2.2.3 Exponentielle Glättung

Mit Hilfe der **exponentiellen Glättung** kann wie beim Verfahren der gleitenden Mittelwerte die glatte Komponente einer Zeitreihe herausgefiltert werden. Zum Zwecke der Glättung werden die Zeitreihenwerte jedoch nicht mehr gleich-, sondern exponentiell gewichtet. Das exponentielle Gewichtsschema weist den weiter zurückliegenden Werten **geometrisch abnehmende Gewichte** zu.

Im Unterschied zur Methode der gleitenden Durchschnitte ist mit dem Verfahren der exponentiellen Glättung jedoch unmittelbar eine Prognosegleichung verbunden, die zur kurzfristigen Vorhersage von Zeitreihen verwendet werden kann.

Bei **trendbehafteten Zeitreihen** ist für die Glättung und Prognose das **Verfahren der exponentiellen Glättung zweiter Ordnung** konzipiert worden. Das Prinzip der exponentiellen Glättung lässt sich jedoch am besten anhand eines Grundmodells ohne Trend (und ohne Saison) veranschaulichen. Man spricht in diesem Fall von einer **exponentiellen Glättung erster Ordnung**.

• Exponentielle Glättung erster Ordnung

<u>Annahme:</u> Zeitreihe (y_t) schwankt um einen konstanten Wert

Vorhersagewert für die Periode n+1 bei Ausnutzung aller verfügbaren Informationen:

(2.26) $$\hat{y}_{n,1} = \bar{y}_n = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Prognosewert für Periode n+2: (2.27) $$\hat{y}_{n+1,1} = \bar{y}_{n+1} = \frac{1}{n+1}\sum_{i=1}^{n+1} y_i$$

Prognosewert $\hat{y}_{t,h}$: t aktuelle Periode, h Prognosehorizont

Neuer Prognosewert in Abhängigkeit des vorhergehenden Prognosewerts:

(2.28) $$\hat{y}_{n+1,1} = \frac{n}{n+1}\bar{y}_n + \frac{1}{n+1}y_{n+1}$$
$$= \frac{n}{n+1}\hat{y}_{n,1} + \frac{1}{n+1}y_{n+1} \quad .$$

Prognosewert der exponentiellen Glättung für Periode t+1:

(2.29) $$\hat{y}_{t,1} = (1-\alpha)\cdot\hat{y}_{t-1,1} + \alpha\cdot y_t, 0 < \alpha < 1$$ (**Rekursionsformel**)

Alter Prognosewert $\hat{y}_{t-1,1}$: Gewicht 1-α, aktueller Beob.wert y_t: Gewicht α

Beispiel: Die Umsätze an Drehkippbeschlägen der Schloss- und Beschlagindustrie schwankten in einem 8-Jahres-Zeitraum bei keinem klar erkennbaren Trend in etwa um einen Wert von 12,8 GE (s. Abb.). Zum Zwecke einer Vorhersage der Entwicklung dieser Variablen bietet sich daher das Verfahren der exponentiellen Glättung erster Ordnung an.

Bei der exponentiellen Glättung muss ein Anfangswert y_0 festgelegt werden, der als Prognosewert für die erste Periode des Beobachtungszeitraums verwendet werden kann. In der Regel reicht es aus, hierzu den Zeitreihenwert für die Periode unmittelbar vor Beginn des Stützbereichs zu wählen, was hier gemacht werden soll:

$$\hat{y}_{0,1} = y_0 = 12752$$

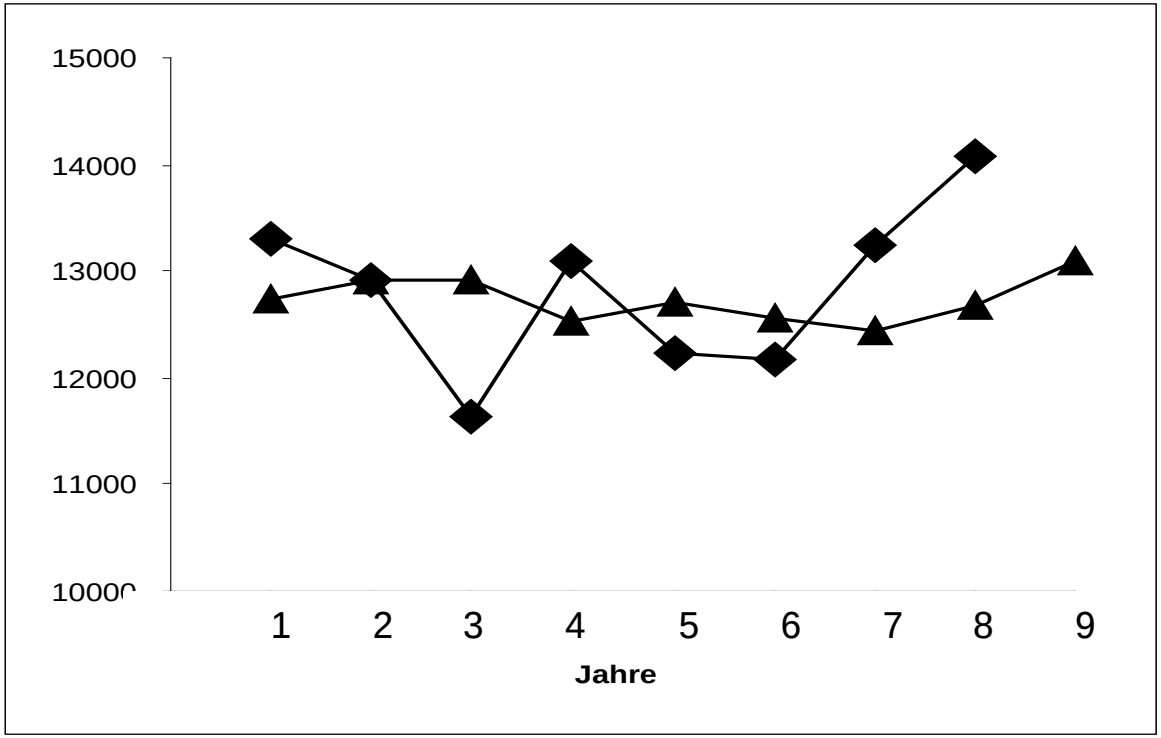
Der Anfangswert entspricht hier dem Umsatz an Drehkippbeschlägen im Jahr 0, der hier als Prognosewert für das Jahr 1 verwendet wird. Unter Verwendung eines Gewichtsfaktors von 0,3 erhält man dann mit der Rekursionsformel (2.29) den Prognosewert

