Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences

CAS Datenanalyse

Kapitel 3: Stochastische Trends und Korrelogramm

Prof. Dr. Raúl Gimeno
FRM,CAIA, PRM

Agenda

- Momente eines stochastischen Prozess
- Stationäre Prozesse
- White-Noise-Prozess
- Martingale-Prozess
- Random-Walk
- Residualkomponente
- Korrelogramm
- Portmanteau –Test
- Ljung-Box Test

Momente eines stochastischen Prozesses

Ein stochastischer Prozess ist eine Folge von Zufallsvariablen, $(X_t)_{t\in T}$ in der T die Menge der Zeitpunkte bezeichnet, für die der Prozess definiert ist.

Mittelwertfunktion μ_t : (1) $\mu_t = E(X_t)$

Varianzfunktion σ_t^2 : (2) $\sigma_t^2 = \mathrm{var}(X_t)$

Autokovarianzfunktion $\gamma_{t,s}$ (ACVF): (3) $\gamma_{t,s} = \mathrm{cov}(X_t, X_s) = E[(X_t - \mu_t)\cdot(X_s - \mu_s)]$

Autokorrelationsfunktion (ACF) (4) $\rho_{t,s} = \gamma_{t,s}/(\sigma_t\sigma_s)$

Interpretation:

- Die Mittelwertfolge (μ_t) gibt die durchschnittliche Zeitfolge an, um die die Realisierungen des Prozesses schwanken → Mittelung über alle Zeitfolgen
- Zeitliches Mittel: Mittelung der einzelnen realisierten Beobachtungen über alle t: $$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$$
- Die Varianzfolge (σ_t^2) gibt für jeden Zeitpunkt t an, in welchem Ausmass die Zufallsvariable X_t um den Wert der Mittelwertfolge (μ_t) streut.

Streng stationärer stochastischer Prozess

Ein stochastischer Prozess X_t heisst schwach stationär, wenn Erwartungswert, Varianz und (Auto)kovarianz des Prozesses unabhängig von t sind.

Schwach stationärer Prozess:

(5) $\mu_t = E(X_t) = \mu$ für alle t → Mittelwertstationär

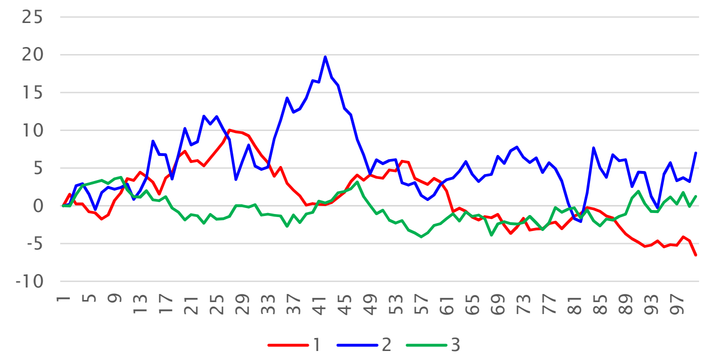
(6) $\mathrm{var}(X_t) = \sigma_t^2 = \sigma^2$ für alle t → Varianzstationär

(7) $\gamma_{t,s} = \mathrm{cov}(X_t, X_s) = \gamma_{t-s}$

Kovarianz zwischen X_t und X_s ist nur von der zeitlichen Differenz abhängig, nicht jedoch von den Zeitpunkten.

- Für die ACF folgt damit
- (8) $\rho_{t,s} = \rho_{t-s} = \gamma_{t-s}/\gamma_0$ mit $\gamma_0 = \sigma^2$
- (8‘) $\rho_{t,s} = \rho_\tau = \gamma_\tau/\gamma_0$ $\tau = t - s$

