F17

Tidsserier

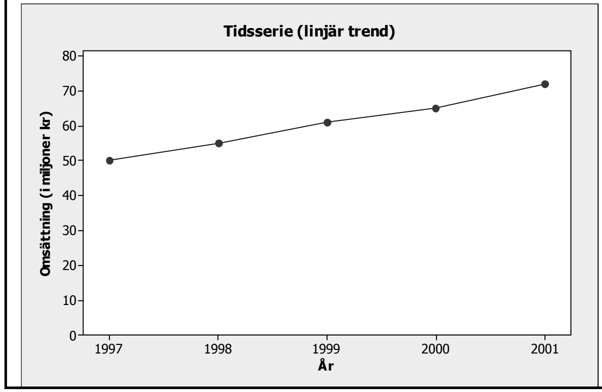
Säsongrensning

Tidsserier, forts från F16

- Vi har en variabel som varierar över tiden
  - Ex folkmängd, omsättning, antal anställda … (beroende variabeln/undersökningsvariabeln)
  - Vi studerar den varje år, halvår, tertial (4 mån), kvartal, månad, vecka, dag
  - Tid kan nu ses som den oberoende variabeln/bakgrundsvariabeln
- Vi har tittat specifikt på två typer av trender
  - Linjär trend
  - Exponentiell trend

Linjär trend

- Den absoluta förändringen av Y då t ökar en enhet är densamma för alla värden på t

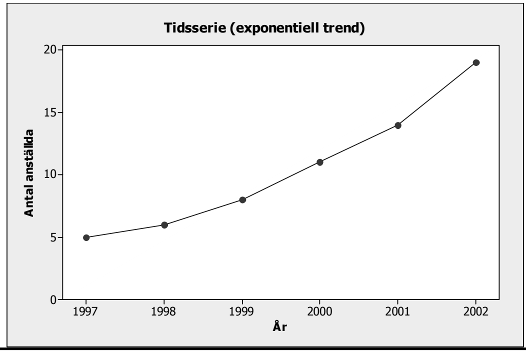


$Y = a + bt$

Ex 702 K&W

Exponentiell trend

- Den *absoluta* förändringen av Y beror på *var* i intervallet vi är
- Den *procentuella* förändringen av Y är densamma för varje ökning av t med 1 enhet



$Y = ab^t$

Ex 703 K&W

Kodning av t

- Om vi centrerar t så blir beräkningarna enklare, oavsett om vi använder en linjär modell eller en exponentiell modell

1

År (x)	1998	1999	2000	2001	2002
t	-2	-1	0	1	2

2

År (x)	1997	1998	1999	2000	2001	2002
t	-5	-3	-1	1	3	5

eller

3

t	-2,5	-1,5	-0,5	0,5	1,5	2,5

$$b = \frac{\sum_{i=1}^{n} y_i t_i}{\sum_{i=1}^{n} t_i^2}$$

$$a = \bar{y}$$

Se K&W s186

Linjär trend

1. Koda om t
2. Använd MK-metoden och skatta a och b med förenklade formler
3. Tolkning av b: genomsnittlig ökning per t-enhet
   1. Om vi använt t-kodning 1 eller 3 ovan: genomsnittlig ökning per år
   2. Om vi använt kodning 2 ovan: genomsnittlig ökning per halvår. För att få genomsnittlig ökning per år: ta 2b

Exponentiell trend

- Vi har ej en linjär funktion
- Gör funktionen linjär genom att logaritmera

$$\log y = \log\left(a \times b^{t}\right)$$

$$\log y = \log a + t \times \log b$$

$$\log y = y' \quad \log a = a' \quad \log b = b'$$

$$y' = a' + b' \times t$$

- Skatta a′ och b′ genom MK-metoden på de loggade y-värdena
- Ta antilogen för lösningarna för att få a och b

Exponentiell trend

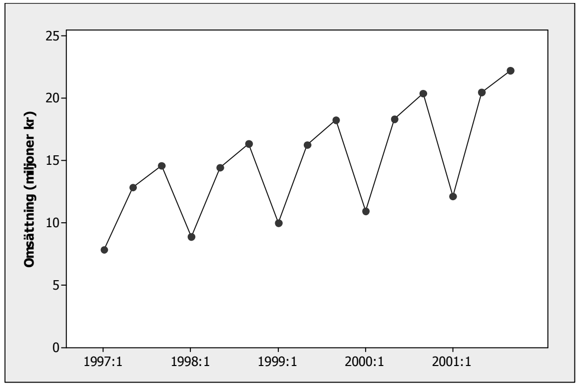
1. Koda om t
2. Logaritmera y-värdena
3. Använd MK-metoden och skatta a′ och b′ med förenklade formler
4. Ta antilog av a′ och b′ för att erhålla a och b
5. Tolkning av b: genomsnittlig procentuell ökning per t-enhet
   1. Om vi använt t-kodning 1 eller 3 ovan: genomsnittlig procentuell ökning per år
   2. Om vi använt kodning 2 ovan: genomsnittlig procentuell ökning per halvår. För att få genomsnittlig procentuell ökning per år: ta b^2