$$\hat{y}_{1,1} = 0{,}7 \cdot \hat{y}_{0,1} + 0{,}3 \cdot y_1 = 0{,}7 \cdot 12752 + 0{,}3 \cdot 13317 = 12922$$

für das Jahr 2. Die Ein-Schritt-Prognosen für den Umsatz an Drehkippbeschlägen können auf diese Weise sukzessive für die Folgejahre bestimmt werden:

Jahr	1	2	3	4	5	6	7	8	9
y_t	13317	12930	11643	13098	12223	12161	13230	14065	
$\hat{y}_{t-1,1}$	12752	12922	12924	12540	12707	12562	12442	12678	13094

Abb.: Drehkippbeschläge und exponentielle Glättung



Prognose im Stützbereich: Ex-post-Prognose
Prognose außerhalb des Stützbereichs: Ex-ante-Prognose

Gewichtungsschema der exponentiellen Glättung

Aktueller Prognosewert: (2.29) $\hat{y}_{t,1} = (1-\alpha)\cdot \hat{y}_{t-1,1} + \alpha \cdot y_t, 0 < \alpha < 1$

Vorheriger Prognosewert: (2.30) $\hat{y}_{t-1,1} = (1-\alpha)\cdot \hat{y}_{t-2,1} + \alpha \cdot y_{t-1}$

Nach Einsetzen von (2.30) in (2.29) erhält man

$$\hat{y}_{t,1} = (1-\alpha)^2 \cdot \hat{y}_{t-2,1} + \alpha \cdot (1-\alpha) \cdot y_{t-1} + \alpha \cdot y_t$$

und nach fortlaufender Substitution der alten Prognosewerte durch

(2.31) $$\hat{y}_{t-i,1} = (1-\alpha)\cdot \hat{y}_{t-i-1,1} + \alpha \cdot y_{t-i}, i = 2,3,\ldots$$

geht die Rekursionsformel (2.29) in die Form

(2.32) $$\hat{y}_{t,1} = \alpha \cdot y_t + (1-\alpha)\cdot \alpha \cdot y_{t-1} + (1-\alpha)^2 \cdot \alpha \cdot y_{t-2} + (1-\alpha)^3 \cdot \alpha \cdot y_{t-3} + \ldots$$

$$= \sum_{i=0}^{\infty} \alpha \cdot (1-\alpha)^i \cdot y_{t-i}$$

über, wenn man den Regress unendlich oft durchführt.

Prognosewert $\hat{y}_{t,1}$ nach (2.32):
Gewogenes arithmetisches Mittel aller zurückliegender Zeitreihenwerte mit geometrisch abnehmenden Gewichten $\alpha(1-\alpha)^i$
(allmähliche Niveauverschiebung wird hierdurch berücksichtigt)

Gewogene Mittelung bei Aufgabe der Vorstellung eines unendlichen Regresses (Beobachtungszeitraum der Länge n):

(2.33) $$\hat{y}_{t,1} = \sum_{i=0}^{n-1} \alpha(1-\alpha)^i \cdot y_{t-i} + (1-\alpha)^n \cdot y_0$$

Der Anfangswert y_0, der mit wachsendem n vernachlässigbar ist da der Faktor $(1-\alpha)^n$ gegen null geht.

Festlegung von y_0 (für Initialisierung des Verfahrens):
Zeitreihenwert oder Mittelwert der Zeitreihenwerte vor Beginn des Stützzeitraums

Fehlerkorrekturformel:

(2.34) $$\hat{y}_{t,1} = \hat{y}_{t-1,1} + \alpha \cdot e_t$$

mit

(2.35) $$e_t = y_t - \hat{y}_{t-1,1}$$

Prognose korrigiert sich quasi selbständig: bei Unterschätzung erfolgt automatisch ein Aufschlag, bei Überschätzung ein Abschlag. Der Prognosefehler e_t wird mit dem Gewicht α berücksichtigt.

Beispiel: Die Anwendung der exponentiellen Glättung soll hier unter Verwendung der Fehlerkorrekturformel (2.34) für die Umsatzdaten der Drehkippbeschläge aufgezeigt werden.

Startwert: $\hat{y}_{0,1} = y_0 = 12752$

Prognosefehler im Jahr 1:

$$e_1 = y_1 - \hat{y}_{0,1} = 13317 - 12752 = 565$$

Ein-Schritt-Prognose für das Jahr 2 mit α=0,3:

$$\hat{y}_{1,1} = \hat{y}_{0,1} + 0{,}3 \cdot e_1 = 12752 + 0{,}3 \cdot 565 = 12922$$

Analog ergeben sich die Prognosewerte für die Folgejahre unter Verwendung der Fehlerkorrekturformel:

Jahr	1	2	3	4	5	6	7	8	9
y_t	13317	12930	11643	13098	12223	12161	13230	14065	
$\hat{y}_{t-1,1}$	12752	12922	12924	12540	12707	12562	12442	12678	13094
e_t	565	8	-1281	558	-484	-401	788	1387	
$0{,}3 \cdot e_t$	170	2	-384	167	-145	-120	236	416	

Bedeutung des Gewichtsfaktors α für die Glättung, Reagibilität und den Einfluss der Zeitreihenwerte

	α klein	α groß
Glättungseffekt der Vorhersage	groß	klein
Reagibilität auf irreguläre Schwankungen	klein	groß
Berücksichtigung neuer Zeitreihenwerte	schwach	stark
Berücksichtigung älterer Zeitreihenwerte	stark	schwach