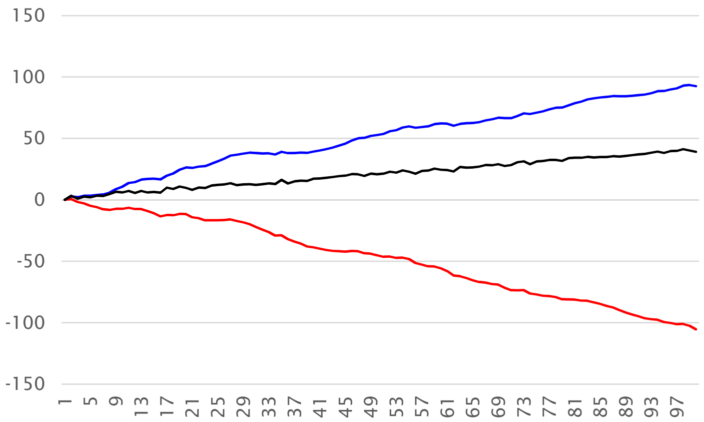
Beispiele für nicht-stationäre Prozesse



$x_t = x_{t-1} + u_{1t}$ $\qquad u_{1t} \approx N(0,4)$

$x_t = x_{t-1} + u_{2t}$ $\qquad u_{2t} \approx N(0,1)$

$x_t = x_{t-1} + u_{2t}$ $\qquad u_{2t} \approx N(0,1)$



$x_t = 1 + x_{t-1} + u_{1t}$

$x_t = 1 + 0.4t + u_{1t}$

$x_t = -1 + x_{t-1} + u_{1t}$

White-Noise-Prozess

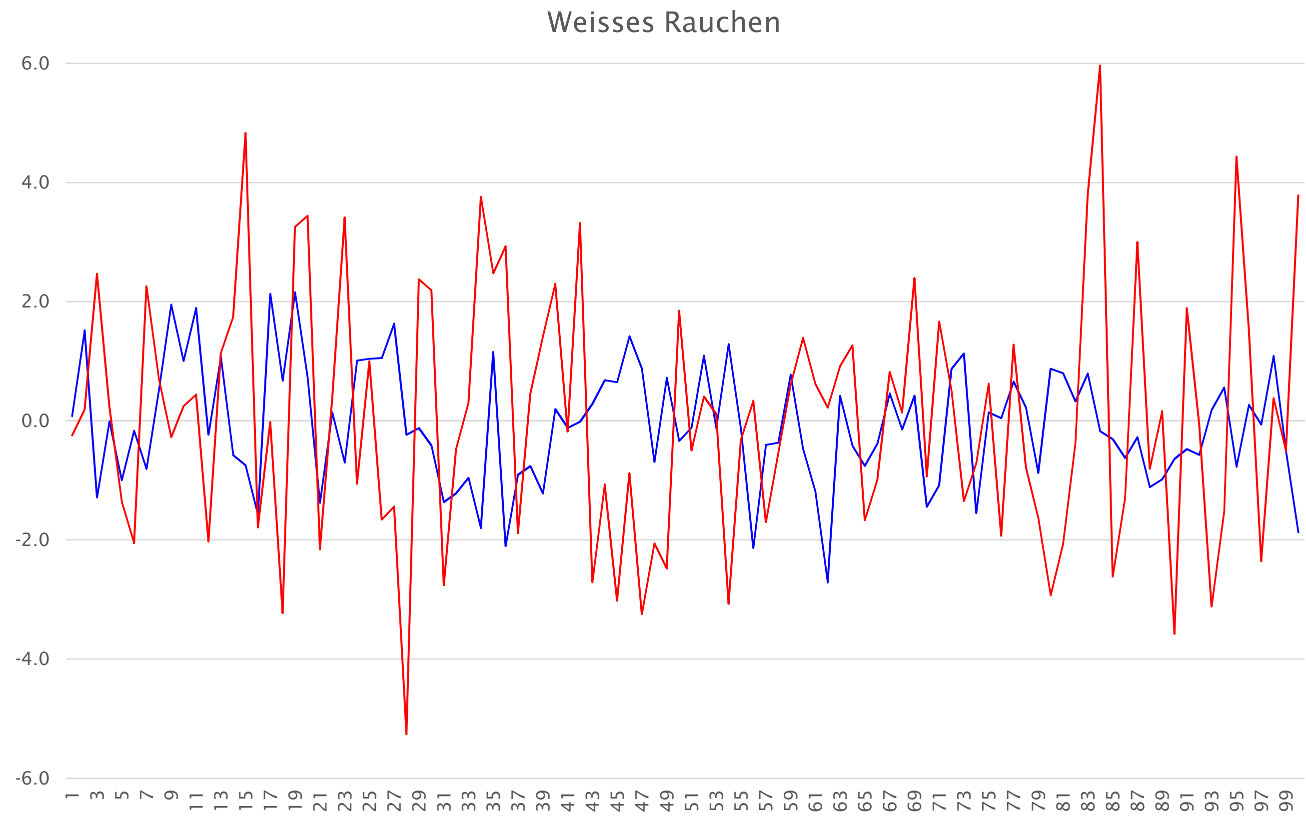
White-Noise-Prozess (weisses Rauschen) U_t, wenn er die Eigenschaften besitzt:

$$E(U_t) = \mu \qquad [i.a.\ wird\ E(U_t) = 0\ gesetzt],$$
$$\operatorname{var}(U_t) = \sigma^2$$
$$\operatorname{cov}(U_t, U_s) = 0 \qquad f\ddot{u}r\ t \neq s$$

Beim White-Noise-Prozess werden allein die beiden ersten Momente betrachtet. Da die Autokovarianzen verschwinden, besteht keine lineare Beziehung zwischen den vergangenen, aktuellen und zukünftigen Realisationen der Zufallsvariablen $U_t \rightarrow U_{t+h}$ $h > 0$ kann auf der Basis eines linearen Zeitreihenmodells nicht prognostiziert werden.

Wenn allerdings höhere Momente des stochastischen Prozesses (U_t) ungleich null sind, könnte (U_{t+h}) ggf. unter Verwendung eines nichtlinearen Modells vorhergesagt werden.

Beispiel



White-Noise-Prozess

Unabhängige und identische Verteilung (i.i.d.)

Eine Zufallsvariable U_t ist unabhängig und identisch verteilt [independent and identically distributed (= i.i.d.)], wenn alle Terme zeitlich unabhängig sind und dieselbe Verteilung haben. Die Dichtefunktionen sind dann für alle t identisch, $f_t(U_t) = f(U_t)$ für alle t,

und die gemeinsame Dichtefunktion $f_{1,2,\ldots,T}(U_1, U_2, \ldots, U_T)$

ist dann gleich dem Produkt der marginalen Dichtefunktionen $f(U_t)$:

$$f_{1,2,\ldots,T}(U_1, U_2, \ldots, U_T) = \prod_{t=1}^{T} f_t(U_t) = [f(U_t)]^T$$

Die Kenntnis vergangener und aktueller Werte von (U_t) liefert dann keine prognostisch verwertbaren Informationen für U_{t+h}, h> 0.