Variationsorsaker

- Trend
- Konjunktur
- Säsong
- Slump

Säsonganalys

- Vi antar nu att utöver den slumpmässiga variationen kring trenden finns också en säsongberoende variation



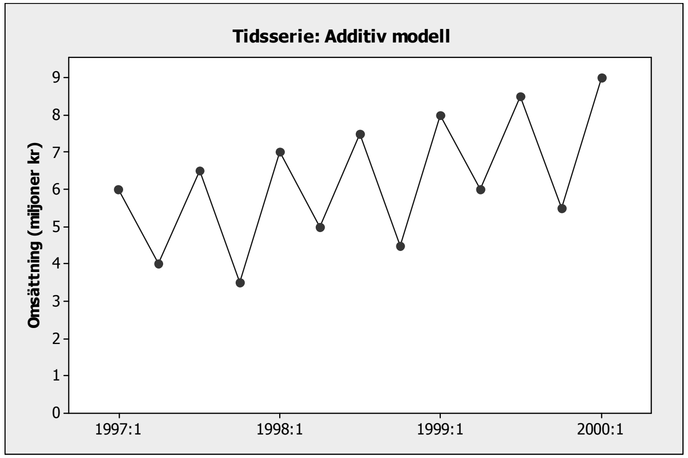
K&W s 196

Additiv modell

- Säsongvariationerna för en viss säsong är desamma i *absoluta tal* för varje år och oberoende av trendens storlek

$$y_{ik} = T_{ik} + S_k + \varepsilon_{ik}$$

Exempel additiv modell

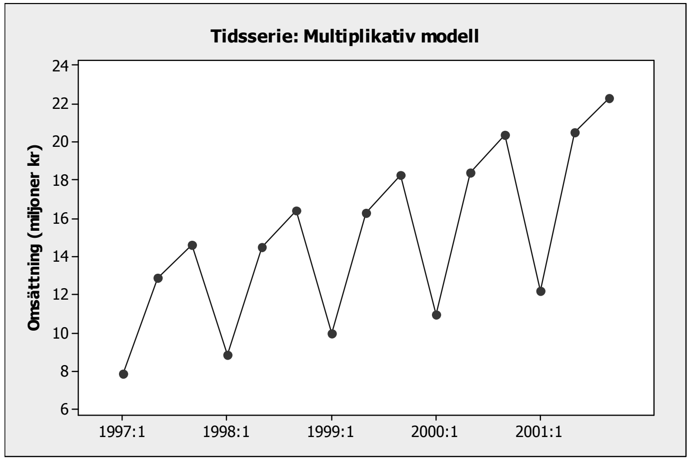


Multiplikativ modell

- Säsongvariationerna för en given säsong avviker med samma procentsats från trenden varje år. De absoluta avvikelserna beror på trendens storlek

$$y_{ik} = T_{ik} \times S_k \times \varepsilon_{ik}$$

Exempel multiplikativ modell



K&W s 196

Vad vill vi göra?

- Vi vill skatta trenden
  - Ett trendvärde som inte är påverkat av säsongvariation
- Vi vill skatta säsongkomponenter
  - Hur stor del av omsättningen (exempelvis) beror på den säsongmässiga variationen

1. Vi säsongrensar först tidsserien för att få trendvärden oberoende av säsong. Vi skattar därmed också säsongkomponenter.
2. Vi skattar trenden (linjär/exponentiell trend modell) på de säsongrensade värdena

Säsongrensning

- Vi kan få flera skattningar av varje säsongkomponent – vi skattar ett medelvärde för varje säsong – S_k
- Additiv modell: Villkor: $\sum_{k=1}^{K} S_k = 0$
- Multiplikativ modell: Villkor: $\sum_{k=1}^{K} S_k = K$
  - Där K är antal säsonger (ex 4 vid kvartalsdata)

Exempel: Uppskattning av trendvärden genom glidande medelvärden

(7,9+12,9+14,6)

(12,9+14,6+8,9)

35,4/3

Period		Obs. värde	3-punktsumma	Medelvärde
1997:	I	7,9		
	II	12,9	35,4	11,800
	III	14,6	36,4	12,133
1998:	I	8,9	38,0	12,667
	II	14,5	39,8	13,267
	III	16,4	40,9	13,633
1999:	I	10,0	42,7	14,233
	II	16,3	44,6	14,867
	III	18,3	45,6	15,200
2000:	I	11,0	47,7	15,900
	II	18,4	49,8	16,600
	III	20,4	51,0	17,000
2001:	I	12,2	53,1	17,700
	II	20,5	55,0	18,333
	III	22,3		

Exempel multiplikativ modell: Uppskattning av säsongkomponenter

$$\frac{12{,}9}{11{,}800} \times 100 \approx 109{,}3$$

- För varje säsong (tertial, kvartal etc.) beräknar vi kvoten mellan det faktiska värdet och trendvärdet