Wahl des Glättungsparameters α

Optimaler Wert für α durch Vergleich der Anpassung alternativer Werte zwischen 0 und 1 in einem Stützbereich

Kriterien: Mean Square Error (MSE) oder Root Mean Square Error (RMSE)

(2.36) $$\mathrm{MSE}(e) = \frac{1}{n}\sum_{t=1}^{n}\left(y_t - \hat{y}_{t-1,1}\right)^2 = \frac{1}{n}\sum_{t=1}^{n} e_t^2$$

(2.37) $$\mathrm{RMSE} = \sqrt{\mathrm{MSE}}$$

In der Praxis wird häufig ein α-Wert zwischen 0,1 und 0,3 gewählt. In diesem Fall sind auch weiter zurückliegende Zeitreihenwerte für die Prognose bedeutsam. Bei einer sich allmählich verändernden zentralen Tendenz einer Zeitreihe, empfiehlt sich die Wahl eines größeren α-Wertes oder der Übergang zu einer exponentiellen Glättung zweiter Ordnung.

Abb.: Verhalten der Gewichtsfunktion $(1-\alpha)^i \cdot \alpha$ bei alternativen Werten von α

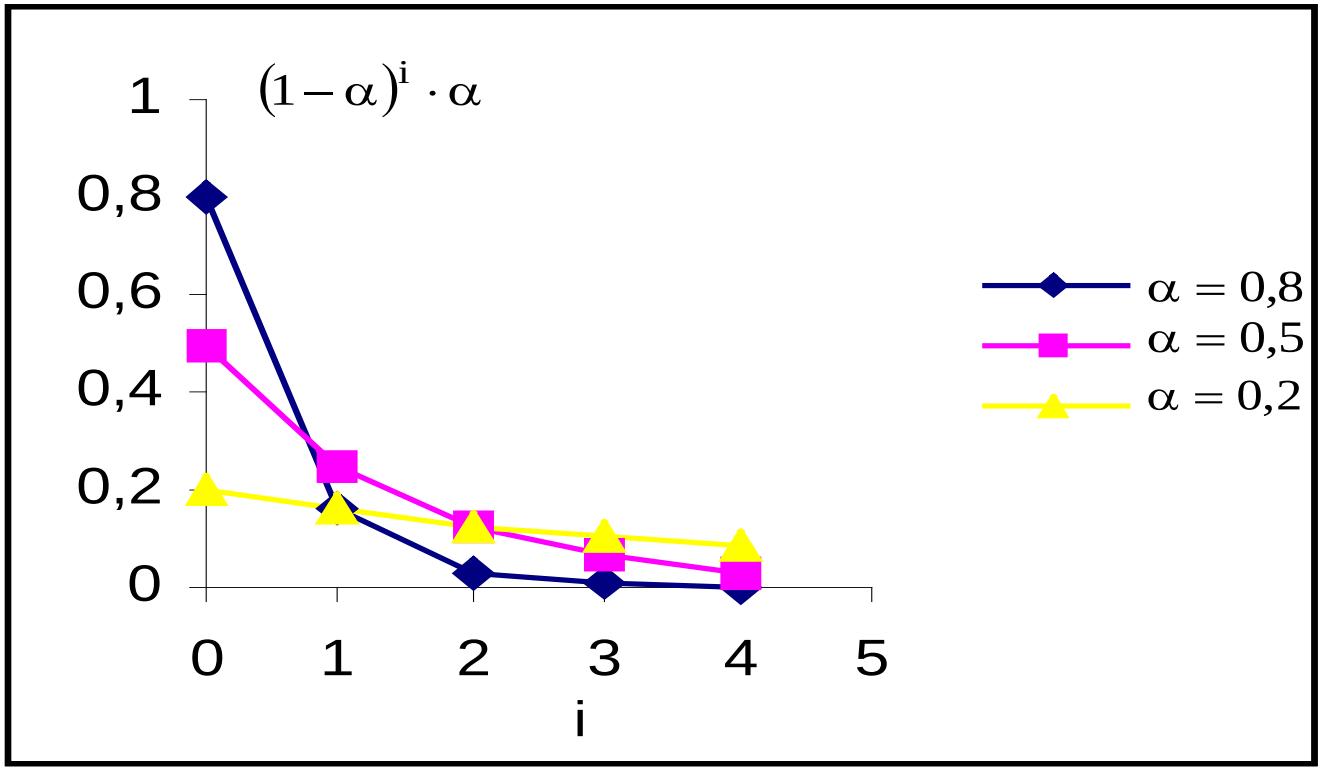
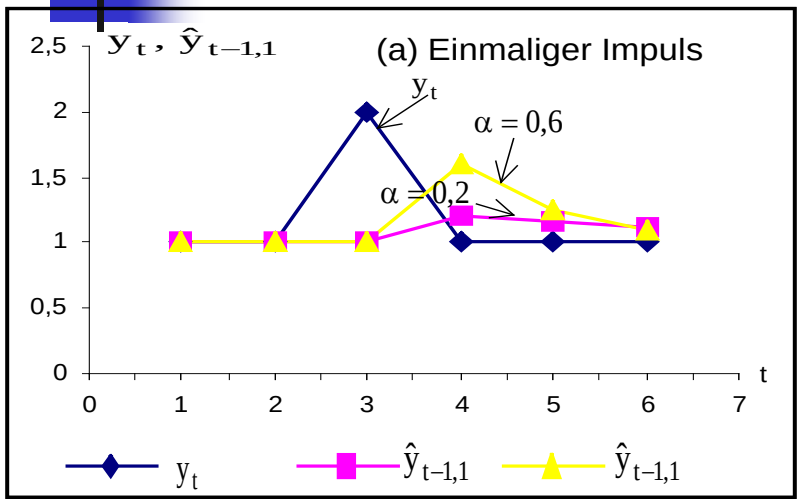
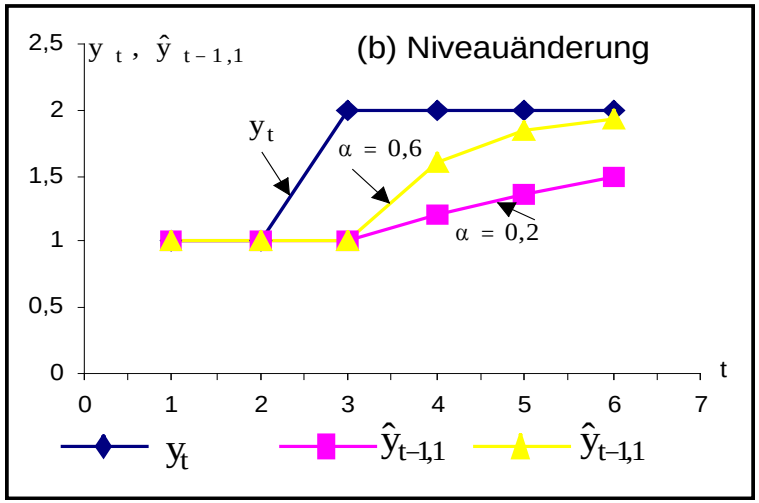
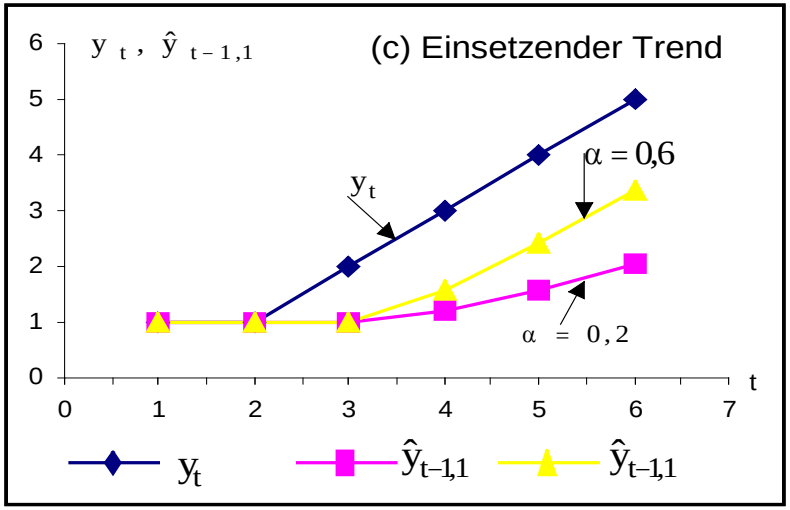


