

Tentamen i Regressions- och tidsserieanalys, 2008-06-09

Skrivtid: kl: 8-12

Hjälpmedel: Räknedosa. Bowerman, B.J., O'Connell, R, Koehler, A.: *Forecasting, Time Series and Regression*. 4th ed. Duxbury, 2005 som inte får innehålla anteckningar men får ha markeringar och flärpar. Formelsamling. Tabeller

Jourhavande lärare: Lotta Hallberg

Redovisa och motivera kort alla dina lösningar

Obs! Skriv namn och personnummer på varje papper du lämnar in.

1

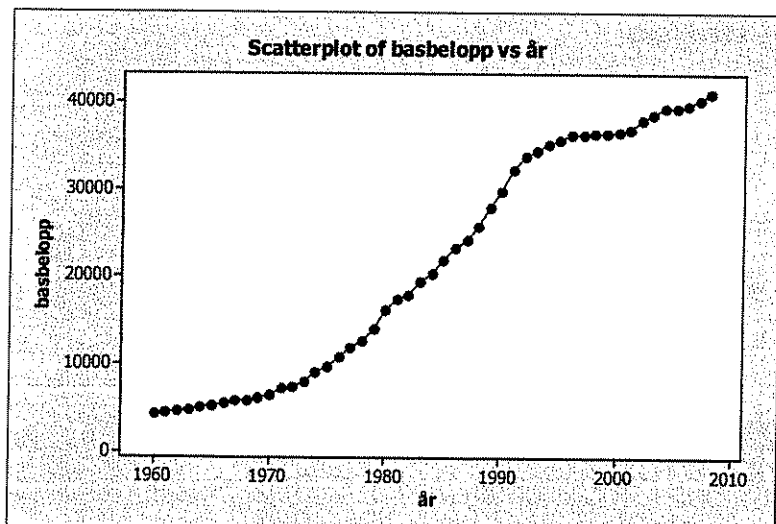
I tabellen nedan visas totalförsäljningen (i löpande priser) för insektsmedel för hästar och för hundar. Även priser för medlet 'Stopp o bit' för hästar från gruppen insektsmedel för hästar samt 'fästinghalsband' för hundar från gruppen insektsmedel för hundar är givna. Data är från åren 2004 till 2006

År	Tot. förs. insektsmedel för hästar	Tot. förs. insektsmedel för hundar	Pris, Stopp o bit	Pris, fästinghalsband
2004	1150	630	257	210
2005	1210	755	267	215
2006	1530	835	285	275

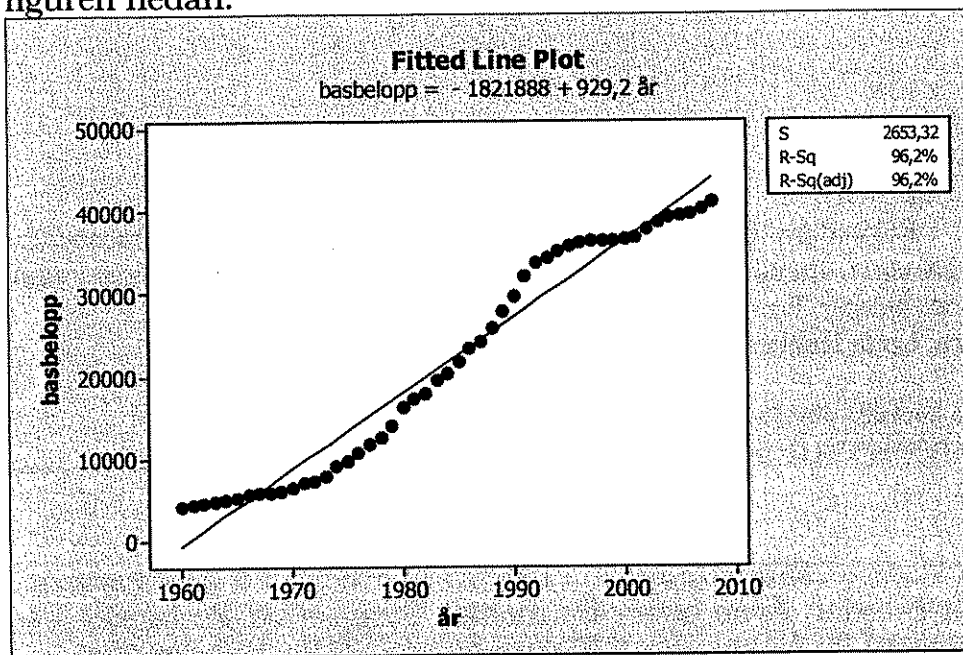
- Använd varorna 'Stopp o bit' och fästinghalsband som representantvaror för sina grupper och beräkna ett kedjeprisindex av Laspeyre-typ för företagets priser med basår 2004. 3p
- Hur har företagets för grupperna insektsmedel för hästar och hundars priser förändrats mellan 2004 och 2006? 1p
- Om man vill se hur prisutvecklingen för varugrupperna insektsmedel för hästar och för hundar har utvecklats gentemot den allmänna prisutvecklingen, hur kan då en lämplig indexserie beräknas som beskriver detta och vad kallas en sådan indexserie? 1p

