Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences

CAS Datenanalyse

Zeitreihenanalyse
Kapitel 2: Trend- und Saisonbereinigung

Prof. Dr. Raúl Gimeno
FRM, CAIA, PRM

Inhaltsverzeichnis

2.1 Phasendurchschnittsverfahren

Zur Bestimmung der Saisonkomponente einer unterjährigen Zeitreihe (y_t) schaltet man die glatte Komponente (g_t) vorab aus.

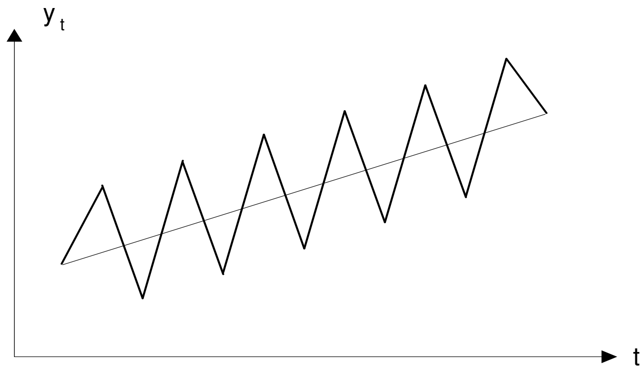
Trendbereinigte Zeitreihenwerte beim additiven Modell:

$$d_t = y_t - g_t = s_t + u_t$$

→ nur noch die Saison- und Restkomponente

Voraussetzung: saisonale Ausschläge sind unabhängig vom Trend

→ Saisonausschläge mit konstanter Amplitude



Bestimmung der Saisonkomponente

Doppelindizierung: Jahr i, Phase (z.B. Monat, Quartal) j

Trendbereinigte Zeitreihe: $d_{ij} = y_{ij} - g_{ij} = s_j + u_{ij}, \quad i = 1, 2, \ldots, k_j; \quad j = 1, 2, \ldots, p$

k_j: Anzahl der Jahre für die Phase j

p: Anzahl der Phasen (bei Quartalsdaten: p=4, bei Monatsdaten: p=12)

g_{ij}: zentrierter gleitender Durchschnitt

Unnormierte Saisonkomponente für Phase j: $s_j^* = \frac{1}{k_j} \sum_i d_{ij}$

Die unnormierten Saisonkomponenten s_j^* summieren sich nicht zu null → erschwert die Interpretation einer Phase als saisonal unter- oder überdurchschnittlich.

Normierungseigenschaft: $\sum_{j=1}^{p} s_j = 0$

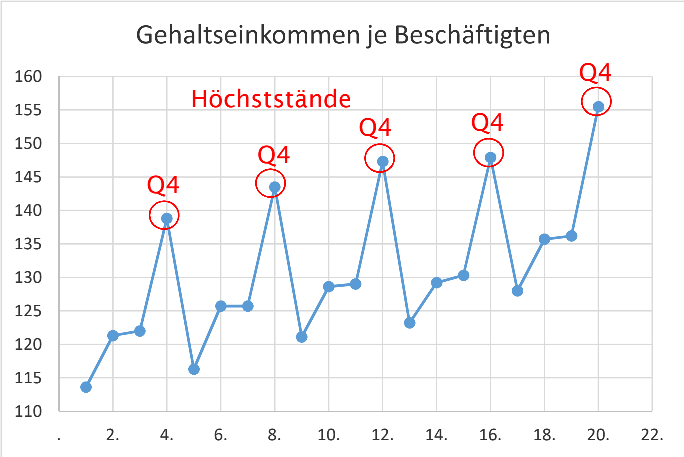
Normierte Saisonkomponente: $s_j = s_j^* - \bar{d}$ $\qquad \bar{d} = \frac{1}{p} \sum_{j=1}^{p} s_j^*$

Saisonbereinigte Zeitreihe: $y_{ij}^* = y_{ij} - s_j$

Bestimmung der Saisonkomponente: Beispiel

Im Zeitreihendiagramm der Gehaltseinkommen je Beschäftigten ist ein klares Saisonmuster erkennbar. Der Index weist jahreszeitlich bedingt jeweils im Q1 eines Jahres einen Tiefstand und im Q4 ein Hoch aus.

Zeitreihenzzerlegung auf der Grundlage des additiven Modells.



Bestimmung der Saisonkomponente: Beispiel

Jahr (i)	Quartal (j)	y_t	$\bar{y}_t^4$	d_{ij}
1986 (1)	Q1 (1)	113.6		
1986 (1)	Q2 (2)	121.3		
1986 (1)	Q3 (3)	122	**124.26**	-2.26
1986 (1)	Q4 (4)	138.8	125.15	13.65
1987 (2)	Q1 (1)	116.3	126.16	-9.86
1987 (2)	Q2 (2)	125.7	127.21	-1.51
1987 (2)	Q3 (3)	125.7	128.40	-2.70
1987 (2)	Q4 (4)	143.5	129.36	14.14
1988 (3)	Q1 (1)	121.1	130.14	-9.04
1988 (3)	Q2 (2)	128.6	131.03	-2.43
1988 (3)	Q3 (3)	129	131.76	-2.76
1988 (3)	Q4 (4)	147.3	132.10	15.20
1989 (4)	Q1 (1)	123.2	132.34	-9.14
1989 (4)	Q2 (2)	129.2	132.58	-3.38
1989 (4)	Q3 (3)	130.3	133.25	-2.95
1989 (4)	Q4 (4)	147.9	134.66	13.24
1990 (5)	Q1 (1)	128	136.21	-8.21
1990 (5)	Q2 (2)	135.7	137.90	-2.20
1990 (5)	Q3 (3)	136.2		
1990 (5)	Q4 (4)	155.5		

Zentrierte gleitende Durchschnitte 4-ter Ordnung

$$\bar{y}_t^4 = \frac{1}{4}\left(\frac{1}{2}y_{t-2} + y_{t-1} + y_t + y_{t+1} + \frac{1}{2}y_{t+2}\right)$$

$$\bar{y}_{86,3} = \frac{1}{4}\left(\frac{113.6}{2} + 121.3 + 122 + 138.8 + \frac{116.3}{2}\right) = 124.26$$

$$d_{86,3} = y_{86,3} - \bar{y}_{86}^4 = 122 - 124.26 = -2.26$$

Unnormierter Saisonfaktor s_1^* für Q1

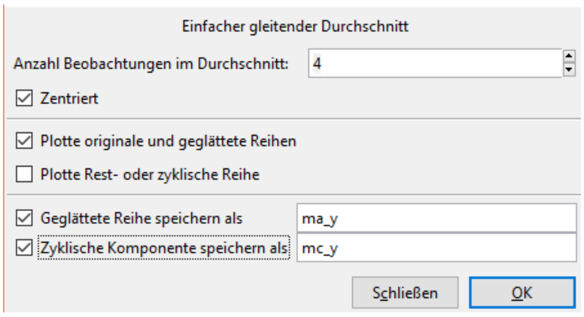
s_1^* = (-9.86 -9.04 -9.14 -8.21)/4 = -9.1

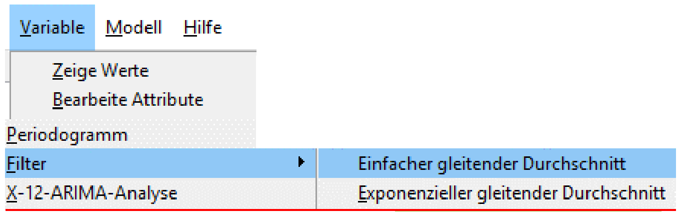
Saisonfaktor s_3^* für Q3:

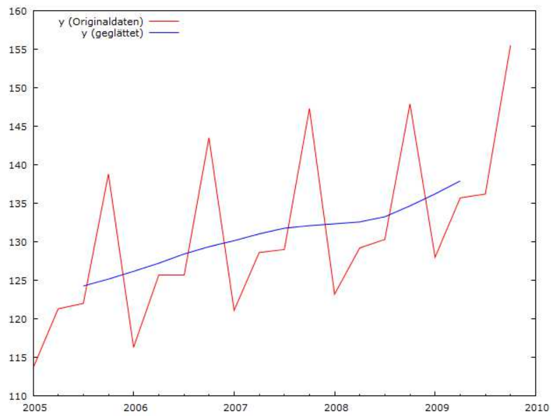
s_3^* = (-2.26 -2.7 -2.76 -2.95)/4 = -2.67

