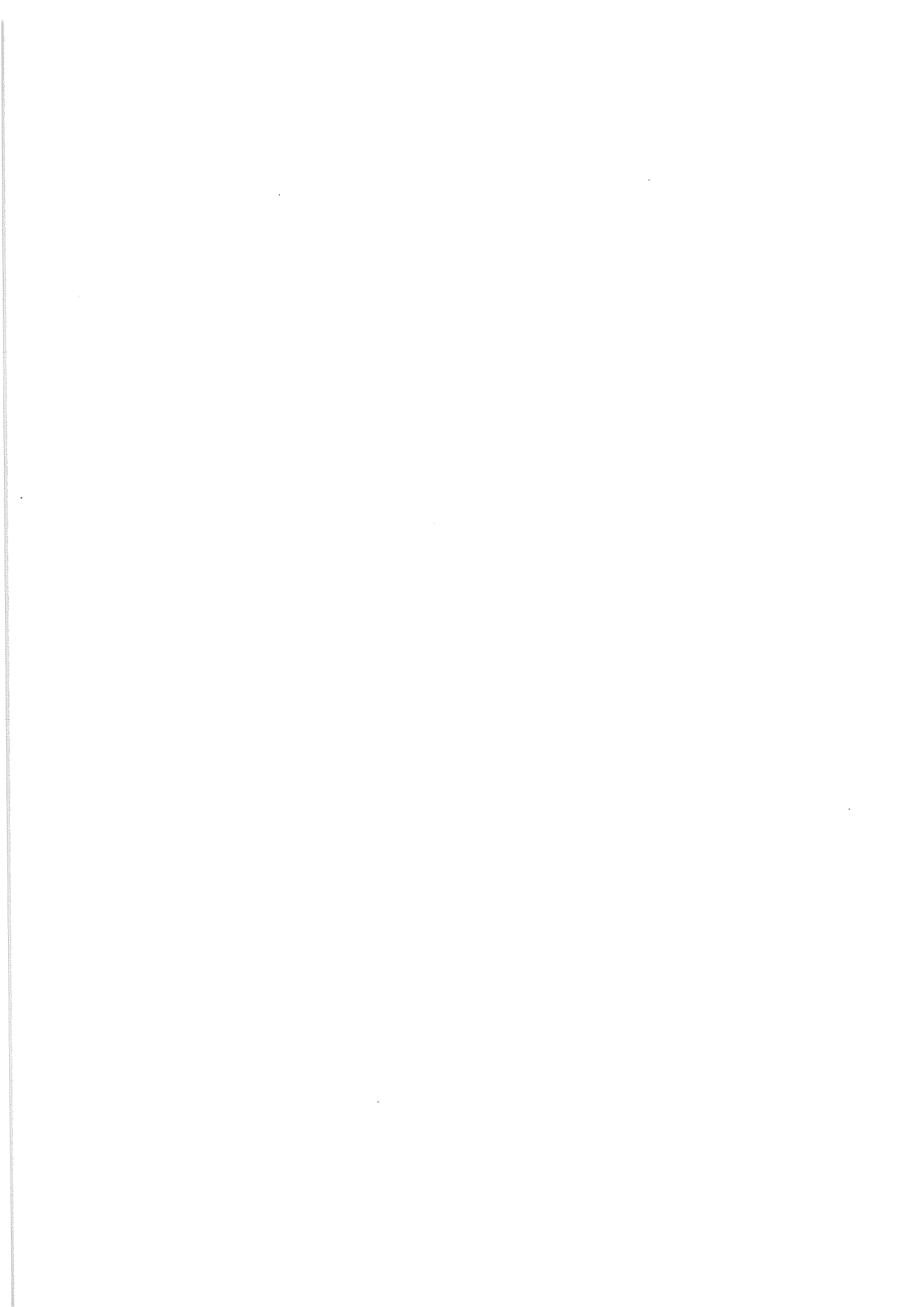


Försättsblad till skriftlig tentamen vid Linköpings universitet



Datum för tentamen	2017-12-08
Sal (3)	<u>TER1(138)</u> TERD(1) TERE(1)
Tid	8-12
Kurskod	732G71
Provkod	TENT
Kursnamn/benämning Provnamn/benämning	Statistik B Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	4
Jour/Kursansvarig Ange vem som besöker salen	Bertil Wegmann
Telefon under skrivtiden	070-1128321
Besöker salen ca klockan	
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Annelie Almquist 013-282934 annelie.almquist@liu.se
Tillåtna hjälpmedel	Ett A4-blad med egna handskrivna anteckningar (båda sidor) samt räknedosa
Övrigt	
Antal exemplar i påsen	