Abb.: Reaktion der Vorhersage auf verschiedene Ereignisse bei alternativem Reaktionsparameter







• Exponentielle Glättung zweiter Ordnung

A. Prinzip des exponent. Glättens 2. Ordnung (einparametriges Schema):
Neben einem Grundwert a_t wird in dem Prognosewert für die Zeit t+1 noch explizit ein Term b_t berücksichtigt, der den Trendanstieg widerspiegeln soll:

(2.38) $\tilde{y}_{t,1} = a_t + b_t$ (Ein-Schritt-Prognose)

Die Berechnung von a_t und b_t erfolgt mittels eines exponentiellen Glättens zweiter Ordnung:

(2.39) $\hat{y}_t = (1-\alpha)\cdot\hat{y}_{t-1} + \alpha\cdot y_t = \hat{y}_{t-1} + \alpha\cdot(y_t - \hat{y}_{t-1})$ (Glättungswert 1. Ordnung)

(2.40) $\hat{\hat{y}}_t = (1-\alpha)\cdot\hat{\hat{y}}_{t-1} + \alpha\cdot\hat{y}_t = \hat{\hat{y}}_{t-1} + \alpha\cdot(\hat{y}_t - \hat{\hat{y}}_{t-1})$ (Glättungswert 2. Ordnung)

Nach der Glättung der Ursprungswerte y_t („Mittelwerte 0. Ordnung“) wird eine analoge Glättung der geglätteten Werte $\hat{y}_t$ („Mittelwerte 1. Ordnung“) durchgeführt. Als Ergebnis erhält man die doppelt geglätteten Werte $\hat{\hat{y}}_t$ („Mittelwerte 2. Ordnung“).
Man erhält a_t und b_t aus den Beziehungen

(2.41) $a_t = \hat{y}_t + (\hat{y}_t - \hat{\hat{y}}_t) = 2\cdot\hat{y}_t - \hat{\hat{y}}_t$ und (2.42) $b_t = \frac{\alpha}{1-\alpha}(\hat{y}_t - \hat{\hat{y}}_t)$

Aus (2.41) und (2.42) erhält man die Anfangsschätzer

(2.43) $\hat{y}_0 = a_0 - \frac{1-\alpha}{\alpha}b_0$ und (2.44) $\hat{\hat{y}}_0 = a_0 - 2\frac{1-\alpha}{\alpha}b_0$

B. Holt-Verfahren (zweiparametriges Schema):

Exponentielle Glättung unter Berücksichtigung des Niveaus und eines Trends
(2 Glättungsparameter: α, β)

Niveau: $a_t = \alpha \cdot y_t + (1 - \alpha)(a_{t-1} + b_{t-1})$
Trend: $b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$

Prognose: $\tilde{y}_{t,h} = a_t + b_t \cdot h + s_{t-j+m}$ (h-Schritt-Prognose)

- **Holt-Winters-Verfahren**

Exponentielle Glättung unter Berücksichtigung des Niveaus, eines Trends und der Saison
(3 Glättungsparameter: α, β, γ, Phase: j z.B. Quartal)

Niveau: $a_t = \alpha(y_t - s_{t-j}) + (1 - \alpha)(a_{t-1} + b_{t-1})$
Trend: $b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$
Saison: $s_t = \gamma(y_t - a_t) + (1 - \gamma)s_{t-j}$