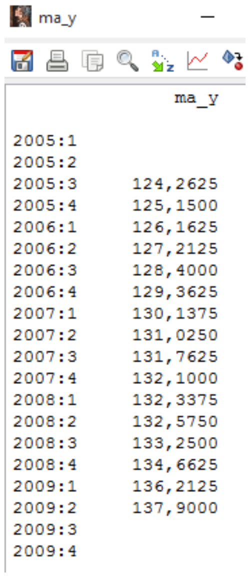
Gretl Output

Menu: Variable/Filter/Einfacher gleitender Durchschnitt









	ma_y
2005:1	
2005:2	
2005:3	124,2625
2005:4	125,1500
2006:1	126,1625
2006:2	127,2125
2006:3	128,4000
2006:4	129,3625
2007:1	130,1375
2007:2	131,0250
2007:3	131,7625
2007:4	132,1000
2008:1	132,3375
2008:2	132,5750
2008:3	133,2500
2008:4	134,6625
2009:1	136,2125
2009:2	137,9000
2009:3	
2009:4	

	mc_y
2005:1	
2005:2	
2005:3	-2,2625
2005:4	13,6500
2006:1	-9,8625
2006:2	-1,5125
2006:3	-2,7000
2006:4	14,1375
2007:1	-9,0375
2007:2	-2,4250
2007:3	-2,7625
2007:4	15,2000
2008:1	-9,1375
2008:2	-3,3750
2008:3	-2,9500
2008:4	13,2375
2009:1	-8,2125
2009:2	-2,2000
2009:3	
2009:4	

Bestimmung der Saisonkomponente

Glatte Komponente → zentrierte gleitende Durchschnitte der Ordnung 4:

$$\bar{y}_t^4 = \frac{1}{4}\left(\frac{1}{2}y_{t-2} + y_{t-1} + y_t + y_{t+1} + \frac{1}{2}y_{t+2}\right)$$

Aus den berechneten trendbereinigten Werte d_{ij} erhält man die unnormierten Saisonkomponente mit k_j = 4 (4 Jahre):

J.Q	d_{ij}
86.Q3	-2.26
86.Q4	13.65
87.Q1	-9.86
87.Q2	-1.51
87.Q3	-2.70
87.Q4	14.14
88.Q1	-9.04
88.Q2	-2.43
88.Q3	-2.76
88.Q4	15.20
89.Q1	-9.14
89.Q2	-3.38
89.Q3	-2.95
89.Q4	13.24
90.Q1	-8.21
90.Q2	-2.20

1\. Quartal $s_1^* = \frac{1}{4}\sum_{i=1}^{4} d_{i1} = \frac{1}{4}(\overset{87}{-9.9}\overset{88}{-9.0}\overset{89}{-9.1}\overset{90}{-8.2}) = -\frac{36.25}{4} = -9.06$

2\. Quartal $s_2^* = \frac{1}{4}\sum_{i=1}^{4} d_{i2} = \frac{1}{4}(-1.5-2.4-3.4-2.2) = -\frac{9.51}{4} = -2.38$

3\. Quartal $s_3^* = \frac{1}{4}\sum_{i=1}^{4} d_{i3} = \frac{1}{4}(-2.3-2.7-2.8-3.0) = -\frac{10.68}{4} = -2.67$

4\. Quartal $s_4^* = \frac{1}{4}\sum_{i=1}^{4} d_{i4} = \frac{1}{4}(13.6+14.1+15.2+13.2) = \frac{56.23}{4} = 14.06$

Bestimmung der Saisonkomponente

Durchschnitt der unnormierten Saisonziffern
für p = 4 (Quartalszahlen):

$$\bar{d} = \frac{1}{p}\sum_{j=1}^{p} s_j^* = \frac{1}{4}[(-9.06)+(-2.38)+(-2.67)+14.06] = \frac{1}{4}(-0.05) = -0.013$$

Normierte Saisonkomponenten: $s_j = s_j^* - \bar{d}$

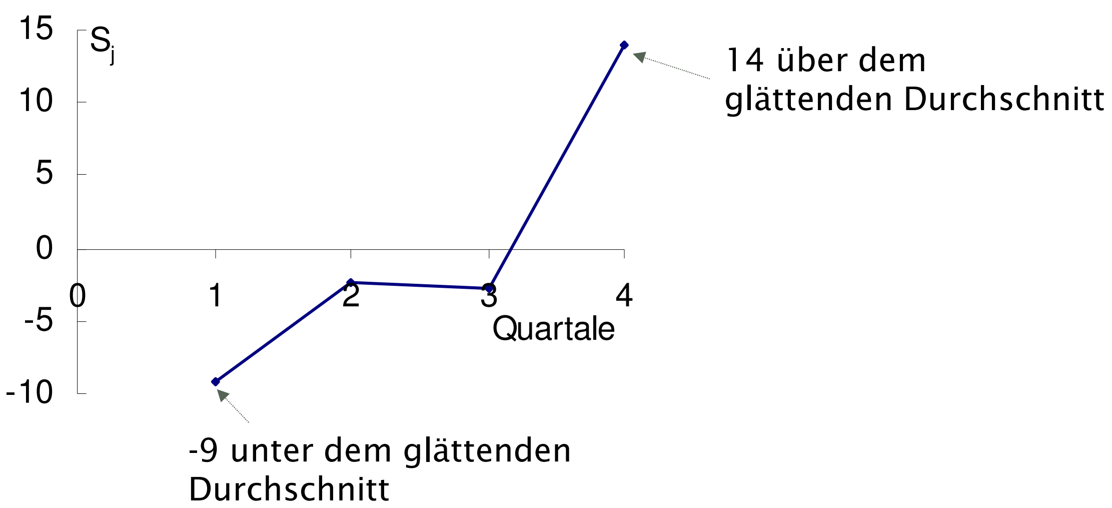
$s_1 = s_1^* - \bar{d}$ = -9.06 -(-0.013) = -9.049
$s_2 = s_2^* - \bar{d}$ = -2.38 -(-0.013) = -2.365
$s_3 = s_3^* - \bar{d}$ = -2.67- (-0.013) = -2.655
$s_4 = s_4^* - \bar{d}$ = 14.06- (-0.013) = 14.070

$\sum_{j=1}^{p} s_j = 0$ Normierungseigenschaft

Bestimmung der Saisonkomponente

Das Saisonprofil s_j^* gibt die Grössenordnung des saisonalen Einflusses in der Zeitreihe der Gehälter je Beschäftigten grafisch wieder.

Saisonprofil der Gehälter je Beschäftigten



Bestimmung der Saisonkomponente: Beispiel

Saisonbereinigte Zeitreihe (y_t^*) der Löhne und Gehälter je Beschäftigten für den gesamten Beobachtungszeitraum

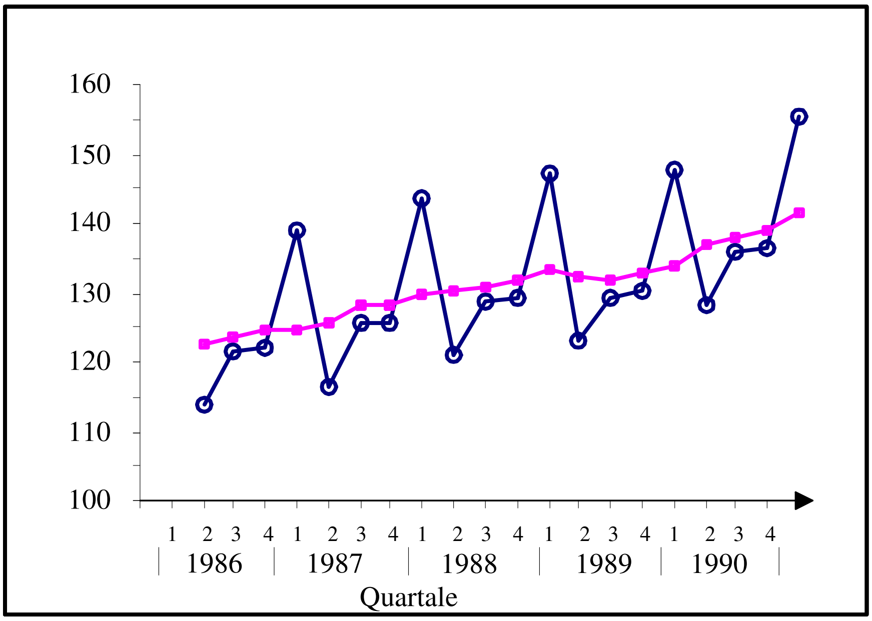
Jahr (i)	Quartal (j)	y_t	s_i	(y_t^*)
1986 (1)	Q1 (1)	113.6	-9.06	122.66
1986 (1)	Q2 (2)	121.3	-2.38	123.68
1986 (1)	Q3 (3)	122	-2.67	124.67
1986 (1)	Q4 (4)	138.8	14.06	124.74
1987 (2)	Q1 (1)	116.3	-9.06	125.36
1987 (2)	Q2 (2)	125.7	-2.38	128.08
1987 (2)	Q3 (3)	125.7	-2.67	128.37
1987 (2)	Q4 (4)	143.5	14.06	129.44
1988 (3)	Q1 (1)	121.1	-9.06	130.16
1988 (3)	Q2 (2)	128.6	-2.38	130.98
1988 (3)	Q3 (3)	129	-2.67	131.67
1988 (3)	Q4 (4)	147.3	14.06	133.24
1989 (4)	Q1 (1)	123.2	-9.06	132.26
1989 (4)	Q2 (2)	129.2	-2.38	131.58
1989 (4)	Q3 (3)	130.3	-2.67	132.97
1989 (4)	Q4 (4)	147.9	14.06	133.84
1990 (5)	Q1 (1)	128	-9.06	137.06
1990 (5)	Q2 (2)	135.7	-2.38	138.08
1990 (5)	Q3 (3)	136.2	-2.67	138.87
1990 (5)	Q4 (4)	155.5	14.06	141.44