Tentamen

Linköpings universitet, Institutionen för datavetenskap, Statistik

Kurskod och namn:	732G71 Statistik B
Datum och tid:	2017-12-08, 8-12
Jourhavande lärare:	Bertil Wegmann
Tillåtna hjälpmedel:	Ett A4-blad med egna handskrivna anteckningar (båda sidor) samt räknedosa.
Betygsgränser:	Tentamen omfattar totalt 20p. Godkänt från 12p, väl godkänt från 16p. Siffrorna i uppgifterna är delvis påhittade.

Redovisa och motivera tydligt alla dina lösningar!

Uppgift 1 (8p)

Ett företag inom telefonförsäljning är intresserade av att undersöka om deras nyskapade test för försäljningsförmåga kan förklara antalet produkter per vecka som telefonförsäljarna säljer. Därför tog man ett slumpmässigt urval av 8 telefonförsäljare och uppmätte antalet testpoäng och antalet sålda produkter per vecka för varje försäljare. Detta gav följande tabell:

Säljare	Testpoäng	Antal sålda produkter per vecka
1	75	24
2	85	28
3	80	19
4	85	21
5	60	12
6	65	15
7	60	13
8	55	10

Använd en enkel linjär regressionsmodell för att förklara antalet sålda produkter per vecka med antalet testpoäng.

- Skatta regressionslinjens skärning och lutning i den enkla linjära regressionsmodellen. Tolka skattningarna i ord. Motivera om tolkningen av skärningen är relevant. (2p)
- På nästa sida finns en reducerad utskrift för den skattade regressionsmodellen i Minitab. Komplettera den saknade informationen i kolumnerna DF, Adj SS (vanliga SS), Adj MS (vanliga MS) och F-value. Redovisa dina beräkningar för dina kompletterande värden. (2p)

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value
Regression				
Error		83,81		
Total		253,88		

- c) Testa på 5 % signifikansnivå om lutningen i regressionsmodellen är signifikant skild från noll. Dra fullständig slutsats av ditt test i ord. (3p)
- d) Beräkna förklaringsgraden från utskriften ovan och tolka den i ord. (1p)

Uppgift 2 (4p)

Följande tabell visar prisutveckling och försäljningsvärde (pris*kvantitet) i basårets priser för ett tillverkningsföretags produkter A och B samt KPI för åren 2013 till 2016:

	PRODUKT A		PRODUKT B		KPI
År	Pris/st	Försäljningsvärde	Pris/st	Försäljningsvärde	
2013	350	13,7	720	35,4	314,1
2014	360	14,5	730	32,6	313,5
2015	358	14,2	725	38,4	313,4
2016	364	14,8	722	42	316,4

- a) Beräkna ett relativprisindex för produkt A:s prisutveckling jämfört med KPI. Använd basår 2013. (1p)
- b) Beräkna ett sammansatt prisindex för produkt A och B. Använd basår 2013. Har Laspeyres eller Paasches viktsystem använts för att beräkna produkternas försäljningsvärden för varje år? Motivera. (3p)

Uppgift 3 (4p)

Ett lands myndighet samlade in antalet ankomna människor till landet (i miljoner) kvartalsvis fr.o.m. kvartal 1 år 2010 t.o.m. kvartal 4 år 2013. I syfte att prognosticera framtida antalet ankomna människor, så skattade man en linjär regressionsmodell (Modell 1) för att förklara antalet ankomna människor med en linjär tidstrend, $t=1,2,\dots,16$, samt dummyvariabler D_1 , D_2 och D_3 för respektive kvartal 1, 2 och 3 för att modellera säsongvariation mellan kvartalen. Dessutom skattade man en enkel linjär regressionsmodell (Modell 2) med endast en linjär tidstrend som jämförelse för att förklara antalet ankomna människor.