2

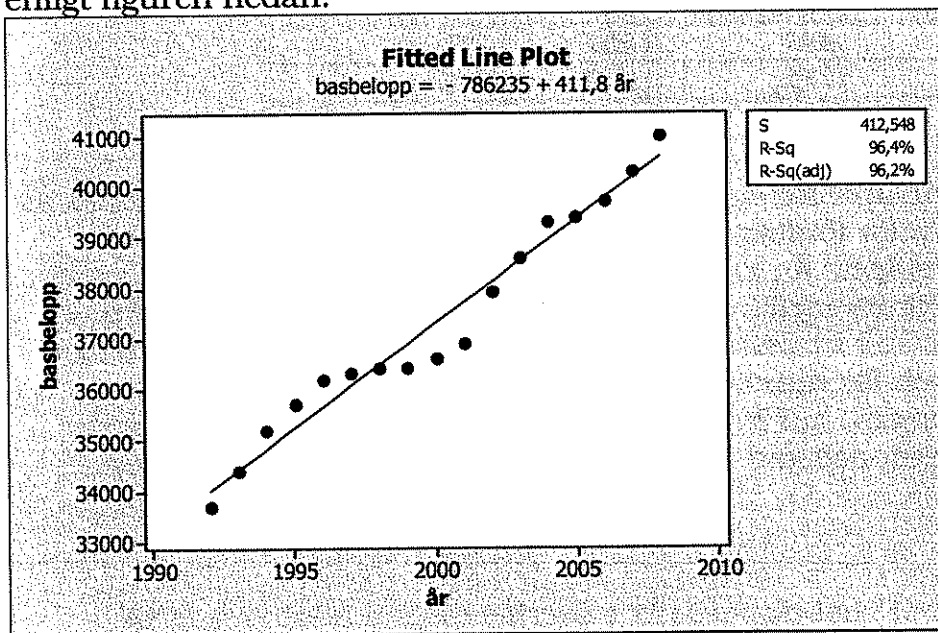
I grafen nedan ser vi basbeloppet i kr i Sverige mellan 1960 och 2008.



Till dessa data har en enkel linjär regressionsmodell anpassats vilket visas i figuren nedan.



Eftersom detta inte alls ser ut att vara en bra modell så har endast observationerna från 1992 och framåt använts och då blev anpassningen enligt figuren nedan.