$= y_{86,1} - s_1$

$= y_{87,3} - s_3$

Additives Modell: Saisonbereinigte Zeitreihe

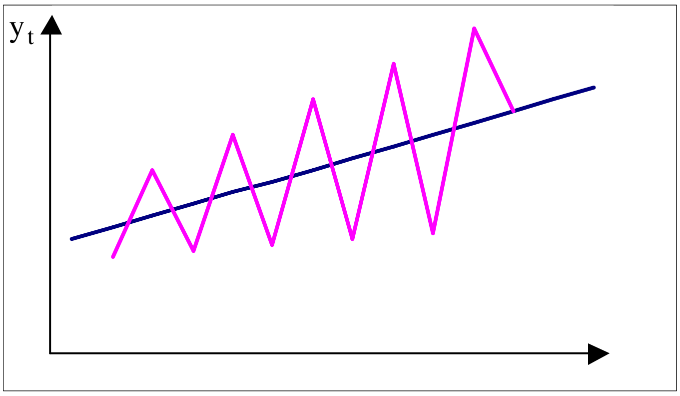
Gehälter je Beschäftigten mit saisonbereinigter Zeitreihe



Multiplikatives Modell

Grundmodell: Saisonausschläge nehmen mit steigendem Trend im Mittel proportional zu.

→ Saisonausschläge mit proportional zunehmender Amplitude



Die Saisonkomponente kann wie im additiven Modell bestimmt werden, wenn man die originäre Zeitreihe (y_t) durch die logarithmierte Zeitreihe ($\log y_t$) ersetzt.

Multiplikatives Modell

Bestimmung der Saisonkomponente und der saisonbereinigten Zeitreihe:

Trendbereinigte Werte: $d_{ij} = \dfrac{y_{ij}}{g_{ij}} = s_j \cdot u_{ij},\ i = 1,2,\dots,k_j, j = 1,2,\dots,p$ → # Phasen

Unnormierte Saisonkomponente: $s_j^* = \dfrac{1}{k_j} \sum_i d_{ij}$

→ wie im additiven Modell unter Verzicht auf eine geometrische Mittelung

Normierungsfaktor (NF): $\bar{d} = p \Big/ \sum_{j=2}^{p} s_j^*$

Normierte Saisonkomponente: $s_j = s_j^* \cdot \bar{d}$

Saisonbereinigte Zeitreihe: $y_{ij}^* = y_{ij} / s_j$

Q1	s_1^*	$s_1 = s_1^* \cdot \bar{d}$	$y_{i1}^* = y_{i1} / s_1$
Q2	s_2^*	$s_2 = s_2^* \cdot \bar{d}$	$y_{i2}^* = y_{i2} / s_2$
Q3	s_3^*	$s_3 = s_3^* \cdot \bar{d}$	$y_{i3}^* = y_{i3} / s_3$
Q4	s_4^*	$s_4 = s_4^* \cdot \bar{d}$	$y_{i4}^* = y_{i4} / s_4$

$\sum_{i=1}^{4} s_i^* \neq 4 \qquad \sum_{i=1}^{4} s_i = 4$

Bestimmung der Saisonkomponente: Beispiel

Jahr (i)	Quartal (j)	y_t	$\bar{y}_t^4$	d_{ij}
1986 (1)	Q1 (1)	113.6		
1986 (1)	Q2 (2)	121.3		
1986 (1)	Q3 (3)	122	**124.26**	0.98
1986 (1)	Q4 (4)	138.8	125.15	1.11
1987 (2)	Q1 (1)	116.3	126.16	0.92
1987 (2)	Q2 (2)	125.7	127.21	0.99
1987 (2)	Q3 (3)	125.7	128.40	0.98
1987 (2)	Q4 (4)	143.5	129.36	1.11
1988 (3)	Q1 (1)	121.1	130.14	0.93
1988 (3)	Q2 (2)	128.6	131.03	0.98
1988 (3)	Q3 (3)	129	131.76	0.98
1988 (3)	Q4 (4)	147.3	132.10	1.12
1989 (4)	Q1 (1)	123.2	132.34	0.93
1989 (4)	Q2 (2)	129.2	132.58	0.97
1989 (4)	Q3 (3)	130.3	133.25	0.98
1989 (4)	Q4 (4)	147.9	134.66	1.10
1990 (5)	Q1 (1)	128	136.21	0.94
1990 (5)	Q2 (2)	135.7	137.90	0.98
1990 (5)	Q3 (3)	136.2		
1990 (5)	Q4 (4)	155.5		

Zentrierte gleitende Durchschnitte 4-ter Ordnung

$$\bar{y}_t^4 = \frac{1}{4}\left(\frac{1}{2}y_{t-2} + y_{t-1} + y_t + y_{t+1} + \frac{1}{2}y_{t+2}\right)$$

$$\bar{y}_{86,3} = \frac{1}{4}\left(\frac{113.6}{2} + 121.3 + 122 + 138.8 + \frac{116.3}{2}\right) = 124.26$$

$$d_{86,3} = y_{86,3} / \bar{y}_{86}^4 = 122/124.26 = 0.98$$

Unnormierter Saisonfaktor s_1^*

s_1^* = (0.92+ 0.93 +0.93+0.94)/4 = 0.93

Saisonfaktor s_3^* für Q3:

s_3^* = 0.98 + 0.98 0.98 +0.98)/4 = 0.98

Bestimmung der Saisonkomponente

Glatte Komponente = zentrierte gleitende Durchschnitte der Ordnung 4:

$$\bar{y}_t^4 = \frac{1}{4}\left(\frac{1}{2}y_{t-2} + y_{t-1} + y_t + y_{t+1} + \frac{1}{2}y_{t+2}\right)$$

Aus den berechneten trendbereinigten Werte d_{ij} erhält man die unnormierten Saisonkomponente mit k_j = 4 (4 Jahre):

87 88 89 90

1. Quartal $s_1^* = \frac{1}{4}\sum_{i=2}^{4} d_{i1} = \frac{1}{4}(0.92 + 0.93 + 0.93 + 0.94) = \frac{3.723}{4} = 0.9308$

2. Quartal $s_2^* = \frac{1}{4}\sum_{i=2}^{4} d_{i2} = \frac{1}{4}(0.99 + 0.98 + 0.97 + 0.98) = \frac{3.9282}{4} = 0.982$

3. Quartal $s_3^* = \frac{1}{4}\sum_{i=1}^{4} d_{i3} = \frac{1}{4}(0.98 + 0.98 + 0.98 + 0.98) = \frac{3.9177}{4} = 0.979$

4. Quartal $s_4^* = \frac{1}{4}\sum_{i=1}^{4} d_{i4} = \frac{1}{4}(1.11 + 1.11 + 1.12 + 1.10) = \frac{4.4317}{4} = 1.1079$

Bestimmung der Saisonkomponente

Durchschnitt der unnormierten Saisonziffern:

$$\bar{d} = \frac{1}{p}\sum_{j=1}^{p} s_j^* = \frac{1}{4}[0.931 + 0.982 + 0.979 + 1.108] = \frac{1}{4}(4.0002) = 1$$

Normierte Saisonkomponenten: $s_j = s_j^* \cdot \bar{d}$

$$s_1 = s_1^* \cdot \bar{d} \quad = 0.928 \times 1 \quad = 0.928$$

$$s_2 = s_2^* \cdot \bar{d} \quad = 0.982 \times 1 \quad = 0.982$$

$$s_3 = s_3^* \cdot \bar{d} \quad = 0.9794 \times 1 \quad = 0.9794$$

$$s_4 = s_4^* \cdot \bar{d} \quad = 1.108 \times 1 \quad = 1.108$$

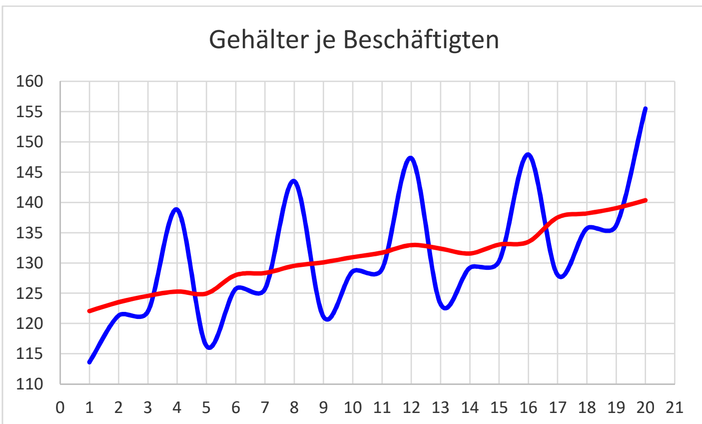
Bestimmung der Saisonkomponente: Beispiel