Martingale-Prozess

Ein stochastischer Prozess (X_t) mit der Eigenschaft $E(X_{t+1}|X_t, X_{t-1}, X_{t-2}, \ldots) = X_t$ (9)

Wenn X_t der bekannte aktuelle Aktienkurs ist, dann würde für die Periode t+1 bei einem Martingale-Prozess unter Kenntnis aller vergangenen Kursinformationen genau derselbe Kurs zu erwarten sein. Als "tomorrow's price" ist stets "today's price" zu erwarten.

Die Verwertung vergangener Kursrealisationen führt zu keiner verbesserten Kursprognose → die beste Kursprognose im Sinne des Mean Square Errors ist stets der aktuelle Kurs.

Alternative Form der Martingale-Eigenschaft: $E(\Delta X_{t+1}|X_t, X_{t-1}, X_{t-2}, \ldots) = 0$

Martingale-Differenz: $\Delta X_{t+1} = X_{t-1} - X_t$

Die Martingale-Differenz hat hinsichtlich des Erwartungswerts und der Kovarianzen dieselben Eigenschaften wie der White-Noise-Prozess, doch ist die Varianz nicht notwendig konstant.

Random Walk

Es sei U_t eine unabhängig identisch verteilte Zufallsvariable (i.i.d.) mit den Parametern $E(U_t) = 0$ und $\mathrm{var}(U_t) = \sigma^2$

$X_t = X_{t-1} + U_t$

Dann folgt die Zufallsvariable X_t einem **Random Walk** (Irrfahrt) → der aktuelle Wert einer Zufallsvariablen ergibt sich aus dem Vorperiodenwert plus einer Realisation einer i.i.d.-Zufallsvariablen.

Auf diese Weise ergibt sich ein Zufallspfad, der nicht-stationär ist, d.h. sich beliebig von seinem Mittelwert entfernen kann → Irrfahrt z.B. bei Aktienkursen.

Um den Erwartungswert und die Varianz des Prozesses zu ermitteln, setzen wir den Anfangswert $X_0 = 0$ ohne Einschränkung der Allgemeinheit.

Dann ergibt sich die Folge (X_t) aus

$$\begin{aligned} X_1 &= U_1 \\ X_2 &= U_1 + U_2 \\ X_3 &= U_1 + U_2 + U_3 \\ &\vdots \\ X_t &= U_1 + U_2 + \ldots + U_t \end{aligned} \qquad (1.10)$$

Random Walk

(11) $\mathrm{E}(X_t) = \mathrm{E}(U_1 + U_2 + \ldots + U_t) = \mathrm{E}(U_1) + \mathrm{E}(U_2) + \ldots + \mathrm{E}(U_t) = 0$

Unbekannte Varianz des Prozesses:

(12) $\mathrm{Var}(X_t) = \mathrm{Var}(U_1 + U_2 + \ldots + U_t) = \mathrm{Var}(U_1) + \mathrm{Var}(U_2) + \ldots + \mathrm{Var}(U_t) = t \cdot \sigma^2$

Für $t \to \infty$, d.h. bei unendlich lang andauerndem Prozess geht die Varianz gegen unendlich.

Random Walk: Erwartungswert zeitlich konstant → mittelwertstationär
Varianz nimmt mit t zu → zeitvariabel → nicht varianzstationär

(13) $\mathrm{cov}(X_t X_s) = t \cdot \sigma^2, \quad 0 \leq t \leq s$ → nicht kovarianzstationär

Kovarianz und die Varianz hängen vom Zeitindex t ab.

Autokorrelationsfunktion des Random Walk (14): $\rho_{t,s} = \sqrt{t/s}$

→Funktion der zeitlichen Differenz |s–t| und von den konkreten Zeitperioden t und s abhängig, für die sie betrachtet wird.

Random Walk

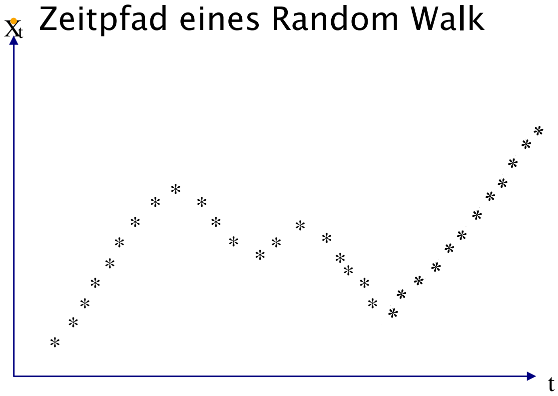
Durch Differenzenbildung lässt sich der nichtstationäre Prozess (X_t) in einen stationären Prozess überführen:

(15) $\Delta X_t = U_t$

→ (ΔX_t) = i.i.d.-Prozess, dessen Parameter mit denen des i.i.d.-Prozesses (U_t) übereinstimmen.

