjälpmedel: *Miniräknare. Kursbok:* Bowerman, O'Connel, Koehler: Forecasting, Time series, and Regression (alla upplagor tillåtna - får innehålla markeringar, understrykningar och flärpar, men inte anteckningar), *A4-blad* med vad som helst skrivet på fram- och baksida (för hand/skrivet på dator/kopierat etc. - inga begränsningar).

Betyg: För godkänt betyg krävs 12 av 20 p. För väl godkänt betyg krävs 16 av 20 p.

Redovisa och motivera kort alla dina lösningar. Lycka till!

Uppgift 1. (5p)

Nedan ges husstorlek (100-tals kvadratfot) samt huspris (100-tals dollar) för ett slumpmässigt urval av hus i ett område i USA.

Hus	1	2	3	4	5
Storlek	18	20	25	22	19
Pris	80	95	104	110	85

- Skatta β_0 och β_1 i en enkel linjär regressionsmodell: $y = \beta_0 + \beta_1 x + \epsilon$ där y =pris och x =storlek. Tolk β_1 . (2p)
- Testa, på 5% signifikansnivå, om lutningen β_1 är 0 i populationen av alla hus i området. Dra en slutsats. (3p)

Uppgift 2. (3p)

Nedan visas uppgifter om priser och antal sålda exemplar av en vara.

År	Pris	Sålt antal
2013	340	1063
2014	360	1428
2015	385	986

- Beräkna ett enkelt prisindex för varan, med basår 2013. Hur stor har den procentuella prisökningen varit mellan 2013 och 2015? (1.5p)
- Beräkna försäljningsvärdena för de tre åren, i 2015 år priser. (1.5p)

Uppgift 3. (7p)

Nedan visas resultat (Minitabutskrift samt grafer) från skattning av en multipel regressionsmodell $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \epsilon$ där y =inkomst per capita (Inkomst), x_1 = procent av befolkningen under 15 år (Ung) och x_2 är en dummyvariabel kodad 1 om Latinamerika och 0 om Afrika, för 52 länder.

- Testa, på 5% signifikansnivå, om modellen som helhet är signifikant. Dra en slutsats. (1.5p)
- Tolka β_2 samt testa, på 5% signifikansnivå, om β_2 är noll i populationen. Dra en slutsats. (1.5p)
- Bilda ett 95%-igt konfidensintervall för β_3 och tolka intervallet. (2p)
- Beskriv vad som visas i graf nr 2 (Versus Fits), dvs beskriv vad det är som visas på X-axeln och Y-axeln och vad grafen används för. (1p)
- Undersök siffrorna i utskriften samt lämpliga grafer. Är det något som gör att du inte litar på resultaten från regressionsanalysen (exempelvis resultaten i uppgift b och c)? (1p)

Regression Analysis: Inkomst versus Ung; Land; Ung*Land

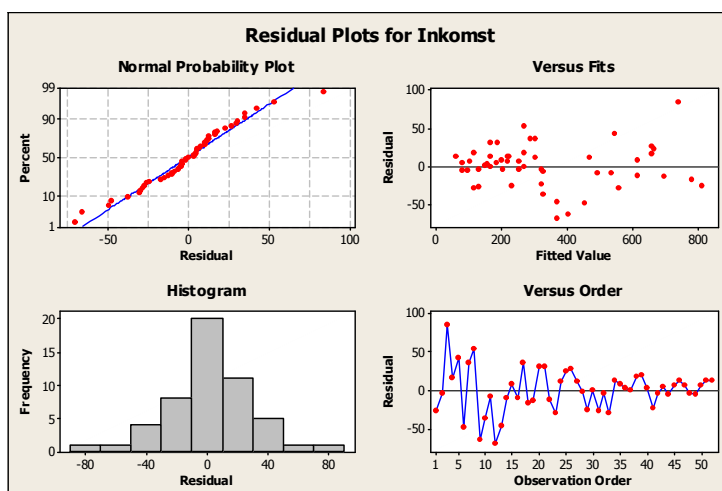
The regression equation is
Inkomst = 898 - 17,4 Ung + 1242 Land - 23,8 Ung*Land

Predictor	Coef	SE Coef	T	P	VIF
Constant	897,52	52,89	16,97	0,000	
Ung	-17,389	1,286	-13,52	0,000	2,078
Land	1242,28	73,08	17,00	0,000	82,947
Ung*Land	-23,764	1,789	-13,28	0,000	82,732

S = 28,9303 R-Sq = 98,2% R-Sq(adj) = 98,1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	2206742	735581	878,87	0,000
Residual Error	48	40174	837		
Total	51	2246916			



Uppgift 4. (5p)

Nedan visas resultat från flera olika tidsserieanalyser av BNP i USA 1981-1991 (kvartalsdata).