Regression Analysis: basbelopp versus år

The regression equation is
 $\text{basbelopp} = -786235 + 412 \text{ år}$

Predictor	Coef	SE Coef
Constant	-786235	40848
år	411,76	20,42

S = 412,548 R-Sq = 96,4% R-Sq(adj) = 96,2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	69176471	69176471	406,45	0,000
Residual Error	15	2552941	170196		
Total	16	71729412			

Uppgifter:

- a) Hur mycket har basbeloppet ökat i snitt per år sedan 1992 enligt modellen ovan? 1p
- b) Beräkna ett 95% konfidensintervall för regressionskoefficienten i modellen ovan. Tolka intervallet. 1p
- c) Pröva om residualerna kan antas vara okorrelerade i modellen ovan. 1p

De 10 sista åren är basbeloppet:

36400, 36600, 36900, 37900, 38600, 39300, 39400, 39700, 40300, 41000

d) Använd dubbel exponentiell utjämning för linjär trend och visa hur de 3 sista åren utjämnas. Du kan använda direkt $l_{2005} = 39109,3$ och $b_{2005} = 286,153$. Gör även prognos för 2009, 2010 samt 2011 då du står vid tidpunkt 2008. Låt $\alpha = \gamma = 0,3$ 3p

e) Visa hur 3 punkters glidande medelvärden beräknas för åren 2005, 2006 samt 2007. 1p

Formler för dubbel exponentiell utjämning (=Holt's trend corrected exponential smoothing):

$$l_T = \alpha \cdot y_T + (1 - \alpha)(l_{T-1} + b_{T-1})$$

$$b_T = \gamma(l_T - l_{T-1}) + (1 - \gamma)b_{T-1}$$

$$\hat{y}_{T+\tau}(T) = l_T + \tau \cdot b_T$$

3

Man vill undersöka hur försäljningen för en viss typ av affär påverkas av antalet hushåll i närområdet samt hur affären är placerad. Under juli ett visst år i en viss stad valde man slumpmässigt ut 5 affärer som låg utmed en gata, 5 affärer som låg i centrala staden samt 5 affärer som låg i ett köpcentrum.

Variablerna är betecknade:

Y= försäljning i 1000-tal dollar

x= antal hushåll i 1000-tal

D1= 1 om affär placerad i köpcentrum och 0 annars

D2= 1 om affär placerad i centrala staden och 0 annars

Så vi ser att den kvalitativa variabeln som beskriver placering av affär är hanterad som två dummyvariabler D1 och D2.

Vidare är:

xD1 och xD2 interaktionstermer mellan x och D1, D2.

Sist kommer ett spridningsdiagram över data

Följande 3 modeller har anpassats:

Modell 1

Regression Analysis: Y versus x; D1; D2; xD1; xD2

The regression equation is

$$Y = 7,9 + 0,921 x + 42,7 D1 + 10,3 D2 - 0,092 xD1 - 0,034 xD2$$

Predictor	Coef	SE Coef	T	P	VIF
Constant	7,90	17,04	0,46	0,654	
x	0,9207	0,1234	7,46	0,000	11,723
D1	42,73	21,50	1,99	0,078	33,340
D2	10,26	21,28	0,48	0,641	32,658
xD1	-0,0917	0,1416	-0,65	0,533	53,014
xD2	-0,0336	0,1382	-0,24	0,813	59,686

S = 6,79953 R-Sq = 98,8% R-Sq(adj) = 98,1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	33296,1	6659,2	144,03	0,000
Residual Error	9	416,1	46,2		
Total	14	33712,2			

Source	DF	Seq SS
x	1	31244,4
D1	1	1940,9
D2	1	83,5
xD1	1	24,6
xD2	1	2,7

Modell 2

Regression Analysis: Y versus x; D1; D2

The regression equation is

$$Y = 15,0 + 0,869 x + 28,4 D1 + 6,86 D2$$

Predictor	Coef	SE Coef	T	P	VIF
Constant	14,978	6,188	2,42	0,034	
x	0,86859	0,04049	21,45	0,000	1,447
D1	28,374	4,461	6,36	0,000	1,646
D2	6,864	4,770	1,44	0,178	1,882

S = 6,34941 R-Sq = 98,7% R-Sq(adj) = 98,3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	33269	11090	275,07	0,000
Residual Error	11	443	40		
Total	14	33712			