Der bedingte Erwartungswert $\mathrm{E}(X_t | X_{t-1}, X_{t-2}, \ldots, X_1)$ ist unabhängig von $X_{t-2}, X_{t-3}, \ldots, X_1$, da sich modellmässig alle Informationen bis zur Periode t–1 im Vorperiodenwert X_{t-1} widerspiegeln: $\mathrm{E}(X_t | X_{t-1}, X_{t-2}, \ldots, X_1) = \mathrm{E}(X_{t-1} + U_t) = X_{t-1} + \mathrm{E}(U_t) = X_{t-1}$ (16)

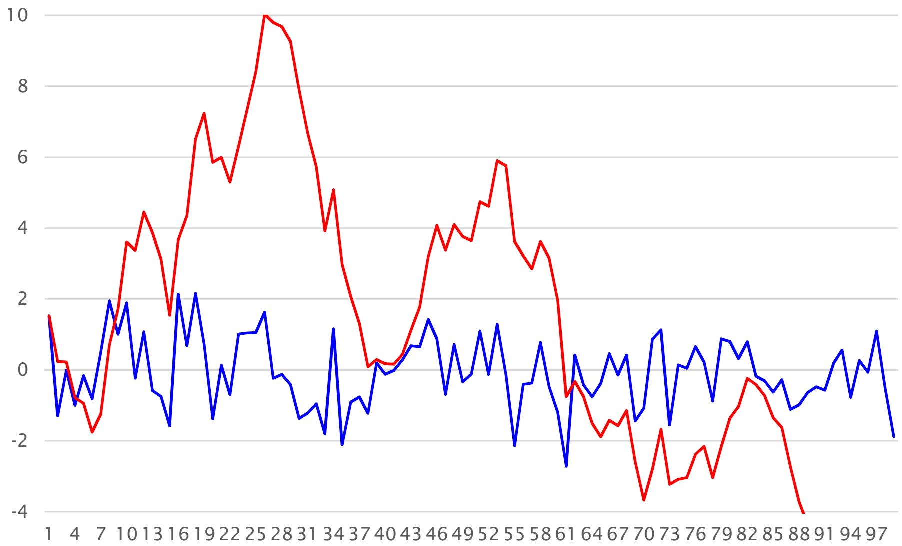
Zeitpfad eines Random Walk



Realisation eines Random Walk in einem Zeitreihendiagramm. Die im Zeitablauf zunehmende Varianz bewirkt, dass sich die Realisationen im Mittel immer weiter voneinander entfernen.

Beispiel

- Random Walk : $X_t = X_{t-1} + U_t$
- Differenz: $\Delta X_t = X_t - X_{t-1}$



Random Walk

$\hat{x}_{n+1}(n)$: Prognosewert für die Periode n+1 unter Berücksichtigung der gesamten Zeitreihenhistorie bis n → durch den bedingten Erwartungswert gegeben.

$$(17) \quad \hat{X}_{n+1}(n) = E(X_{n+1} | X_n, X_{n-1}, \ldots, X_1) = E(X_n + U_{n+1}) = X_n + E(U_{n+1}) = X_n$$

Die Ein-Schritt-Prognose entspricht genau dem letzten bekannten Zeitreihenwert. Analog ergibt sich der Prognosewert $\hat{x}_{n+2}(n)$ aus

$$(18) \quad \hat{X}_{n+2}(n) = E(X_{n+2} | X_n, X_{n-1}, \ldots, X_1) = E(X_{n+1} + U_{n+2})$$
$$= E(X_n + U_{n+1} + U_{n+2}) = X_n + E(U_{n+1}) + E(U_{n+2}) = X_n$$

Durch Verallgemeinerung : $\hat{X}_{n+h}(n) = E(X_{n+h} | X_n, X_{n-1}, \ldots, X_1) = X_n$ (19)

Der h-Schritt-Prognosewert stimmt damit bei einem Random Walk mit dem Ein-Schritt-Prognosewert überein.

Gleichwohl ist die h-Schritt-Prognose mit zunehmendem Zeithorizont mit einer immer grösseren Unsicherheit behaftet → zunehmende Varianz des Prognosefehlers. Bei der Ein-Schritt-Prognose lautet der **Prognosefehler $e_{n,1}$**

$$e_{n+1} = X_{n+1} - \hat{X}_{n+1}(n) = X_n + U_{n+1} - X_n = U_{n+1}$$

Random Walk

Prognosevarianz: $\operatorname{var}(e_{n,1}) = \operatorname{var}(U_{n+1}) = \sigma^2$

Der 2-Schritt-Prognosefehler:

$$e_{n,2} = X_{n+2} - \hat{X}_{n+2}(n) = X_n + U_{n+1} + U_{n+2} - X_n = U_{n+1} + U_{n+2}$$

Prognosevarianz:

$$\operatorname{var}(e_{n,2}) = \operatorname{var}(U_{t+1} + U_{t+2}) = \operatorname{var}(U_{t+1}) + \operatorname{var}(U_{t+2}) + \left[2 \cdot \underbrace{\operatorname{cov}(U_{t+1}, U_{t+2})}_{=0}\right] = \sigma^2 + \sigma^2 = 2\sigma^2 \qquad (20)$$

Varianz eines h-Schritt-Prognosefehlers $e_{n,h}$:

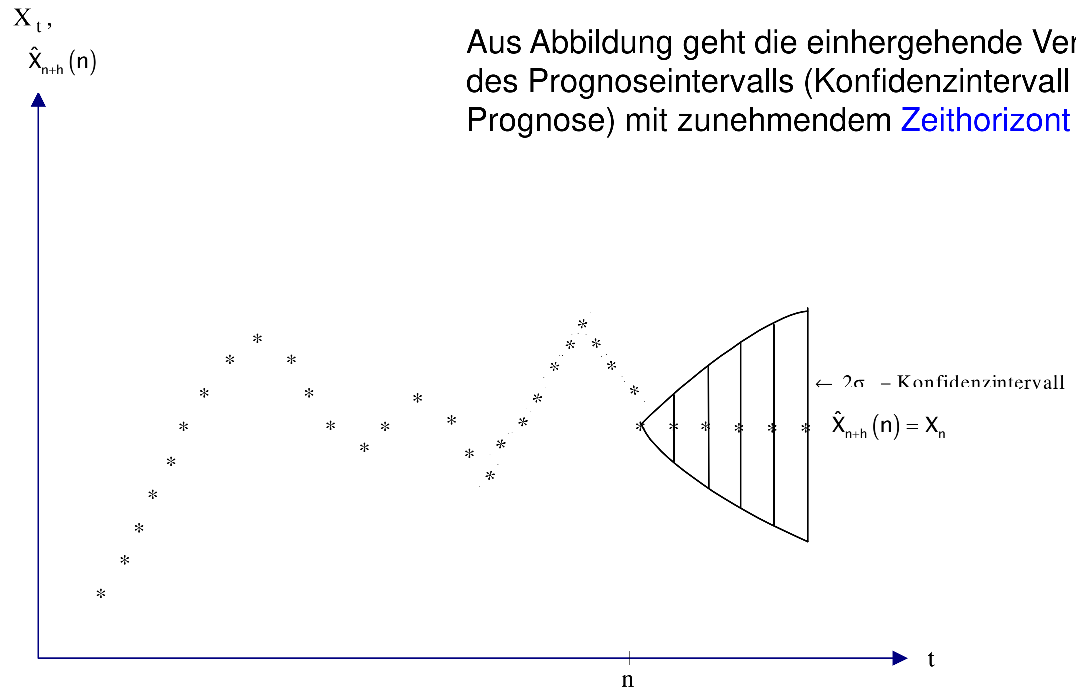
(21) $\operatorname{var}(e_{n,h}) = h \cdot \sigma^2$

→ Standardfehler der Prognose vergrössert sich in jeder Periode um den Faktor $\sqrt{h}$:

(22) $\sigma_e = \sqrt{h}\sigma$

Random Walk

Abbildung: Prognoseintervall



Aus Abbildung geht die einhergehende Verbreiterung des Prognoseintervalls (Konfidenzintervall der Prognose) mit zunehmendem Zeithorizont hervor

Random Walk

Bei einem Random Walk ohne Drift ist z.B. die erwartete Änderung eines Aktienkurses gegeben die Informationen bis zur Periode t-1 stets gleich null:

$$\mathrm{E}\left(\Delta x_t \middle| x_{t-1}, x_{t-2}, \ldots, x_1\right) = \mathrm{E}\left(\underbrace{x_{t-1} + u_t}_{x_t} - x_{t-1}\right) = \mathrm{E}\left(u_t\right) = 0$$

Trendmässig steigende Aktienkurse können durch Einführung eines Driftparameters berücksichtigt werden:

(23) $x_t = b + x_{t-1} + u_t$ (stochastischer Trend)

Die bedingt erwartete Kursänderung einer Aktie würde dann genau dem **Driftparameter** b entsprechen: $\mathrm{E}\left(\Delta x_t \middle| x_{t-1}, x_{t-2}, \ldots, x_1\right) = \mathrm{E}\left(b + x_{t-1} + u_t - x_{t-1}\right) = b + \mathrm{E}\left(u_t\right) = b$

Alternative Formulierungen der Random Walk-Hypothese beziehen sich auf eine Lockerung der Annahmen hinsichtlich des Zufallsprozesses (U_t).