- Tolka utskriften och grafen för analys 1. (2p)
- Vad är det som skiljer analyserna 4 och 5 åt? Vilken av dessa (4 och 5) tycker du är mest lämplig för denna tidsserie? Motivera. (1.5p)
- Vilken av alla fem analyserna tycker du är mest lämplig för denna tidsserie? Motivera. Kan du komma på någon orsak till att den analys du har valt skulle vara mest lämplig? (1.5p)

Analys 1

Time Series Decomposition for GDP

Multiplicative Model

Data GDP
Length 44
NMissing 0

Fitted Trend Equation

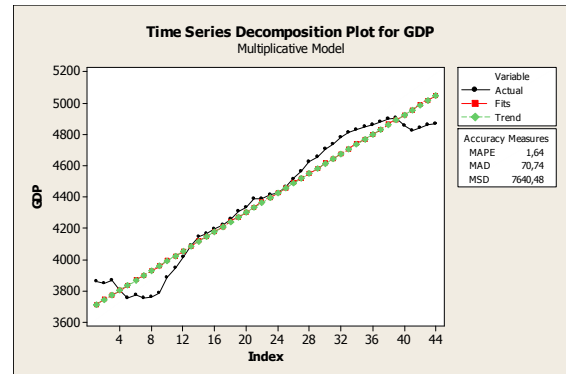
$$Y_t = 3678,6 + 31,1 \cdot t$$

Seasonal Indices

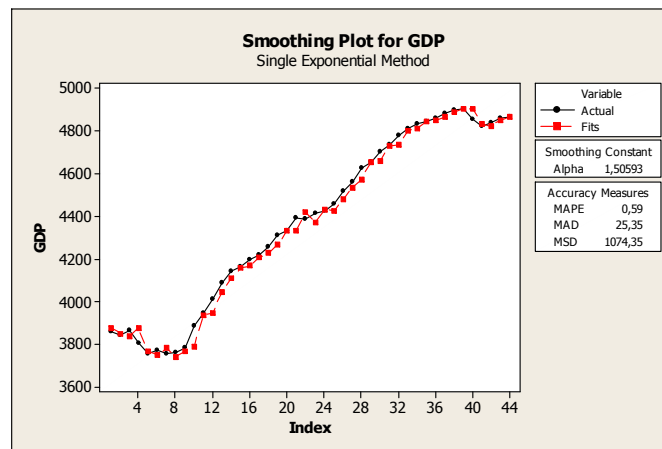
Period	Index
1	0,99926
2	1,00136
3	1,00013
4	0,99924

Accuracy Measures

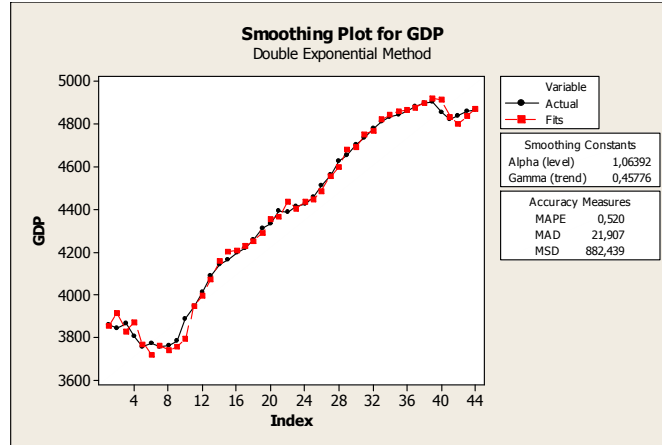
MAPE	1,64
MAD	70,74
MSD	7640,48



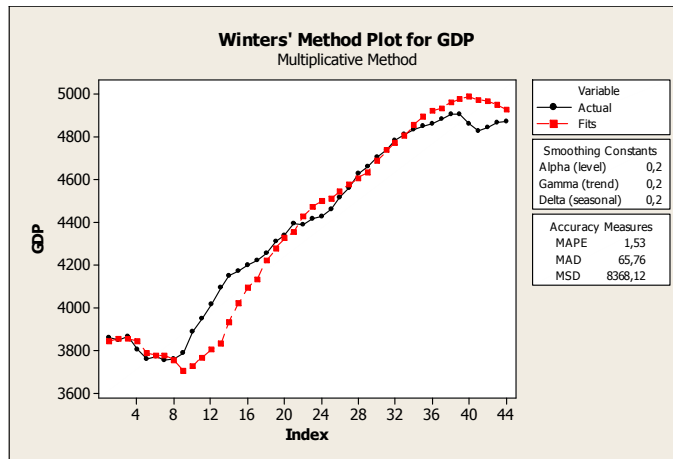
Analys 2



Analys 3



Analys 4



Analys 5