	Tertial		
År	I	II	III
1997		109,3	120,3
1998	70,3	109,3	120,3
1999	70,3	109,6	120,4
2000	69,2	110,8	120,0
2001	68,9	111,8	
Medelvärde	69,675	110,16	120,25
Säsongindex	69,7	110,1	120,2

$$69{,}675 \times 0{,}9997167 \approx 69{,}7$$

Medelvärdena summerar ej till 300 - vi justerar genom att multiplicera varje medelvärde med kvoten

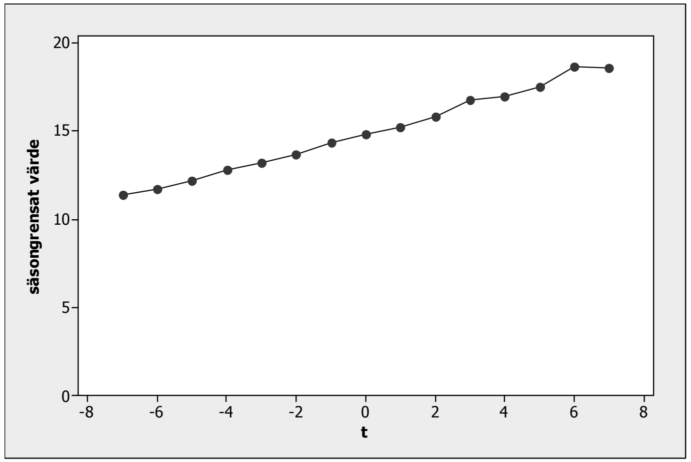
$$\frac{300}{69{,}675 + 110{,}16 + 120{,}25} \approx 0{,}9997167$$

Tolkning av säsongkomponenter

Säsongindex	69,7	110,1	120,2

- Under första tertialet ligger omsättningen *på grund av att det är lågsäsong* drygt 30 procent under det beräknade trendvärdet
- Under tredje tertialet ligger omsättningen *på grund av att det är högsäsong* drygt 20% procent över det beräknade trendvärdet

Säsongrensade värden



Vi kan nu anpassa en trendmodell med MK-metoden

- I detta fall verkar en linjär modell vara bäst. Vi kan koda t så att andra tertialet 1999=0.

```
■ The regression equation is
■ Omsättning = 14,9 + 0,540 t


■ Predictor     Coef  SE Coef       T      P
■ Constant   14,8865   0,0500  297,52  0,000
■ t          0,54043  0,01158   46,67  0,000


■ S = 0,193785   R-Sq = 99,4%   R-Sq(adj) = 99,4%
```

Prognoser

- Omsättning = **14,9 + 0,54 t**
- Säsongkomponenter: **69,7** **110,1** **120,2**

- Säg att vi vill göra en prognos för andra tertialet 2002
- **t=9**
- Prognos för *trendvärde*: Omsättning = 14,9 + 0,54x9 = 19,76
- Prognos för *observerat värde*:
- Omsättning = (14,9 + 0,54x9) x 1,101= 19,841 x 1,101 = 21,76
- Om vi istället hade uppskattat en *exponentiell modell* hade prognosformen sett ut: Omsättning$_{ik}$ = (a + b^{t}) x S$_{k}$

Exempel: Antal sjukdagar (i tusental), se ”Något om säsongrensning…”

Period		Obs. värde	4-punktsumma	2-punktsumma	2-punktsmedelvärde
1988:	IV	103			
1989:	I	86	324		
	II	71	305	629	78,625
	III	64	304	609	76,125
	IV	84	298	602	75,25
1990:	I	85	298	596	74,5
	II	65	298	596	74,5
	III	64	294	592	74
	IV	84	284	578	72,25
1991:	I	81	272	556	69,5
	II	55	253	525	65,625
	III	52			
	IV	65			

Säsongkomponenter för additiv modell

71-78,625=-7,625

	Kvartal			
År	I	II	III	IV
1989		-7,625	-12,125	8,75
1990	10,5	-9,5	-10	11,75
1991	11,5	-10,625		
Medelvärde	11,0	-9,25	-11,0625	10,25
Säsongindex	10,77	-9,48	-11,29	10,02

$$11-\frac{11-9{,}25-11{,}0625+10{,}25}{4}=10{,}77$$

Under kvartal 1 ligger antal sjukanmälningar *över trenden* med i genomsnitt 10770 st

Under kvartal 2 *under trenden* med i genomsnitt 9480 st

Prognoser

- Linjär trend:
  - Antal $sjukdagar_{ik}=a+bt+S_k$
- Exponentiell trend:
  - Antal $sjukdagar_{ik}=axb^t+S_k$

Vikter för glidande medelvärden

- Kvartalsdata: $\frac{1}{8},\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{8}$
  - 5-leds
- Halvårsdata: $\frac{1}{4},\frac{1}{2},\frac{1}{4}$
  - 3-leds
- Tertialdata: $\frac{1}{3},\frac{1}{3},\frac{1}{3}$
  - 3-leds
- Veckodata: $\frac{1}{7},\frac{1}{7},\frac{1}{7},\frac{1}{7},\frac{1}{7},\frac{1}{7},\frac{1}{7}$
  - 7-leds