Saisonbereinigte Zeitreihe (y_t^*) der Löhne und Gehälter je Beschäftigten für den gesamten Beobachtungszeitraum

Jahr (i)	Quartal (j)	y_t	s_j	(y_t^*)
1986 (1)	Q1 (1)	113.6	0.9308	122.05
1986 (1)	Q2 (2)	121.3	0.9820	123.52
1986 (1)	Q3 (3)	122	0.9794	124.56
1986 (1)	**Q4 (4)**	138.8	**1.1079**	125.28
1987 (2)	Q1 (1)	116.3	0.9308	124.95
1987 (2)	Q2 (2)	125.7	0.9820	128.00
1987 (2)	Q3 (3)	125.7	0.9794	128.34
1987 (2)	**Q4 (4)**	143.5	**1.1079**	129.52
1988 (3)	Q1 (1)	121.1	0.9308	130.11
1988 (3)	Q2 (2)	128.6	0.9820	130.95
1988 (3)	Q3 (3)	129	0.9794	131.71
1988 (3)	**Q4 (4)**	147.3	**1.1079**	132.95
1989 (4)	Q1 (1)	123.2	0.9308	132.36
1989 (4)	Q2 (2)	129.2	0.9820	131.56
1989 (4)	Q3 (3)	130.3	0.9794	133.04
1989 (4)	**Q4 (4)**	147.9	**1.1079**	133.49
1990 (5)	Q1 (1)	128	0.9308	137.52
1990 (5)	Q2 (2)	135.7	0.9820	138.18
1990 (5)	Q3 (3)	136.2	0.9794	139.06
1990 (5)	**Q4 (4)**	155.5	**1.1079**	140.35

$= y_{86,1} \,/\, s_1$

$= y_{87,3} \,/\, s_3$



2.2 Regressionsverfahren

Saison-Dummies für den saisonalen Einfluss.

A. Regressionsmodell mit Interzept

Anzahl der Saison-Dummies = p-1

Grund: Spalten von p Saison-Dummies in der Beobachtungsmatrix (**X**) summieren zu 1 → exakte Multikollinearität → Regressionsmodell ist nicht mehr schätzbar

Basis: Quartalsdaten → p = 4

Schätzung: linearer Trend + Saisonkomponente

Regression: $y_t = b_0 + b_1 T + b_2 D_2 + b_3 D_3 + b_4 D_4 + u_t$

D_j = Saison-Dummy für das j-te Quartal:

$$D_j = \begin{cases} 1, \text{ falls Quartal } j \\ 0, \text{ sonst} \end{cases} \quad j=2,3,4$$

Regressionsverfahren

Referenzquartal = 1. Quartal → Koeffizienten $ß_2$, $ß_3$ und $ß_4$ geben die Saisoneinflüsse der Quartale 2, 3 und 4 relativ zum 1. Quartal an.

Regressionsmodell in Matrizenform: **y** = **Xß** + **u**

y: Tx1-Vektor der Zeitreihenwerte (bei kompletten Jahren: T=p·k)

X: Tx5-Beobachtungsmatrix („X") bei Quartalsdaten

ß: 5x1-Vektor der Regressionskoeffizienten, $\mathbf{ß} = (ß_0 \;\; ß_1 \;\; ß_2 \;\; ß_3 \;\; ß_4)^T$

u: Tx1-Vektor der Störvariablen

Struktur der Beobachtungsmatrix **X:**

x_0: Scheinvariable

x_1: Trendvariable

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 1 & 3 & 0 & 1 & 0 \\ 1 & 4 & 0 & 0 & 1 \\ 1 & 5 & 0 & 0 & 0 \\ 1 & 6 & 1 & 0 & 0 \\ 1 & 7 & 0 & 1 & 0 \\ 1 & 8 & 0 & 0 & 1 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ x_0 & x_1 & D_2 & D_3 & D_4 \end{bmatrix}$$

OLS-Schätzer $\mathbf{b} = (\mathbf{X'X})^{-1}\mathbf{X'y}$ enthält den Trendkoeffizienten β_1 sowie die gesuchten Regressionsschätzer für den Saisoneinfluss der Phasen (hier: Quartale) 2, 3, 4 relativ zur 1. Phase.

Regressionsverfahren

Bestimmung der **Saisonkomponente (**normierte Saisonziffern) S_1, S_2, S_3, S_4 für alle vier Quartale:

$$S_j = \begin{cases} -\bar{d} \text{ für } j = 1 \text{(Referenzphase)} \\ b_j - \bar{d} \quad \text{für } j = 2,\, 3,\, 4 \end{cases} \qquad \bar{d} = \sum_{j=2}^{4} b_j \Big/ 4 \qquad \text{(1. Phase: Referenzphase)}$$

Beispiel: Löhne und Gehälter je Beschäftigten
(Quartalsdaten für 5 Jahre: T= k·p = 5·4=20)

Regression: Lohnindex = 112.63 + 0.8675T+ 6.79D_2 + 6.47D_3 + 23.56D_4

t-Werte	(158.0)	(17.6)	(8.6)	(8.13)	(29.36)

Arithm. Mittel der Saisonkoeffizienten (einschl. 0 für Referenzphase):

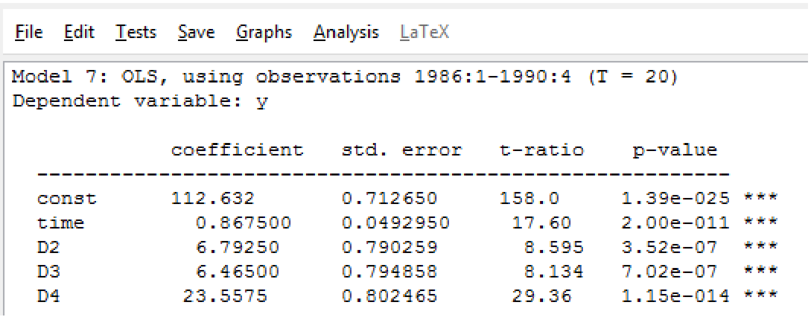
$$\bar{d} = (0 + 6.8 + 6.5 + 23.6)/4 = 36.805/4 = 9.2$$

Normierte Saisonkomponente:

S_1 = 0 – 9.2 = -9.2 $\qquad$ S_2 = 6.79 – 9.2 = -2.41

S_3 = 6.47 – 9.2 = -2.73 $\qquad$ S_4 = 23.56 – 9.2 = 14.35

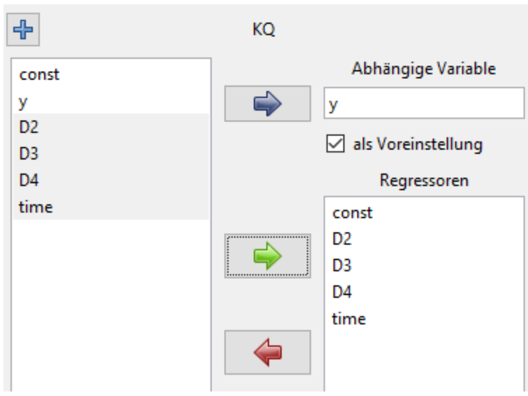
Gretl

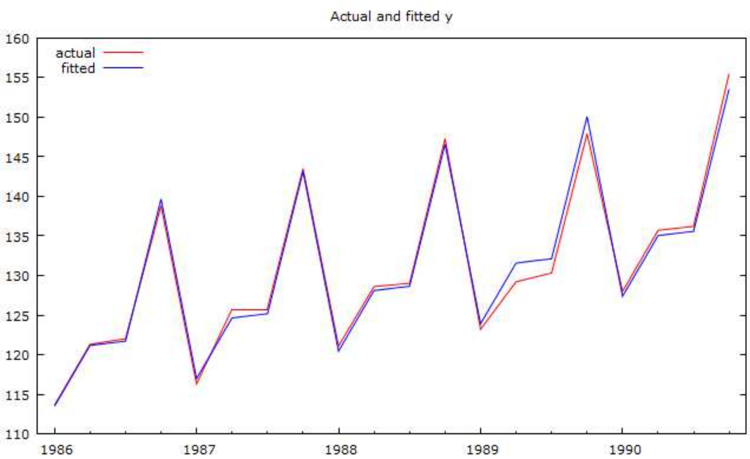


```
File  Edit  Tests  Save  Graphs  Analysis  LaTeX

Model 7: OLS, using observations 1986:1-1990:4 (T = 20)
Dependent variable: y

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const      112.632       0.712650     158.0      1.39e-025 ***
  time         0.867500    0.0492950     17.60     2.00e-011 ***
  D2           6.79250     0.790259       8.595    3.52e-07  ***
  D3           6.46500     0.794858       8.134    7.02e-07  ***
  D4          23.5575      0.802465      29.36     1.15e-014 ***
```





Regressionsmodell ohne Interzept

B. Regressionsmodell ohne Interzept

Mit p Saison-Dummies können die Saisonkoeffizienten aller p Phasen eines Jahres ohne Bezug zu einer Referenzphase geschätzt werden.