Prognose: $\tilde{y}_{t,h} = a_t + b_t \cdot h + s_{t-j+m}$ (h-Schritt-Prognose)

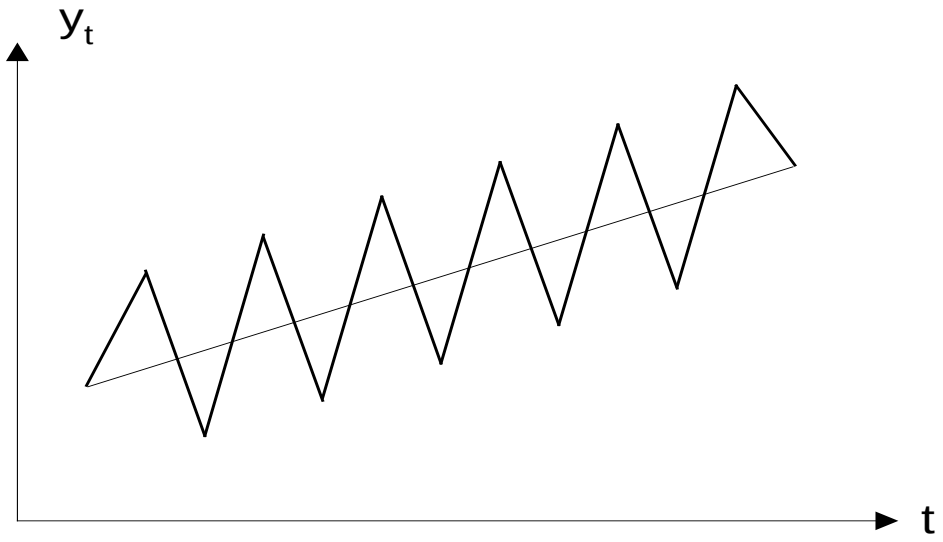
2.3 Saisonkomponente und Saisonbereinigung
2.3.1 Phasendurchschnittsverfahren

Zur Bestimmung der Saisonkomponente einer unterjährigen Zeitreihe (y_t) schaltet man die glatte Komponente vorab aus. Beim additiven Modell liegen dann die trendbereinigten (und konjunkturbereinigten) Zeitreihenwerte

(2.45) $$d_t = y_t - g_t = s_t + u_t$$

vor, die nur noch die Saison- und Restkomponente enthalten. Voraussetzung hierfür ist, dass die saisonalen Ausschläge unabhängig vom Trend der Zeitreihe sind (s. Abb.).

Abb.: Saisonausschläge mit konstanter Amplitude



Bestimmung der Saisonkomponente

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

Trendbereinigte Zeitreihe:

(2.46) $d_{ij} = y_{ij} - g_{ij} = s_j + u_{ij}, \quad i = 1, 2, \ldots, k_j; \quad j = 1, 2, \ldots, p$

k_j: Anzahl der Jahre für die Phase j
p: Anzahl der Phasen
bei Quartalsdaten: p=4, bei Monatsdaten: p=12

Unnormierte Saisonkomponente:

(2.47) $s_j^* = \frac{1}{k_j} \sum_i d_{ij}$

Die unnormierten Saisonkomponenten s_j^* summieren sich allgemein nicht zu null. Diese Tatsache kann die Interpretation einer Phase als saisonal unter- oder über-durchschnittlich erschweren. Für die Saisonkomponente s_j fordert man daher

(2.48) $\sum_{j=1}^{p} s_j = 0 \, (\text{Normierung})$

Normierte Saisonkomponente:

(2.49) $s_j = s_j^* - \bar{d}$ mit (2.50) $\bar{d} = \frac{1}{p} \sum_{j=1}^{p} s_j^*$

Saisonbereinigte Zeitreihe: (2.51) $y_{ij}^* = y_{ij} - s_j$

Beispiel: Im Zeitreihendiagramm der Lohn- und Gehaltseinkommen je Beschäftigten ist ein klares Saisonmuster erkennbar. Der Index weist jahreszeitlich bedingt jeweils im ersten Quartal eines Jahres einen Tiefstand und im vierten Quartal ein Hoch aus. Außerdem sind keinerlei Anhaltspunkte dafür erkennbar, dass die saisonalen Ausschläge mit wachsendem Trend zunehmen. Daher lässt sich eine Zeitreihenzerlegung adäquat auf der Grundlage des additiven Modells vornehmen.

Jahr (i)	Quartal (j)	y_{ij}	$\bar{y}^4_{ij}$	d_{ij}
1	I. (1)	113,6		
1	II. (2)	121,3		
1	III. (3)	122,0	124,3	-2,3
1	IV. (4)	138,8	125,2	13,6
2	I. (1)	116,3	126,2	-9,9
2	II. (2)	125,7	127,2	-1,5
2	III. (3)	125,7	128,4	-2,7
2	IV. (4)	143,5	129,4	14,1
3	I. (1)	121,1	130,1	-9,0
3	II. (2)	128,6	131,0	-2,4
3	III. (3)	129,0	131,8	-2,8
3	IV. (4)	147,3	132,1	15,2
4	I. (1)	123,2	132,3	-9,1
4	II. (2)	129,2	132,6	-3,4
4	III. (3)	130,3	133,3	-3,0
4	IV. (4)	147,9	134,7	13,2
5	I. (1)	128,0	136,2	-8,2
5	II. (2)	135,7	137,9	-2,2
5	III. (3)	136,2		
5	IV. (4)	155,5		

Zur Bestimmung der glatten Komponente haben wir zentrierte gleitende Durchschnitte der Ordnung 4 verwendet, die bereits bei der Behandlung der Methode der gleitenden Durchschnitte (Abschn. 2.2.2) ermittelt worden sind.