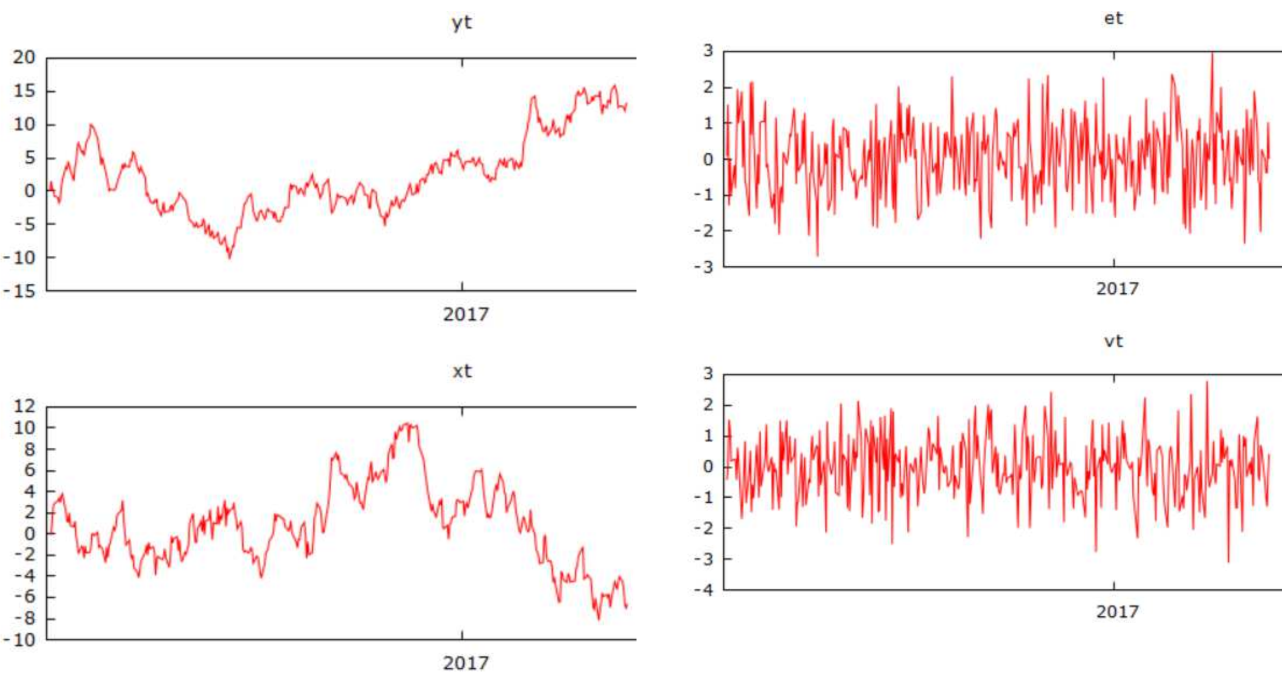
Unabhängige, aber nicht identisch verteilte Innovationen (U_t): Prozesse mit heteroskedastischen Varianzen.

Unabhängigkeitsannahme durch die Unkorreliertheitsannahme ersetzt, was einem White-Noise-Prozess für U_t impliziert.

Spurious-Regressionsproblem

Zwei unabhängige Random Walks

- $x_t = x_{t-1} + u_t$ $u \approx N(0,1)$
- $y_t = y_{t-1} + v_t$ $v \approx N(0,1)$



Spurious-Regressionsproblem

Random walk als Regressoren!

Modell 1: $y_t = \beta_1 + \beta_2 x_t + u_t$

Abhängige Variable: yt

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	2,95576	0,273878	10,79	9,77e-024	***
xt	-0,726998	0,0694755	-10,46	1,43e-022	***

Mittel d. abh. Var.	2,359334	Stdabw. d. abh. Var.	5,830937
Summe d. quad. Res.	9507,922	Stdfehler d. Regress.	5,117873
R-Quadrat	0,231741	Korrigiertes R-Quadrat	0,229625

H_0: $b_2 = 0$

Modell 2: $x_t = \alpha_1 + \alpha_2 y_t + v_t$

Abhängige Variable: xt

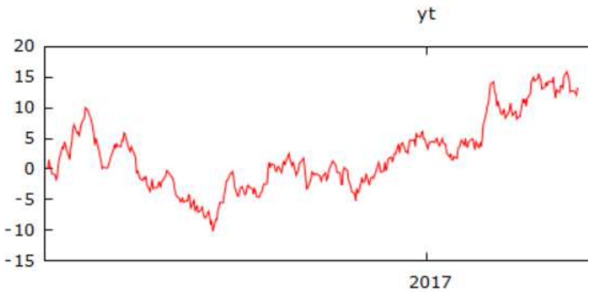
	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	1,57246	0,191390	8,216	3,75e-015	***
yt	-0,318765	0,0304627	-10,46	1,43e-022	***

Mittel d. abh. Var.	0,820392	Stdabw. d. abh. Var.	3,861063
Summe d. quad. Res.	4168,913	Stdfehler d. Regress.	3,388895
R-Quadrat	0,231741	Korrigiertes R-Quadrat	0,229625

H_0: $a_2 = 0$

Die Nullhypothese wird verworfen

Random Walk



Abhängige Variable: yt

Modell 1: $y_t = \beta_1 + \beta_2 time + u_t$

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	-3,63324	0,493399	-7,364	1,21e-012	***
time	0,0327463	0,00233655	14,01	5,70e-036	***

Mittel d. abh. Var.	2,359334	Stdabw. d. abh. Var.	5,830937
Summe d. quad. Res.	8030,640	Stdfehler d. Regress.	4,703507
R-Quadrat	0,351108	Korrigiertes R-Quadrat	0,349321

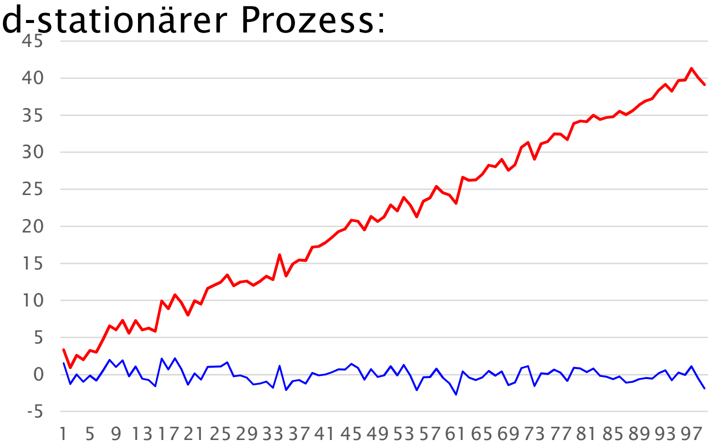
Das Modell ist fehlspezifiziert, da ein deterministischer Trend und eine konstante Varianz unterstellt wird. Random Walk hat keine konstante Varianz!