Regressionsmodell mit Quartalsdaten (p=4):

$$y_t = \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \beta_4 D_4 + \beta_5 T + u_t$$

Saison-Dummies D_j :

$$D_j = \begin{cases} 1, \text{ falls Quartal } j \\ 0, \text{ sonst} \end{cases} \qquad j=1,2,3,4$$

Die Beobachtungsmatrix („Designmatrix“) **X** hat hier die Struktur:

x_1: Trendvariable T

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \\ 1 & 0 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 & 6 \\ 0 & 0 & 1 & 0 & 7 \\ 0 & 0 & 0 & 1 & 8 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ D_1 & D_2 & D_3 & D_4 & x_1 \end{bmatrix}$$

Regressionsmodell ohne Interzept

Bestimmung der (normierten) Saisonkomponente $S_1, S_2, ..., S_p$ als Durchschnitts-wert der geschätzten Saisonkoeffizienten:

$$\bar{d} = \left(\sum_{j=1}^{p} b_j\right) \Big/ p$$

Normierte Saisonkomponenten: $S_j = b_j - \bar{d}$ j=1,2,3,4

Beispiel: Gehälter je Beschäftigten

(Quartalsdaten für 5 Jahre: T= k·p= 5·4=20)

Lohnindex = $112.6 \cdot D_1 + 119.43 D_2 + 119.1 D_3 + 136.19 D_4 + 0.8675 T$

Arithmetisches Mittel der Saisonkoeffizienten:

$$\bar{d} = (112.6 + 119.43 + 119.1 + 136.19)/4 = 487.3/4 = 121.83$$

Normierte Saisonkomponenten: $\sum_{j=1}^{4} s_j = 0$

S_1 = 112.6 – 121.83 = -9.23 S_2= 119.43 – 121.83 = -2.41

S_3 = 119.1 – 121.83 = -2.74 S_4 = 136.19 – 121.83 = 14.37

Modelle für die Saisonkomponente

Modell für Daten mit Saisonschwankungen; $y_t = T_t + S_t + e_t$

S_t: periodische Funktion

$$\left.\begin{array}{l}\text{Für Monatsdaten: } S_{t+12} = S_t \\ \text{Für Quartalsdaten: } S_{t+4} = S_t\end{array}\right\} \text{ allgemein: } S_{t+a} = S_t$$

Als Modell für die Saisonschwankungen werden oft trigonometrische Funktionen verwendet:

$S_t = \beta_1 \cos(\omega_1 t) + \gamma_1 \sin(\omega_1 t) \qquad t = 1,2,\ldots$ wobei $\omega_1 = 2\pi/a$

Gewünschte Eigenschaft: $S_{t+a} = S_t$ für alle t

Beispiele für Monatsdaten

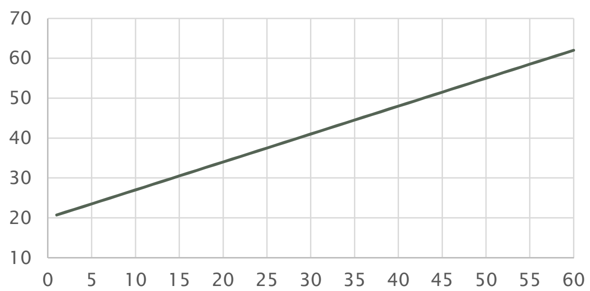
Das Modell für S_t hängt von zwei Parametern ab:

β_1 bestimmt die Amplitude des Saisonschwankungen (Höhe der Wellen)

γ_1: bestimmt die Phase (die Lage des Höhenpunktes der Welle)

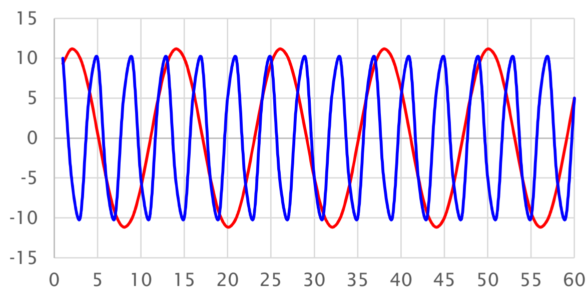
Verschiedene Modelle

Trend: $T_t = 20 + 0.07t$



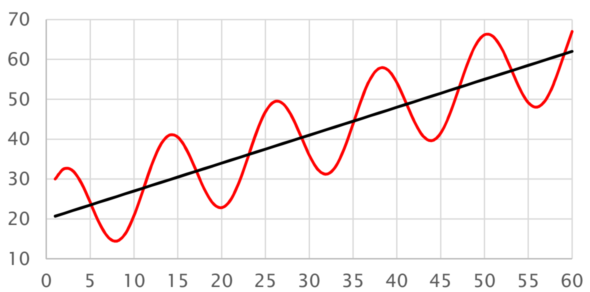
$T_t = 20 + 0.7t$

Saisonkomponente 1 und 2

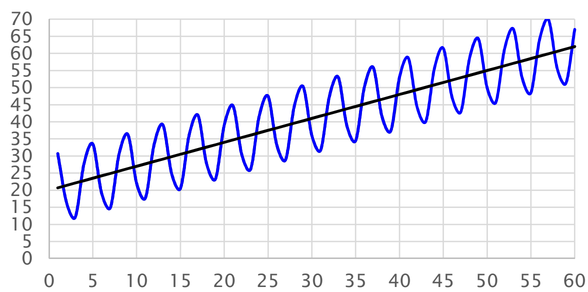


$S_t = 5\cos(\omega_1 t) + 10\sin(\omega_1 t)$ a = 12 und a = 4

Trend + Saisonkomponente 1



Trend + Saisonkomponente 2



Modelle für die Saisonkomponente

Beispiel für eine Zeitreihe von Monatsdaten, die linear wächst und wellförmige Schwankungen aufweist:

$T_t = 20 + 0.7t$ und $S_t = 5\cos(\omega_1 t) + 10\sin(\omega_1 t)$ wobei a = 12 und a = 4

Modell mit 4 Saisonparametern:

$$S_t = \beta_1\cos(\omega_1 t) + \gamma_1\sin(\omega_1 t) + \beta_2\cos(\omega_2 t) + \gamma_2\sin(\omega_2 t)$$

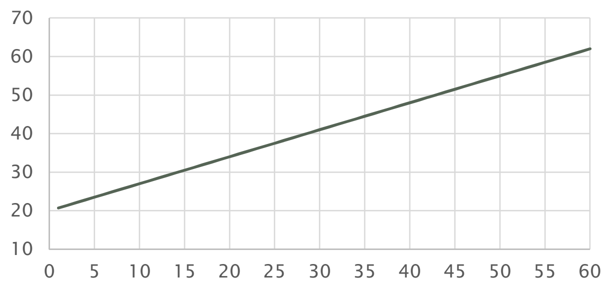
$$S_t = \underbrace{5\cos(\omega_1 t) + 10\sin(\omega_1 t)}_{\text{1. Komponente}} + \underbrace{4\cos(\omega_2 t) + 6\sin(\omega_2 t)}_{\text{2. Komponente}} \quad \text{wobei } \omega_2 = 4\pi/a$$

Trigonometrische Funktion sind in der Lage, komplizierte Saison-schwankungen zu modellieren.