Aus den in der Arbeitstabelle berechneten trendbereinigten Werte d_{ij} erhält man die unnormierte Saisonkomponente:

$$s_1^* = \frac{1}{4}\sum_{i=2}^{5} d_{i1} = \frac{1}{4}(-9{,}9 - 9{,}0 - 9{,}1 - 8{,}2) = -\frac{36{,}2}{4} = -9{,}1,$$

$$s_2^* = \frac{1}{4}\sum_{i=2}^{5} d_{i2} = \frac{1}{4}(-1{,}5 - 2{,}4 - 3{,}4 - 2{,}2) = -\frac{9{,}5}{4} = -2{,}4,$$

$$s_3^* = \frac{1}{4}\sum_{i=1}^{4} d_{i3} = \frac{1}{4}(-2{,}3 - 2{,}7 - 2{,}8 - 3{,}0) = -\frac{10{,}8}{4} = -2{,}7,$$

$$s_4^* = \frac{1}{4}\sum_{i=1}^{4} d_{i4} = \frac{1}{4}(13{,}6 + 14{,}1 + 15{,}2 + 13{,}2) = \frac{56{,}1}{4} = 14{,}0.$$

Durchschnitt der unmormierten Saisonziffern:

$$\bar{d} = \frac{1}{p}\sum_{j=1}^{p} s_j^* = \frac{1}{4}[(-9{,}1) + (-2{,}4) + (-2{,}7) + 14{,}0] = \frac{1}{4}(-0{,}2) = -0{,}05$$

Normierte Saisonkomponente:

$$s_1 = s_1^* - \bar{d} = -9{,}1 - (-0{,}05) = -9{,}05$$

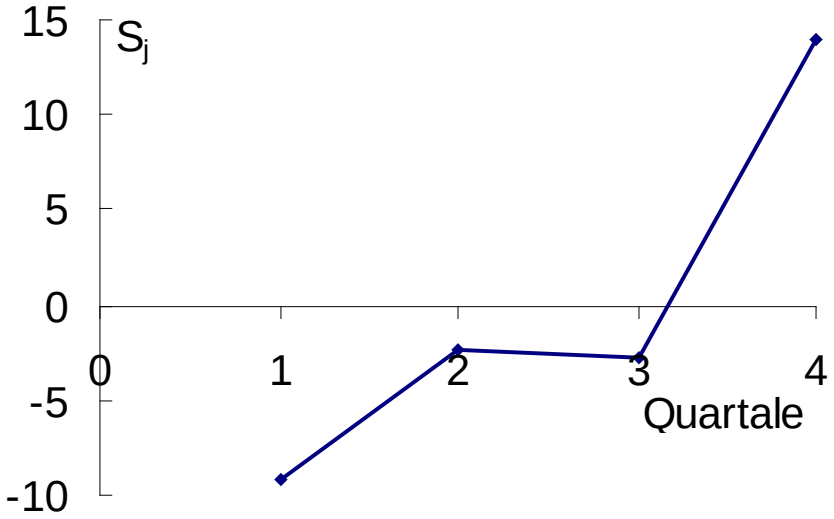
$$s_2 = s_2^* - \bar{d} = -2{,}4 - (-0{,}05) = -2{,}35$$

$$s_3 = s_3^* - \bar{d} = -2{,}7 - (-0{,}05) = -2{,}65$$

$$s_4 = s_4^* - \bar{d} = 14{,}0 - (-0{,}05) = 14{,}05$$

Das Saisonprofil gibt die Größenordnung des saisonalen Einflusses in der Zeitreihe der Löhne und Gehälter je Beschäftigten grafisch wieder.

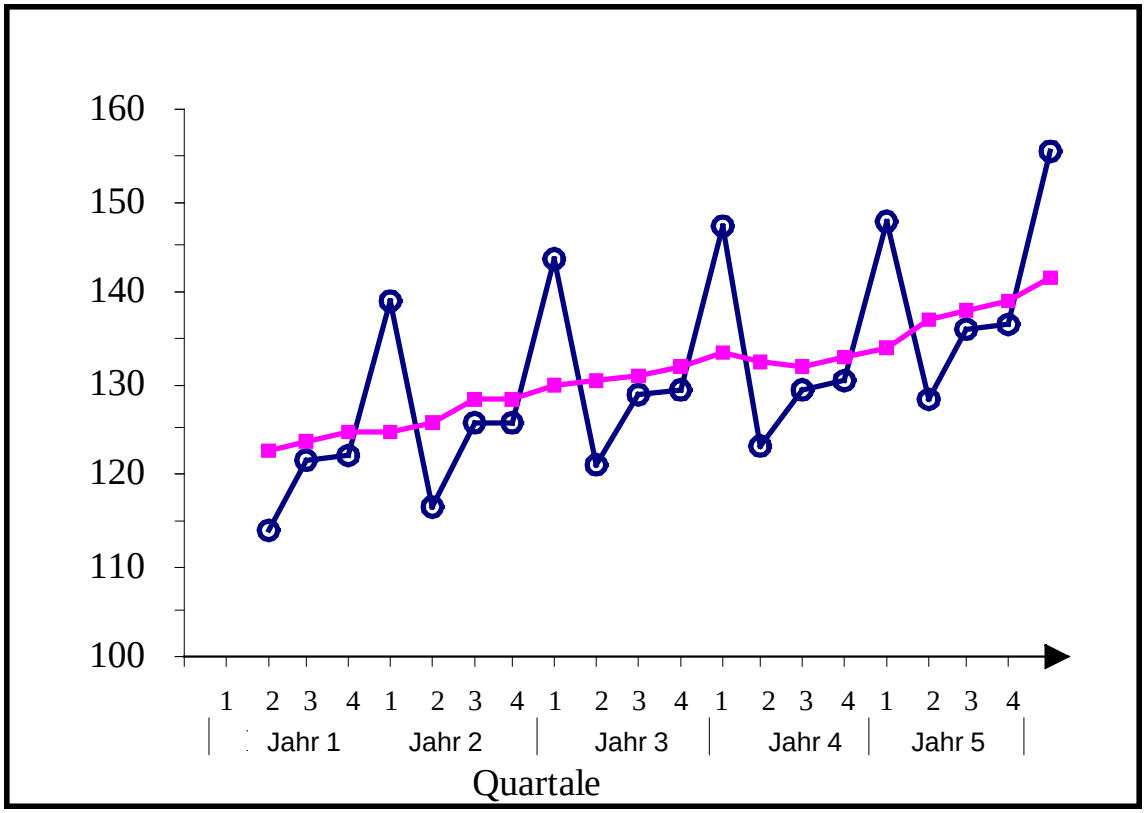
Abb.: Saisonprofil der Löhne und Gehälter je Beschäftigten