Problem: Ein stochastischer Trend wurde durch eine deterministische Komponente modelliert.

Eliminieren eines Trends

- Ziel: Y_t in einen stationären Prozess überführen
- Differenz-stationären Prozess = I(1)-Prozess:
- Durch das Bilden von ersten Differenzen Δy_t →stationärer Prozess Ordnung der Integration = 1
- Trend-stationärer Prozess I(2)-Prozess: Durch Subtrahieren eines deterministischen Trends → stationärer Prozess Δy_t

- Y_t ist integriert von der Ordnung d: Δ^d stationärer Prozess: $y_t \approx I(d)$
- Trend-stationärer Prozess:



Residualkomponente und Korrelogramm

Residualkomponente: Restkomponente u_t einer Zeitreihe (y_t) nach Ausschaltung der systematischen Komponenten.

Wenn die Residualkomponente keine Systematik mehr enthält, d.h. nur noch zufällig schwankt, müssen die Autokorrelationen approximativ gleich 0 sein.

Autokorrelation 1. Ordnung: Korrelation der aktuellen Residuen mit den um eine Periode verzögerten Residuen.

Das Ausmass der noch in den Residuen enthaltenen systematischen Schwankungen wird durch den Autokorrelationskoeffizienten gemessen.

Empirische Autokorrelationskoeffizient: Verhältnis der Kovarianz s_{xy} zum Produkt der Standardabweichungen s_x und s_y der beiden Merkmale X und Y $r = s_{xy} / (s_x \cdot s_y)$.

Korrelationskoeffizient

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

Empirischer Autokorrelationskoeffizient

Für eine gegebene Zeitreihe (x_1, x_2, …, x_T)

Autokorrelationskoeffizient 1. Ordnung:
Korrelation zwischen den T-1 Wertepaaren
(x_2, x_1), (x_3, x_2), …, (x_T, x_{T-1}): Distanz = 1

$$r_1 = \frac{\sum_{t=2}^{T}(x_t-\bar{x})(x_{t-1}-\bar{x})}{\sum_{t=1}^{T}(x_t-\bar{x})^2}$$

Autokorrelationskoeffizient 2. Ordnung:
Korrelation zwischen den T-2 Wertepaaren
(x_3, x_1), (x_4, x_2), …, (x_n, x_{T-2}): Distanz = 2

$$r_2 = \frac{\sum_{t=3}^{T}(x_t-\bar{x})(x_{t-2}-\bar{x})}{\sum_{t=1}^{T}(x_t-\bar{x})^2}$$

Autokorrelationskoeffizient k-ter Ordnung:
Korrelation zwischen den T-k Wertepaaren
(x_{k+1}, x_1), (x_{k+2}, x_2), …, (x_T, x_{T-k}): Distanz = k

$$r_k = \frac{\sum_{t=k+1}^{T}(x_t-\bar{x})(x_{t-k}-\bar{x})}{\sum_{t=1}^{T}(x_t-\bar{x})^2}$$

Autokorrelationskoeffizient 1. Ordnung

Beispiel: Autokorrelationskoeffizienten 1. Ordnung für die Zeitreihe des Kfz-Bestandes.

	t	x_t	$x_t - \bar{x}$	x_{t-1}	$x_{t-1} - \bar{x}$	$(x_t - \bar{x})^2$	$(x_t - x)(x_{t-1} - x)$
	1	27'116	-3'033		1	9'201'785	-3'033.4
	2	27'858	-2'291	27'116	-3'033	5'250'718	6'950'969.4
	3	28'452	-1'697	27'858	-2'291	2'881'318	3'889'599.6
	4	29'122	-1'027	28'452	-1'697	1'055'642	1'744'029.9
	5	29'905	-244	29'122	-1'027	59'753	251'153.1
	6	30'618	469	29'905	-244	219'544	-114'535.8
	7	31'748	1'599	30'618	469	2'555'380	749'012.1
	8	32'762	2'613	31'748	1'599	6'825'447	4'176'315.2
	9	33'764	3'615	32'762	2'613	13'065'012	9'443'227.2
Summe	45	271'345				41'114'598	27'086'737

$\bar{x}$ = 30'149

$$r_1 = \frac{\sum_{t=2}^{n}(x_t-\bar{x})(x_{t-1}-\bar{x})}{\sum_{t=1}^{n}(x_t-\bar{x})^2} = \frac{27'086'737}{41'114'598} = 0.659$$

Autokorrelationskoeffizient 2. Ordnung

Beispiel: Autokorrelationskoeffizienten 2. Ordnung für die Zeitreihe des Kfz-Bestands.