På nästföljande 2 sidor finner du utskrifter för dem skattade regressionsmodellerna från Minitab.

- a) Testa på 5 % signifikansnivå om man kan dra slutsatsen att säsongvariationen i antalet ankomna människor över kvartalen är signifikant. Sätt upp fullständiga hypoteser för ditt test och dra fullständig slutsats av ditt test i ord. (2p)
- b) Oavsett slutsats i uppgift a), beräkna en prognos med tillhörande 99 % prognosintervall för antalet ankomna människor i kvartal 1 år 2014 med hjälp av Modell 1. Tolk prognosen och prognosintervallet i ord. Distance value = 0.5625. (2p)

MODELL 1:

Regression Analysis: Ankomna versus t; D1; D2; D3

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	4	10996,8	2749,20	42,47	0,000
t	1	1776,6	1776,61	27,44	0,000
D1	1	367,5	367,49	5,68	0,036
D2	1	7389,0	7389,00	114,13	0,000
D3	1	1019,1	1019,11	15,74	0,002
Error	11	712,1	64,74		
Total	15	11708,9			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
8,04610	93,92%	91,71%	86,49%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	247,44	6,03	41,00	0,000	
t	2,356	0,450	5,24	0,000	1,06
D1	-13,93	5,85	-2,38	0,036	1,58
D2	-61,54	5,76	-10,68	0,000	1,54
D3	-22,64	5,71	-3,97	0,002	1,51

Regression Equation

$$\text{Ankomna} = 247,44 + 2,356 t - 13,93 D1 - 61,54 D2 - 22,64 D3$$

Fits and Diagnostics for Unusual Observations

Obs	Ankomna	Fit	Resid	Std Resid
2	204,00	190,61	13,39	2,08 R

R Large residual

MODELL 2:

Regression Analysis: Ankomna versus t

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	2725	2724,7	4,25	0,058
t	1	2725	2724,7	4,25	0,058
Error	14	8984	641,7		
Total	15	11709			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
25,3324	23,27%	17,79%	0,00%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	218,9	13,3	16,48	0,000	
T	2,83	1,37	2,06	0,058	1,00

Regression Equation

$$\text{Ankomna} = 218,9 + 2,83 t$$

Uppgift 4 (4p)

Nu ska den skattade regressionsmodellen för antalet ankomna människor till landet i uppgift 3 (Modell 1) jämföras med resultaten på nästföljande sidor från Minitab för klassiska komponentuppdelningar och Holt-Winters metod för antalet ankomna människor till landet.

- a) Rangordna de 5 modellerna gentemot varandra utifrån hur bra respektive modell skattar datamaterialet. Vilken av modellerna verkar vara bäst? Motivera. (1p)
- b) Tolka alla säsongskomponenter för både additiv och multiplikativ klassisk komponentuppdelning. (2p)
- c) Beräkna en prognos för kvartal 1 år 2014 med hjälp av den additiva klassiska komponentuppdelningsmetoden. Tolka prognosen i ord. (1p)