Source	DF	Seq SS
x	1	31244
D1	1	1941
D2	1	83

Modell 3

Regression Analysis: Y versus x

The regression equation is

$$Y = 14,9 + 0,937 x$$

Predictor	Coef	SE Coef	T	P	VIF
Constant	14,87	13,13	1,13	0,278	
x	0,93712	0,07305	12,83	0,000	1,000

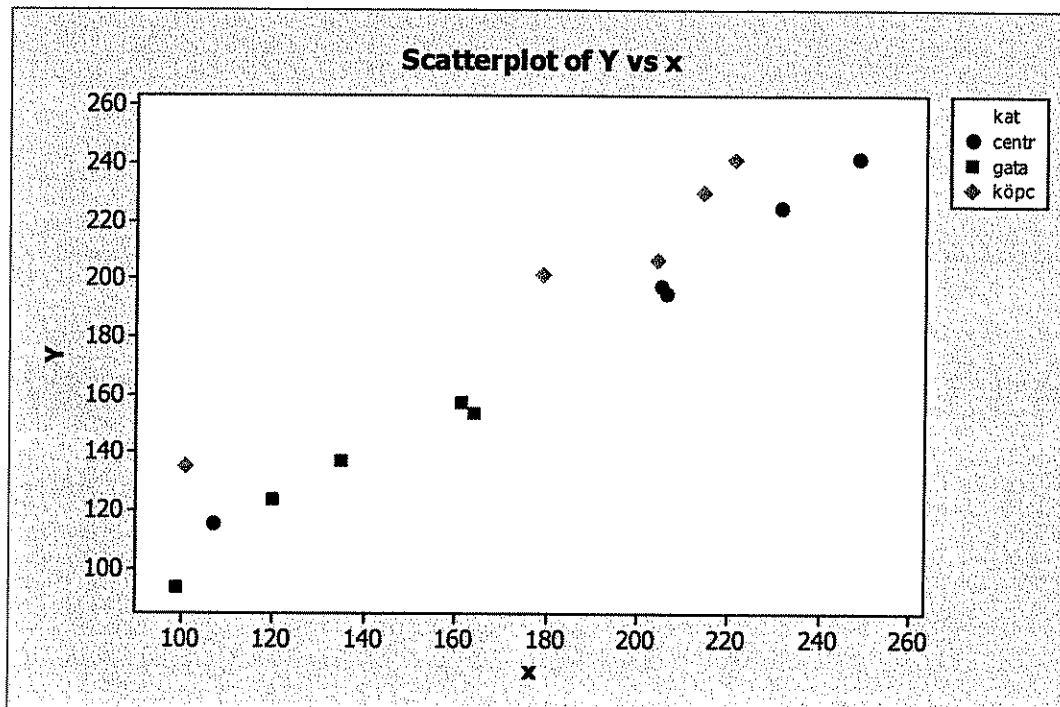
S = 13,7779 R-Sq = 92,7% R-Sq(adj) = 92,1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	31244	31244	164,59	0,000
Residual Error	13	2468	190		
Total	14	33712			

Uppgifter:

- Pröva med ett test om de båda interaktionstermerna i modell 1 kan tas bort. 2p
- Skriv upp modellen för försäljningen i en affär placerad i ett köpcentrum med ett godtyckligt antal hushåll med hjälp av modell 1. 1,5p
- Jämför VIF i de tre modellerna och förklara vilka värden som är bra och vilka som är dåliga. Vad mäter VIF och hur beräknas de? 2p
- Vilken av modellerna är bäst om man använder justerad R^2 som ett jämförande mått? 0,5p
- Tolka regressionskoefficienten framför D2. 0,5p
- Prediktera försäljningen för en affär som har 150000 hushåll i sin närhet och ligger i ett köpcentrum enligt modell 2. 1p
- Förklara orimligheten i de negativa tecknen framför xD1 och xD2 i modell 1. 0,5p



(

(

(

(