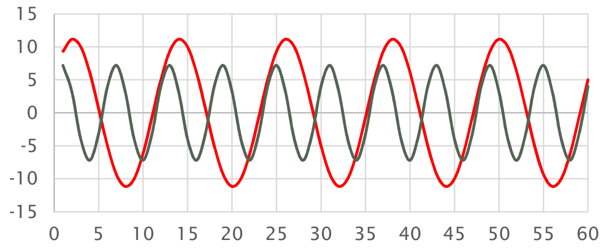
Im Allgemeinen: $S_t = \sum_{j=1}^{q}\left(\beta_j\cos(\omega_j t) + \gamma_j\sin(\omega_j t)\right)$ wobei $\omega_j = 2\pi j/a$ und $q \le a/2$

Verschiedene Modelle

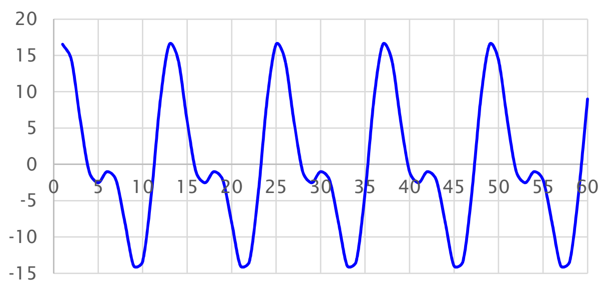
Trend: $T_t = 20 + 0.07t$



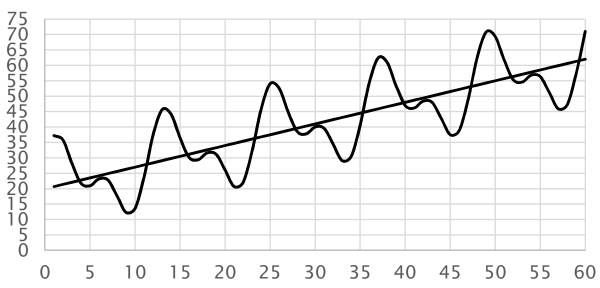
Saisonkomponente 1.Teil und 2.Teil



Saisonkomponente 1+2



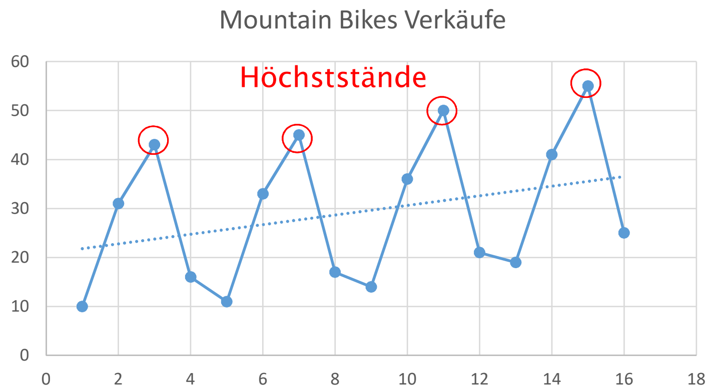
Trend + Saisonkomponenten



2.3 Holt-Winter Verfahren

Die Methode von Holt-Winters ist zur Prognose von Zeitreihen mit **linearem Trend** und **Saison** geeignet.
Es treten hier 3 Glättungsparameter auf:
α, γ und δ mit $0 < \alpha, \gamma, \delta < 1$

Jahr	Quartal	Yt
1	1	10
1	2	31
1	3	43
1	4	16
2	1	11
2	2	33
2	3	45
2	4	17
3	1	14
3	2	36
3	3	50
3	4	21
4	1	19
4	2	41
4	3	55
4	4	25



Positiver Trend über 4 Jahre
Saisonkomponente bleibt konstant
→ additives Holt-Winters Modell

Holt-Winters Modell mit additiver Saison

Im additiven Modell ist die Saison vom **Niveau** der Zeitreihe unabhängig.
Zeitreihenmodell: $y_t = (\beta_0 + \beta_1 t) + S_t + u_t$

Niveau: $L_t = \alpha(Y_t - S_{t-p}) + (1-\alpha)(L_{t-1} + b_{t-1})$

Trend: $b_t = \gamma(L_t - L_{t-1}) + (1-\gamma)b_{t-1}$ (Wachstumsrate)

Saison: $S_t = \delta(Y_t - L_t) + (1-\delta)S_{t-p}$

p: Anzahl von Saisons (Phasen) innerhalb eines Jahres (p = 12 für Monatsdaten, p = 4 für Quartalsdaten).

h-Schritte-Prognose zum Zeitpunkt t: $\hat{y}_{t+h}(t) = L_t + h b_t + S_{t+h-p}$
Neu ist die **Saisongleichung** → exponentielle Glättung der $(Y_t - L_t)$ Komponente (Beobachtung minus geglättetem Niveau)
Niveau L ergibt sich aus einer exponentiellen Glättung der saisonbereinigten Werte:
$L_t = \alpha(Y_t - S_{t-p}) + (1-\alpha)\,\hat{y}_t(t-1)$

Holt-Winters Modell mit additiver Saison: Beispiel

Anfangswerte: L_0, b_0, und S_{-3}, S_{-2}. S_{-1}.

OLS Schätzung für die Hälfte der Zeitreihen (16 Daten)

Wenigstens 4 Jahre!

Schritt 1: Anfangswerte mittels Regression

Regressionsgleichung: $\hat{y}_t$ = 20.85 + 0.9809t

Interzept: L_0 = 20.85 und

Steigung: b_0 = 0.9809

```
              Koeffizient   Std.-fehler
  --------------------------------------
  const        20,8500       7,49172
  time          0,980882     0,774775

Mittel d. abh. Var.     29,18750    Stdal
Summe d. quad. Res.     2857,313    Stdf
R-Quadrat               0,102726    Korr
F(1, 14)                1,602813    P-We
Log-Likelihood         -64,18340    Akai
Schwarz-Kriterium       133,9120    Hann
rho                    -0,075002    Durb
```


Holt-Winters Modell mit additiver Saison: Beispiel

Schritt 2: Bestimmung der Saisonkomponente

Schätzungen der Trendkomponente (N = 16 = 4x4Q):

$\hat{y}_1$ = 20.85 + 0.9809(1) = 21.8309

$\hat{y}_2$ = 20.85 + 0.9809(2) = 22.8118

...

$\hat{y}_{16}$ = 20.85 + 0.9809(16) = 36.5444

Trendbereinigte Zeitreihe: $S_t = y_t - \hat{y}_t$

$S_1 = y_1 - \hat{y}_1$ = 10 – 21.8309 = -11.8309

$S_2 = y_2 - \hat{y}_2$ = 31 – 22.8118 = 8.1882

....

$S_{16} = y_{16} - \hat{y}_{16}$ = 25 – 36.5444 = -11.5444

Berechnung der durchschnittlichen Saisonkomponente

$\bar{S}_1 = \frac{1}{4}(S_1 + S_5 + S_9 + S_{13})$ = (-11.8309 -14.7545 -15.6781 -14.601)/4 = -14.2163

$\bar{S}_4 = \frac{1}{4}(S_4 + S_8 + S_{12} + S_{16})$ = (-8.773 -11.69 -11.62 -11.54)/4 = -10.908

Holt-Winters Modell: Startwerte

Parameter
$\alpha = 0.2$
$\gamma = 0.1$
$\delta = 0.1$
$p = 4$

Zeit	Y_t	Niveau L_t	Wachstum b_t	Saison-Komponente
-3	Q1			-14.2162
-2	Q2			6.5529
-1	Q3	Regression		18.5721
0	Q4	**20.85**	**0.9809**	-10.9088
1	10	22.3079	1.0286	-14.0254
2	31	23.5586	1.0508	6.6418
3	43	24.5731	1.0472	18.5575

$$S_t = \delta(Y_t - L_t) + (1-\delta)S_{t-p}$$

$$S_1 = 0.1(Y_1 - L_1) + 0.9S_{-3} =$$
$$= 0.1(10 - 22.3079) + 0.9(-14.21)$$
$$= -14.025$$

$$L_t = \alpha\underbrace{(Y_t - S_{t-p})}_{\text{Saisonbereinigte Beob.}} + (1-\alpha)(L_{t-1} + b_{t-1})$$

Saisonbereinigte Beob.

L_0 = Interzept der Regression

$$L_1 = 0.2(Y_1 - S_{-3}) + 0.8(L_0 + b_0) =$$
$$= 0.2(10 + 14.21) + 0.8(20.85 + 0.98)$$
$$= 22.307$$

$$b_t = \gamma\underbrace{(L_t - L_{t-1})}_{\text{Niveauänderung}} + (1-\gamma)b_{t-1}$$

Niveauänderung

b_0 = Steigung aus Regression

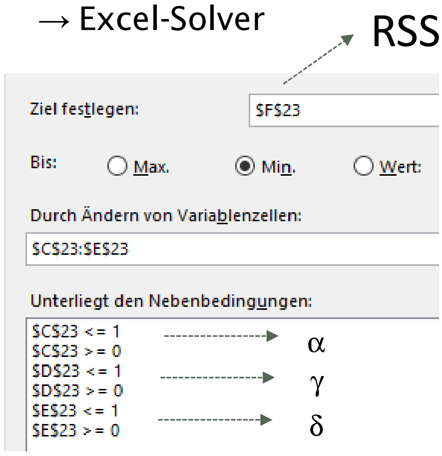
$$b_1 = 0.1(L_1 - L_0) + 0.9b_0 =$$
$$= 0.1(22.3079 - 20.85) + 0.9(0.98)$$
$$= 1.028$$