Unter Verwendung der Saisonziffern s_j lässt sich die saisonbereinigte Zeitreihe $\left(y_{ij}^*\right)$ der Löhne und Gehälter je Beschäftigten für den gesamten Beobachtungszeitraum bestimmen:

Jahr (i)	Quartal (j)	y_{ij}	s_j	y_{ij}^*
1	I.(1)	113,6	-9,05	122,65
1	II.(2)	121,3	-2,35	123,65
1	III.(3)	122,0	-2,65	124,65
1	IV.(4)	138,8	14,05	124,75
2	I.(1)	116,3	-9,05	125,35
2	II.(2)	125,7	-2,35	128,05
2	III.(3)	125,7	-2,65	128,35
2	IV.(4)	143,5	14,05	129,45
3	I.(1)	121,1	-9,05	130,15
3	II.(2)	128,6	-2,35	130,95
3	III.(3)	129,0	-2,65	131,65
3	IV.(4)	147,3	14,05	133,25
4	I.(1)	123,2	-9,05	132,25
4	II.(2)	129,2	-2,35	131,55
4	III.(3)	130,3	-2,65	132,95
4	IV.(4)	147,9	14,05	133,85
5	I.(1)	128,0	-9,05	137,05
5	II.(2)	135,7	-2,35	138,05
5	III.(3)	136,2	-2,65	138,85
5	IV.(4)	155,5	14,05	141,45

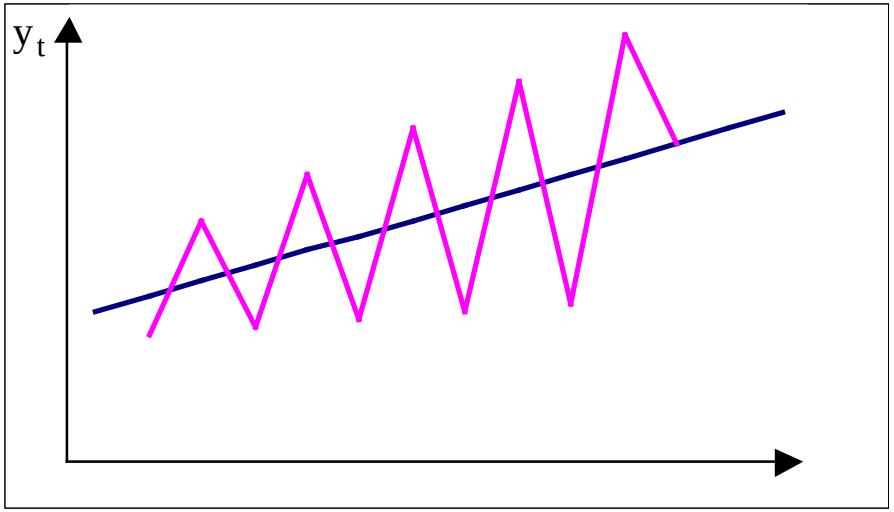
Abb.: Löhne und Gehälter je Beschäftigten mit saisonbereinigter Zeitreihe



Multiplikatives Modell

Das **multiplikative Modell** ist für ökonomische Zeitreihen adäquat, deren Saisonausschläge mit steigendem Trend im Mittel proportional zunehmen.

Abb.: Saisonausschläge mit proportional zunehmender Amplitude



Da die Komponenten in den Logarithmen additiv sind, kann die Saisonkomponente im Prinzip wie im additiven Modell bestimmt werden, wenn man die originäre Zeitreihe (y_t) durch die logarithmierte Zeitreihe ($\log y_t$) ersetzt.

Aus der originären Zeitreihe (y_t) lassen sich die **Saisonkomponente** und die **saison-bereinigte Zeitreihe** wie folgt bestimmen.

Trendbereinigte Werte: (2.52) $d_{ij} = \frac{y_{ij}}{g_{ij}} = s_j \cdot u_{ij},\ i = 1,2,\ldots,k_j,\ j = 1,2,\ldots,p$

Unnormierte Saisonkomponente: (2.47) $s_j^* = \frac{1}{k_j} \sum_i d_{ij}$

(gewöhnlich wie im additiven Modell unter Verzicht auf eine geometrische Mittelung)

Für die normierte Saisonkomponente gilt approx. die Eigenschaft

(2.52) $s_1 \cdot s_2 \cdot \ldots \cdot s_p = 1$ (Normierung)

Normierte Saisonkomponente:

(2.53) $s_j = \frac{s_j^*}{\bar{d}}$

$\bar{d}$ wird häufig wie beim additiven Modell als arithmetisches Mittel nach Gl. (2.50) berechnet. Sachlogisch sit jedoch eine geometrische Mittelung $\bar{d} = \prod_{j=1}^{p} s_j^*$ vorzuziehen.

Saisonbereinigte Zeitreihe:

(2.54) $y_{ij}^* = \frac{y_{ij}}{s_j}$

2.3.2 Regressionsverfahren

Mit dem Regresssionsverfahren wird der saisonale Einfluss mittels sog. **Saison-Dummies** (0,1-Variablen) bestimmt. In der Regel wird mit einem Regressionsmodell jedoch nicht allein die Saisonkomponente ermittelt, sondern simultan z.B. mit dem Trend oder exogenen Variablen.