Time Series Decomposition for Ankomna

Method

Model type Additive Model

Data Ankomna

Length 16

NMissing 0

Fitted Trend Equation

$$Y_t = 223,11 + 2,332 \times t$$

Seasonal Indices

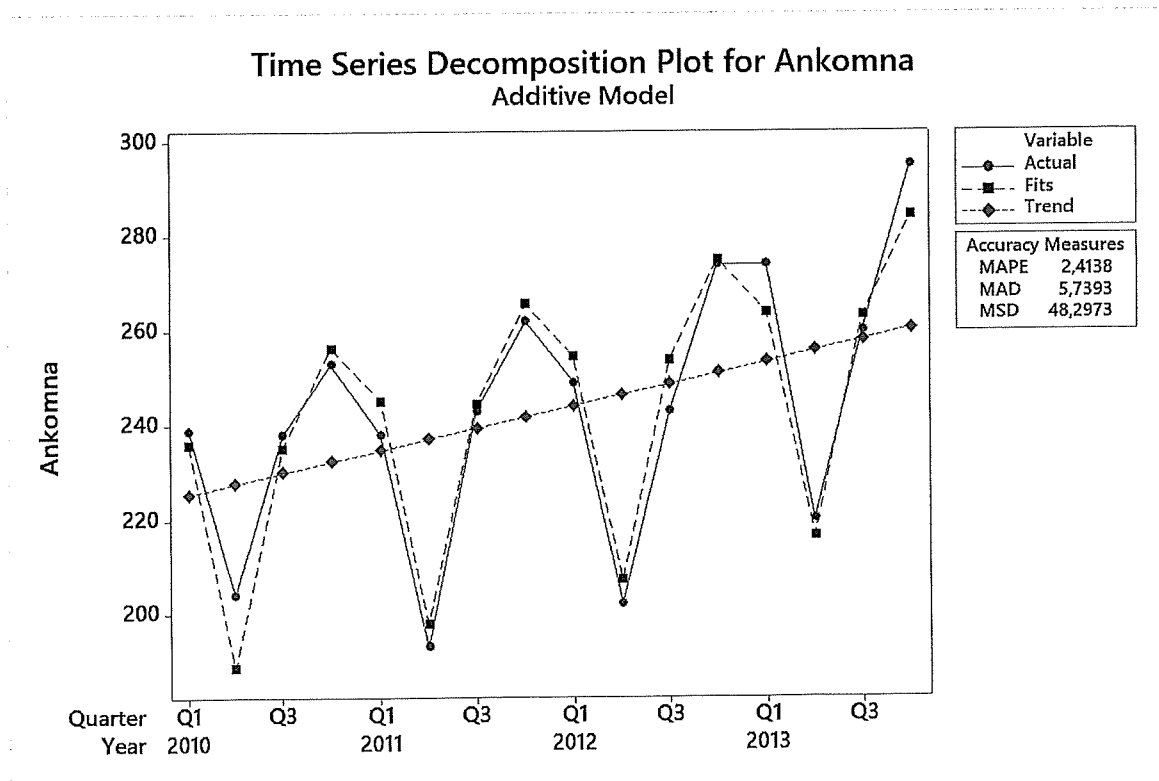
Period	Index
1	10,375
2	-39,250
3	5,000
4	23,875