Holt-Winters Verfahren: Optimale Glättungsparameter

α	γ	δ	SSE	MSE	σ
0.56	0	0	18.7974	1.4460	1.2025

Zeit	Y_t	Niveau	Wachstum	Saison Komponente	Prognose $\hat{y}_t(t-1)$
-3				-14.2162	
-2				6.5529	
-1				18.5721	
0		20.85	0.9809	-10.9088	
1	10	23.1682	0.9809	-14.2162	7.6147
2	31	24.3161	0.9809	6.5529	30.7020
3	43	24.8098	0.9809	18.5721	43.8691
4	16	26.4175	0.9809	-10.9088	14.8818
5	11	26.1750	0.9809	-14.2162	13.1823
6	33	26.7585	0.9809	6.5529	33.7088
7	45	27.0041	0.9809	18.5721	46.3114
8	17	27.9423	0.9809	-10.9088	17.0762
9	14	28.5268	0.9809	-14.2162	14.7070
10	36	29.4737	0.9809	6.5529	36.0606
11	50	31.0003	0.9809	18.5721	49.0266
12	21	31.9406	0.9809	-10.9088	21.0723
13	19	33.0867	0.9809	-14.2162	18.7053
14	41	34.2803	0.9809	6.5529	40.6205
15	55	35.9153	0.9809	18.5721	53.8333
16	25	36.3426	0.9809	-10.9088	25.9874

Schritt 3: Finde die optimalen Werte für α, γ und δ
→ Min RSS!
→ Excel-Solver



$$\alpha = 0.506$$
$$\gamma = 0$$
$$\delta = 0$$

Holt-Winters Modell: h-Schritte-Prognose

Schritt 3: Berechnung der Prognose mit $(\alpha, \gamma, \delta) = (0.2, 0.1, 0.1)$

$\hat{y}_{t+h}(t) = L_t + h b_t + S_{t+h-p}$

$t = 0$, $h = 1$:

$\hat{y}_{0,1} = \hat{y}_0 = L_0 + b_0 + S_{1-4} = 20.85 + 0.9809 - 14.2162 = 7.6147$

h-Schritte-Prognose

In Periode 16 (letzte Beobachtung), die Einschritt-Prognose für Periode 17:

$\hat{y}_{17}(16) = L_{16} + b_{16} + S_{17-4} = 36.1813 + 0.9544 - 14.216 = 22.86$

In Periode 16, die Zwei-Schritt-Prognose für Periode 18:

$\hat{y}_{18}(16) = L_{16} + 2b_{16} + S_{18-4} = 36.1813 + 2(0.9544) + 6.524 = 44.61$

Holt-Winters Verfahren: h-Schritte-Prognose

In Periode 16, die Drei-Schritte-Prognose

$\hat{y}_{16,3} = \hat{y}_{19}(16) = L_{16} + 3b_{16} + S_{19-4} = 36.1813 + 3(0.9544) + 18.5721 = 57.62$

Wenn $y_{17} = 33$ beobachtet wird,

- Glättungsparameter α, γ, δ aktualisieren, indem RSS minimiert wird oder
- nächste Trend- L_{17} und Niveauschätzung b_{17} berechnen

$L_{17} = \alpha(y_{17} - S_{t-p}) + (1-\alpha)(L_{16} + b_{16})$
$= 0.2(33+14.269) + 0.8(36.1813 + 0.9544) = 39.1624$

$b_{17} = \gamma(L_{17} - L_{16}) + (1-\gamma)b_{16}$
$= 0.1(39.1624 - 36.1813) + 0.9(0.9544) = 1.157$

Holt-Winters Modell mit multiplikativer Saison

Im multiplikativen Modell ist der Saisoneffekt **proportional** zum Niveau der Reihe. Der Trend selbst ist linear.

Zeitreihenmodell: $y_t = (\beta_0 + \beta_1 t)S_t \times u_t$

Niveau: $L_t = \alpha(Y_t/S_{t-p}) + (1 - \alpha)(L_{t-1} + b_{t-1})$

Trend: $b_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)b_{t-1}$

Saison: $S_t = \delta(Y_t/L_t) + (1 - \delta)S_{t-p}$

p: Anzahl von Saisons innerhalb eines Jahres (p = 12 für Monatsdaten, p = 4 für Quartalsdaten).

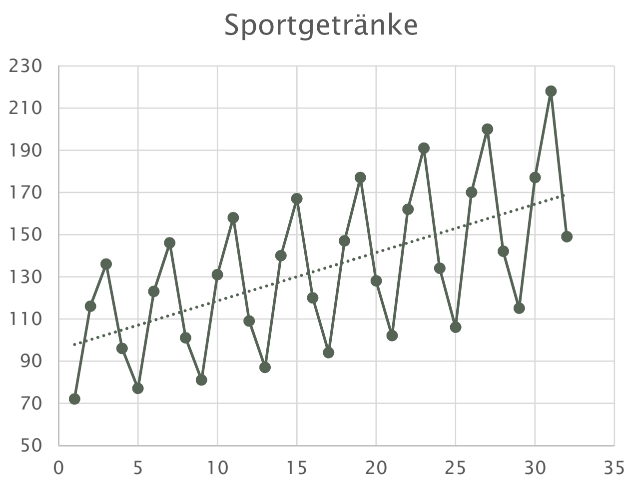
h-Schritte Prognose zum Zeitpunkt t: $\hat{y}_{t+h}(t) = (L_T + hb_t)S_{t+h-p}$

Die Saisongleichung beschreibt eine exponentielle Glättung des Quotienten (Y_t/L_t).

Beispiel: Sportgetränke

Beobachtungen:
- ✓ Linearer Aufwärtstrend
- ✓ Saisonale Schwankungen nehmen über die Zeit zu

Quartal	1	2	3	4	5	6	7	8
1	72	77	81	87	94	102	106	115
2	116	123	131	140	147	162	170	177
3	136	146	158	167	177	191	200	218
4	96	101	109	120	128	134	142	149



Anfangswerte

Anfangswerte: L_0, b_0, und die Saisonkomponente S_{-3}, S_{-2}, S_{-1}, S_0.

OLS Schätzung für die Hälfte der Zeitreihen (16 Daten)
Wenigstens 4 Jahre!

Schritt 1: Anfangswerte mittels Regression
Regressionsgleichung:

$\hat{y}_t = 95.25 + 2.4706t$

Interzept: $L_0 = 95.25$ und
Steigung: $b_0 = 2.4706$

```
                    Koeffizient   Std.-fehler
  ----------------------------------------
  const          95,2500        14,4648
  time            2,47059        1,49591

Mittel d. abh. Var.      116,2500   Stdabw
Summe d. quad. Res.      10651,71   Stdfeh
R-Quadrat                0,163062   Korrig
F(1, 14)                 2,727649   P-Wert
Log-Likelihood          -74,71011   Akaike
Schwarz-Kriterium        154,9654   Hannan
rho                     -0,031579   Durbin
```


Holt-Winters Modell mit multiplikativer Saison: Beispiel

Schritt 2: Bestimmung der Saisonkomponente
Schätzungen der Trendkomponente:

$\hat{y}_1 = 95.25 + 2.4706(1) = 97.7206$

$\hat{y}_2 = 95.25 + 2.4706\,(2) = 100.1912$

...

$\hat{y}_{16} = 95.25 + 2.4706\,(16) = 134.7794$

Trendbereinigte Zeitreihe: $S_t = y_t / \hat{y}_t$

$S_1 = y_1 / \hat{y}_1 = 72/97.7206 = 0.7368$

$S_2 = y_2 / \hat{y}_2 = 116/100.1912 = 1.1578$

....

$S_{16} = y_{16} / \hat{y}_{16} = 120/134.7794 = 0.8903$

Berechnung der durchschnittlichen Saisonkomponenten:

$\bar{S}_1 = (S_1 + S_5 + S_9 + \ldots + S_{29})/8 = (0.7368+0.7156+0.6894+\ldots+0.689)/8 = 0.695$

$\bar{S}_2 = (S_2 + S_6 + S_{10} + \ldots + S_{30})/8 = (1.1578+1.1174+1.0921+\ldots+1.045)/8 = 1.086$

$\bar{S}_3 = (S_3 + S_7 + S_{11} + \ldots + S_{31})/8 = (1.13247+1.2906+1.2622+\ldots+1.2686)/8 = 1.272$

...