A. Regressionsmodell mit absolutem Glied

In einem **Regressionsmodell mit absolutem Glied** muss die **Anzahl der Saison-Dummies** gleich **p-1** gesetzt werden. Der Grund liegt darin, dass sich die Spalten von p Saison-Dummies in der Beobachtungsmatrix („Designmatrix“) **X** zum Wert von 1 der Scheinvariablen summieren, die dem absoluten Glied zugrunde liegt. In diesem Fall entstünde exakte Multikollinearität, so dass das Regressionsmodell nicht mehr schätzbar wäre.

Bei Einbeziehung eines linearen Trends lässt sich die Saisonkomponente auf der Basis von für Quartalsdaten aus dem Regressionsmodell

(2.55) $y_t = ß_1 + ß_2 \cdot D_2 + ß_3 \cdot D_3 + ß_4 \cdot D_4 + ß_5 \cdot t + u_t$.

bestimmen. D_j ist hierin eine Saison-Dummy für das j-te Quartal:

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

Hier ist das 1. Quartal als **Referenzquartal** gewählt worden. In diesem Fall geben die Koeffizienten $ß_2$, $ß_3$ und $ß_4$ geben die Saisoneinflüsse der Quartale 2, 3 und 4 relativ zum 1. Quartal an.

Regressionsmodell (2.55) in Matrizenform: (2.57) $\mathbf{y} = \mathbf{X} \cdot \mathbf{ß} + \mathbf{u}$

y: nx1-Vektor der Zeitreihenwerte (bei kompletten Jahren: n=p·k)
X: nx5-Beobachtungsmatrix („Designmatrix“) bei Quartalsdaten
ß: 5x1-Vektor der Regressionskoeffizienten, $\mathbf{ß} = (ß_1 \;\; ß_2 \;\; ß_3 \;\; ß_4 \;\; ß_5)'$
u: nx1-Vektor der Störvariablen

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

x_1: Scheinvariable
x_2: Trendvariable

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 2 \\ 1 & 0 & 1 & 0 & 3 \\ 1 & 0 & 0 & 1 & 4 \\ 1 & 0 & 0 & 0 & 5 \\ 1 & 1 & 0 & 0 & 6 \\ 1 & 0 & 1 & 0 & 7 \\ 1 & 0 & 0 & 1 & 8 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ x_1 & D_2 & D_3 & D_4 & x_2 \end{bmatrix}$$

Der **Kleinst-Quadrate-Schätzer** (OLS-Schätzer, ordinary least squares estimator)

(2.58) $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$

enthält den Trendkoeffizienten sowie die gesuchten Regressionsschätzer für den Saisoneinfluss der Phasen (hier: Quartale) 2, 3, 4 relativ zur 1. Phase.

Man erhält die (normierte) **Saisonkomponente**, d.h. normierte Saisonziffern S_1, S_2, ..., S_p für alle vier Quartale wie folgt:

(2.59) $$S_j = \begin{cases} -\bar{d} \text{ für } j = 1 (\text{Referenzphase}) \\ \hat{\beta}_j - \bar{d} \text{ für } j = 2, \ldots, p \end{cases}$$

(2.60) $$\bar{d} = \frac{\left(\sum_{j=2}^{p} \hat{\beta}_j \right)}{p}$$ (1. Phase: Referenzphase)

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

Regressionsgleichung: $\widehat{\text{Lohn}}$ = 112,6 + 6,8 ·D2 + 6,5 ·D3 + 23,6 ·D4 + 0,8675 ·LTREND

R^2=0,990 t-Werte → (158,0) (8,595) (8,134) (29,356) (17,598)

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

$\bar{d}$ =(0 + 6,8 + 6,5 + 23,6)/4 = 36,9/4 = 9,2

(Normierte) **Saisonkomponente**:

S_1 = 0 – 9,2 = -9,2; S_2= 6,8 – 9,2 = -2,4; S_3 = 6,5 – 9,2 = -2,7; S_4 = 23,6 – 9,2 = 14,4

B. Regressionsmodell ohne absolutes Glied

In einem Regressionsmodell ohne absolutes Glied können unter Verwendung von p Sasion-Dummies die Saisonkoeffizienten aller p Phasen eines Jahres ohne Bezug zu einer Referenzphase geschätzt werden. Für Quartalsdaten (p=4) liegt der Regressionsschätzung der Saisonkoeffizienten bei Einbeziehung eines linearen Trends das Regressionsmodell

(2.61) $y_t = ß_1 \cdot D_1 + ß_2 \cdot D_2 + ß_3 \cdot D_3 + ß_4 \cdot D_4 + ß_5 \cdot t + u_t$

zugrunde. Die Saison-Dummies D_j sind hierbei durch

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

definiert.

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

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

x_1: Trendvariable

Zur Bestimmung der (normierten) **Saisonkomponente**, d.h. der normierte Saisonziffern S_1, S_2, ..., S_p bestimmt man den Durchschnittschnittswert der geschätzten Saisonkoeffizienten:

(2.63) $$\bar{d} = \left(\sum_{j=1}^{p} \hat{\beta}_j \right) \Big/ p$$

Die (normierte) Saisonkomponente ergibt sich aus den geschätzten Saisonkoeffizienten $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3 \text{ und } \hat{\beta}_4$ nach Subtraktion des Durchschnitts $\bar{d}$:

(2.64) $$S_j = \hat{\beta}_j - \bar{d}\ , \qquad j=1,2,3,4.$$

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

Regr.gleich.: $\widehat{\text{Lohn}}$ = 112,6·D1 + 119,4 ·D2 + 119,1 ·D3 + 136,2 ·D4 + 0,8675 ·LTREND

R^2=0,990, t-Werte → (158,0) (160,4) (153,1) (167,5) (17,598)

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

$\bar{d}$ =(112,6 + 119,4 + 119,1 + 136,2)/4 = 487,3/4 = 121,8

(Normierte) **Saisonkomponente**:

S_1 = 112,6 – 121,8 = -9,2; S_2= 119,4 – 121,8 = -2,4;

S_3 = 119,1 – 121,8 = -2,7; S_4 = 136,2 – 121,8 = 14,4