Accuracy Measures

MAPE 2,4138

MAD 5,7393

MSD 48,2973



Time Series Decomposition for Ankomna

Method

Model type Multiplicative Model

Data Ankomna

Length 16

NMissing 0

Fitted Trend Equation

$$Y_t = 223,37 + 2,297 \times t$$

Seasonal Indices

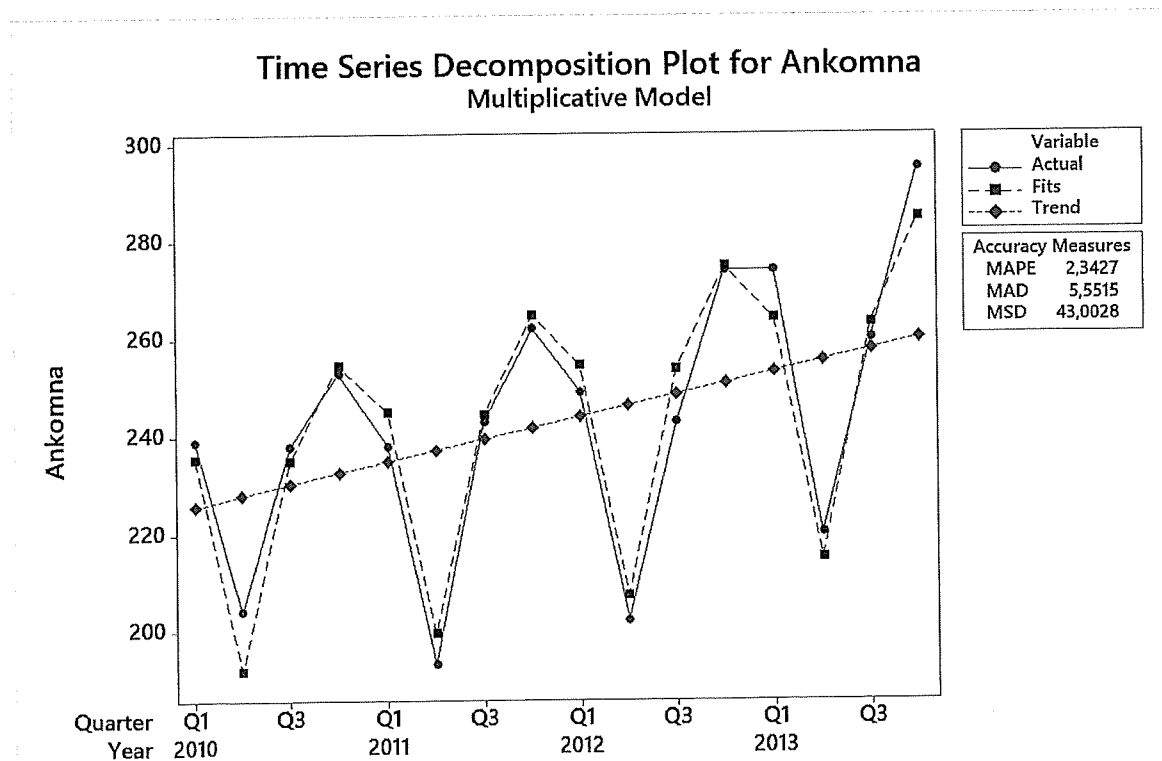
Period	Index
1	1,04304
2	0,84089
3	1,02100
4	1,09507