Holt-Winters Modell mit multiplikativer Saison: Beispiel

$$\sum_{i=1}^{4} \bar{S}_i = 0.695 + 1.086 + 1.272 + 0.878 = 3.9327$$

Normierungsfaktor: $NF = p \Big/ \sum_{i=1}^{4} \bar{S}_i = 4 / 3.9327 = 1.01708$

Normierte Saisonkomponente: $S_i = \bar{S}_i \cdot NF$

Durchschnitt der Saisonkomponenten = 1.017

Saisonkomponente für Initialisierung:

$S_{-3} = S_{1-4} = \bar{S}_1 \cdot NF = S_1 = 0.695(1.017) = 0.7078$

$S_{-2} = S_{2-4} = \bar{S}_2 \cdot NF = S_2 = 1.086(1.017) = 1.1050$

$S_{-1} = S_{3-4} = \bar{S}_3 \cdot NF = S_3 = 1.2937(1.017) = 1.2941$

$S_0 = S_{4-4} = \bar{S}_4 \cdot NF = S_4 = 0.878(1.017) = 0.8930$

Holt-Winters Modell mit multiplikativer Saison: Beispiel

Schritt 3: Berechnung der 1-Schritt-Prognose

1-Schritt-Prognose: $\hat{y}_{t+h}(t) = (L_t + hb_t)\, S_{t+h-p}$ mit $h = 1$

$t = 0,\ p = 4$: $\hat{y}_1(0) = (L_0 + b_0)S_{1-4} = (95.25+2.4706)0.7078 = 69.165$

Werte: $t = 1$, $h = 1$, $p = 4$, $\alpha = 0.2$, $\gamma = 0.1$ und $\delta = 0.1$

$L_1 = \alpha(y_1/S_{-3}) + (1-\alpha)(L_0 + b_0)$

$= 0.2(72/0.7078) + 0.8(95.25 + 2.4706) = 98.521$

$b_1 = \gamma(L_1 - L_0) + (1-\gamma)b_0 = 0.1(98.521-95.25) + 0.9(2.4706) = 2.5507$

$S_1 = \delta(y_1/L_1) + (1-\delta)S_{1-4} = 0.1(72/98.521) + 0.9(0.7078) = 0.701$

Prognose: $\hat{y}_2(1) = (L_1 + b_1)S_{2-4} = (98.521 + 2.5507)(1.105) = 111.685$

$L_2 = \alpha(y_2/S_{-2}) + (1-\alpha)(L_1 + b_1) = 0.2(116/1.105) + 0.8(98.521+2.5507) = 101.853$

$b_2 = \gamma(L_2 - L_1) + (1-\gamma)b_1 = 0.1(101.853-98.521) + 0.9(2.5507) = 2.6288$

$S_2 = \delta(y_2/L_2) + (1-\delta)S_{2-4} = 0.1(116/101.853) + 0.9(1.105) = 1.1084$

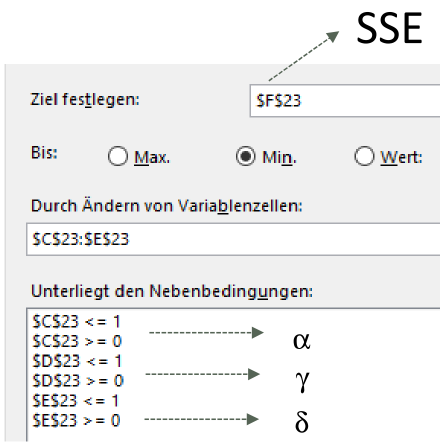
Prognose: $\hat{y}_3(2) = (L_2 + b_2)S_{3-4} = (101.853 + 2.6288)(1.2937) = 135.213$

Holt-Winters Modell: Optimale Glättungsparameter

α	γ	δ	SSE	MSE
0.2	0.1	0.10000	177.3537	5.9118

Zeit	y_t	Niveau	Trend	Saisonfaktor	$\hat{y}_{t,1}$ Forecast	Prognosefehler e_t	Quadrierte e_t^2
-3				0.7062			
-2				1.1114			
-1				1.2937			
0		**95.25**	**2.4706**	0.8886	69.0103		
1	72	98.5673	2.5553	0.7086	112.3876	2.9897	8.9384
2	116	101.7726	2.6203	1.1142	135.0531	3.6124	13.0492
3	136	104.5392	2.6349	1.2944	95.2335	0.9469	0.8966
4	96	107.3467	2.6522	0.8892	77.9481	0.7665	0.5876
5	77	109.7313	2.6254	0.7079	125.1923	-0.9481	0.8990
6	123	111.9632	2.5861	1.1127	148.2754	-2.1923	4.8062
7	146	114.1977	2.5509	1.2928	103.8078	-2.2754	5.1774
8	101	116.1170	2.4877	0.8872	83.9645	-2.8078	7.8838
9	81	117.7672	2.4040	0.7059	133.7113	-2.9645	8.7883
10	131	119.6839	2.3553	1.1109	157.7760	-2.7113	7.3513
11	158	122.0738	2.3587	1.2930	110.3993	0.2240	0.0502
12	109	124.1171	2.3272	0.8863	89.2597	-1.3993	1.9581
13	87	125.8040	2.2632	0.7045	142.2648	-2.2597	5.1064
14	140	127.6594	2.2224	1.1094	167.9343	-2.2648	5.1295
15	167	129.7373	2.2079	1.2924	116.9458	-0.9343	0.8729
16	120	132.6344	2.2768	0.8882	95.0429	3.0542	9.3284

Finde die optimalen Werte für α, γ und δ → Min SSE!



Ergebnis
$\alpha = 0.394$
$\gamma = 0.025$
$\delta = 0$

Holt-Winters Modell: h-Schritte-Prognose

h-Schritte-Prognose: $\hat{y}_{t+h}(t) = (L_t + hb_t)S_{t+h-p}$

In Periode 32 (letzte Beobachtung), die 1-Schritt-Prognose für Periode 33

$\hat{y}_{33}(32) = (L_{32} + b_{32})S_{33-4} = (167.779 + 2.238)(0.7057) = 119.99$

In Periode 32, die Drei-Schritte-Prognose für Periode 35

$\hat{y}_{35}(32) = (L_{32} + 3b_{32})S_{35-4} = (167.779 + 3(2.238))1.293 = 225.615$

Wenn y_{33} = 123 beobachtet wird, Schätzer α γ und d aktualisieren oder L33 und b33 rekursiv schätzen:

$L_{33} = \alpha(y_{33}/S_{33-4}) + (1-\alpha)(L_{32} + b_{32})$
$= 0.2(123/0.7057) + 0.8(167.77 + 2.238) = 170.87$

$b_{33} = \gamma(L_{33} - L_{32}) + (1-\gamma)b_{32}$
$= 0.1(170.87 - 167.77) + 0.9(2.2387) = 2.324$

Holt und Holt-Winters Glättungsverfahren

- Diese Glättungsethoden sind einfach und benötigen höchstens drei Parameter.
- Sie lassen sich leicht automatisieren, sodass sie auf eine grosse Zahl von Zeitreihen (zB in grossen Lagerhaltungsproblemen) anwendbar werden.
- Die Zielfunktion (minimale Fehlerquadratsumme) kann bei Vorliegen von Ausreissern zB durch die Minimierung des MAE_p, mean absolute percentage error (MAPE), ersetzt werden.

$$MAPE = \frac{100}{T}\sum_{t=1}^{T}\left|\frac{y_t - y_t(t-1)}{y_t}\right|$$

- Holt-Winters für exponentielles Wachstum und multiplikative Saisons in die Zeitreihe zuerst logarithmieren und dann Holt-Winters mit additiver Saison anwenden.

2.4 Hodrick-Prescott-Filter

HP-Filter: Verfahren der Makroökonomie um Konjunkturzyklen zu analysieren.
Ziel: Zeitreihe ausgleichen, so dass diese weniger abhängig von kurzfristigen Schwankungen ist.
HP-Filter separiert den Trend von der zyklischen Komponente.
Zeitreihenzerlegung: $y_t = \tau_t + c_t$ τ: Trend c: zyklische Komponente

Trendkomponente τ minimiert den folgenden Ausdruck:

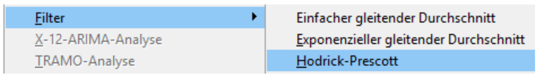
$$\sum_{t=1}^{T}(y_t-\tau_t)^2+\lambda\sum_{t=2}^{T-1}\left[(\tau_{t+1}-\tau_t)-(\tau_t-\tau_{t-1})\right]^2.$$

Erster Term: Summe der quadrierten Abweichungen (y_t-τ_t) zwischen Trendkomponente und Wert der Zeitreihe y_t :

Zweiter Term: Vielfaches λ der Summe der quadrierten zweiten Differenzen der Trend-Komponente. Dieser Term *bestraft* die Variation in der Trend-Komponente → Glattheit der Trendfunktion.
Je höher λ, desto grösser die Bestrafung und glatter die Trendfunktion.

Hodrick-Prescott-Filter

Quartalszahlen: $\lambda = 1600$
Monatszahlen: $\lambda = 14‘400$



Der Hodrick-Prescott-Filter kann den Trend flexibler ausgleichen als ein einfacher linearer Filter.

Kritikpunkten:

- HP-Filter weist an den Rändern eine Verzerrung auf, da für die Randwerte Schätzwerte eingesetzt werden müssen, um Differenzen bilden zu können.
- Der Filter kann per Konstruktion längere Trendabweichungen nicht ausweisen → durch die Wahl von λ auf eine bestimmte maximale Länge des Konjunkturzyklus festgelegt. Eine Rezession von mehr als drei Jahren würde bereits als Abwärtstrend gewertet.
- λ bleibt ein frei gewählter Parameter, der keine theoretische Fundierung aufweist.