Accuracy Measures

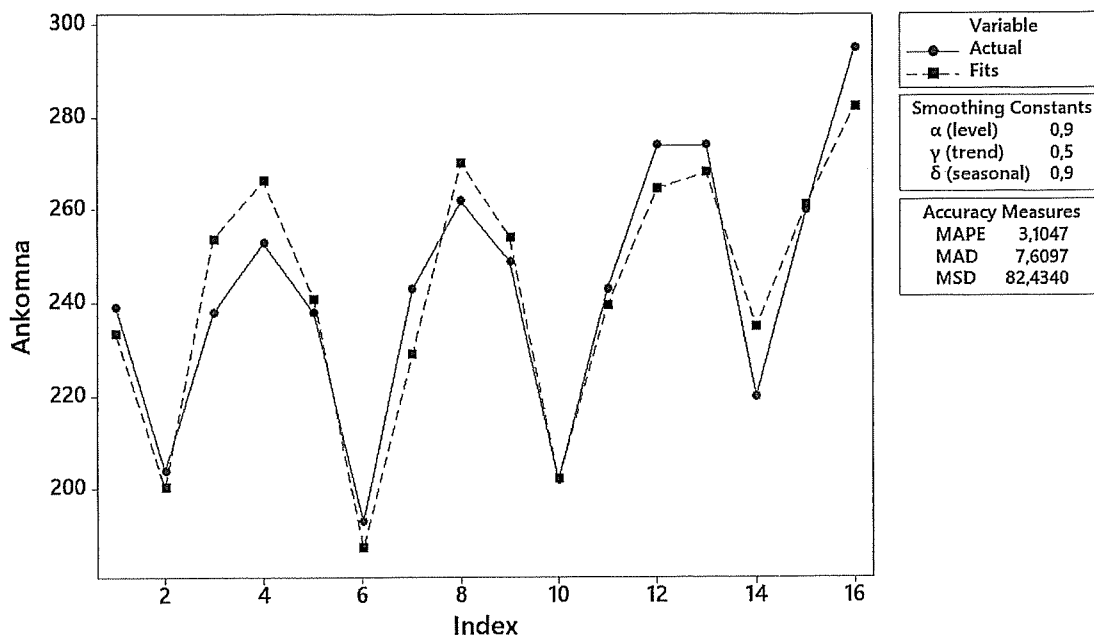
MAPE 2,3427

MAD 5,5515

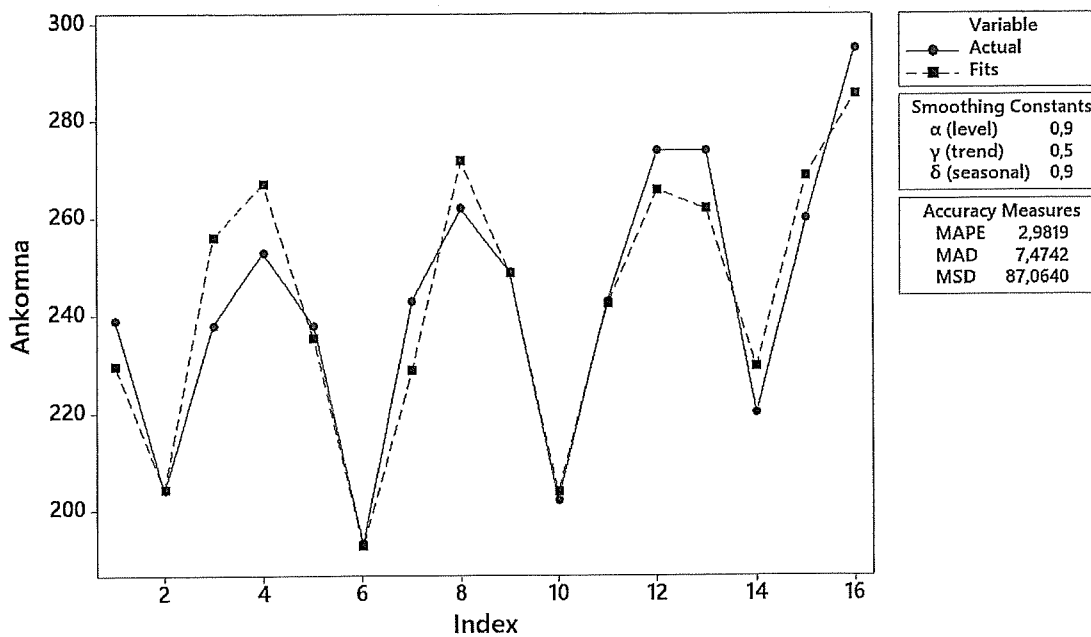
MSD 43,0028



Winters' Method Plot for Ankomna
Additive Method



Winters' Method Plot for Ankomna
Multiplicative Method



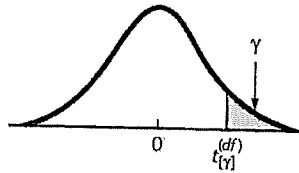


TABLE A2 A t-Table: Values of $t_{(df)}^{(\gamma)}$

<i>df</i>	$t_{[.10]}^{(df)}$	$t_{[.05]}^{(df)}$	$t_{[.025]}^{(df)}$	$t_{[.01]}^{(df)}$	$t_{[.005]}^{(df)}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
inf.	1.282	1.645	1.960	2.326	2.576

Source: From "Table of Percentage Points of the t-Distribution," by Maxine Merrington, *Biometrika* 32 (1941), 300. Reproduced by permission of the *Biometrika* Trustees.

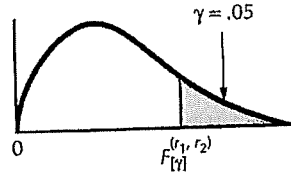


TABLE A3 An F-Table: Values of $F_{[\gamma]}$

Denominator Degrees of Freedom, r_2	Numerator Degrees of Freedom, r_1								
	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Source: From "Tables of Percentage Points of the Inverted Beta (F)-Distribution," by Maxine Merrington and Catherine M. Thompson, *Biometrika* 33 (1943), 73-88. Reproduced by permission of the *Biometrika* Trustees.