t	x_t	$x_t - \bar{x}$	x_{t-2}	$x_{t-2} - \bar{x}$	$(x_t - \bar{x})^2$	$(x_t - \bar{x})(x_{t-2} - \bar{x})$
1	27'116	-3'033			9'201'785	
2	27'858	-2'291			5'250'718	
3	28'452	-1'697	27'116	-3'033	2'881'318	5'149'103.4
4	29'122	-1'027	27'858	-2'291	1'055'642	2'354'331.9
5	29'905	-244	28'452	-1'697	59'753	414'930.9
6	30'618	469	29'122	-1'027	219'544	-481'414.8
7	31'748	1'599	29'905	-244	2'555'380	-390'758.0
8	32'762	2'613	30'618	469	6'825'447	1'224'127.4
9	33'764	3'615	31'748	1'599	13'065'012	5'778'067.9
45	271'345				41'114'598	14'048'389

$\bar{x} = 30'149$

$$r_2 = \frac{\sum_{t=3}^{n}(x_t - \bar{x})(x_{t-2} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} = \frac{14'048'389}{41'114'598} = 0.342$$

Autokorrelationskoeffizient

Autokorrelationskoeffizient 1. Ordnung = Verhältnis der Autokovarianz 1. Ordnung, γ_1, zur Varianz der Zeitreihe der Residuen, γ_0 :

$$r_1 = \gamma_1 / \gamma_0$$

Autokovarianz 1. Ordnung: $\gamma_1 = \frac{1}{T}\sum_{t=2}^{T} u_t \cdot u_{t-1}$

Varianz von u_t: $\gamma_0 = \frac{1}{T}\sum_{t=1}^{T} u_t^2$

Autokorrelationskoeffizienten k-ter Ordnung, (r_k) für k =0,1,2…

$$r_k = \gamma_k / \gamma_0 \qquad \gamma_k = \frac{1}{T}\sum_{t=k+1}^{T} u_t \cdot u_{t-k}$$

Autokorrelation 1. Ordnung

Beispiel: Überprüfung der Residuen des linearen Trendmodells der Zeitreihe des Kfz-Bestandes auf Autokorrelation 1.Ordnung

t	y_t	$\hat{y}_t$	u_t	u_t^2	u_{t-1}	$u_t \cdot u_{t-1}$
1	27116	26856,6	259.4	67288,36	-	-
2	27858	27679,8	178.2	31755,24	259.4	46225,08
3	28452	28503,0	-51.0	2601,00	178.2	-9088,2
4	29122	29326,2	-204.2	41697,64	-51.0	10414,2
5	29905	30149,4	-244.4	59731,36	-204.2	49906,48
6	30618	30972,6	-354.6	125741,16	-244.4	86664,24
7	31748	31795,8	-47.8	2284,84	-354.6	16949,88
8	32762	32619,0	143.0	20449,00	-47.8	-6835,4
9	33764	33442,2	321.8	103555,24	143.0	46017,4
Summe	271‘345	271‘345.0	0.4≈0	455‘103.84		240‘253.68

$$r_1 = \sum_{t=2}^{T} u_t \cdot u_{t-1} / \sum_{t=1}^{T} u_t^2 = \frac{240253.68}{455103.84} = 0.528$$

Autokorrelation 2. Ordnung

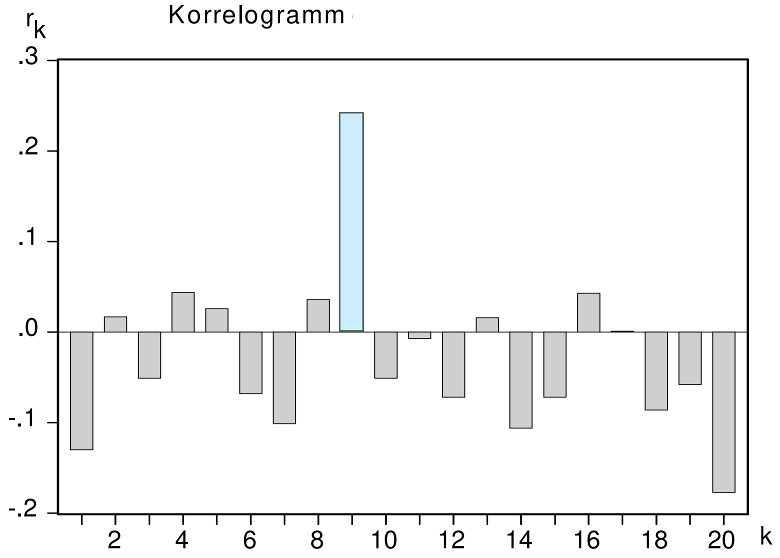
t	y_t	$\hat{y}_t$	u_t	u_t^2	u_{t-2}	$u_t \cdot u_{t-2}$
1	27116	26856,6	259.4	67288,36	-	-
2	27858	27679,8	178.2	31755,24	-	-
3	28452	28503,0	-51.0	2601,00	259.4	-13229,40
4	29122	29326,2	-204.2	41697,64	178.2	-36388,44
5	29905	30149,4	-244.4	59731,36	-51.0	12464,40
6	30618	30972,6	-354.6	125741,16	-204.2	72409,32
7	31748	31795,8	-47.8	2284,84	-244.4	11682,32
8	32762	32619,0	143.0	20449,00	-354.6	-50707,80
9	33764	33442,2	321.8	103555,24	-47.8	-15382,04
Summe	271‘345	271345.0	0.4≈0	455‘103.84		-19‘151.64

$$r_2 = \sum_{t=3}^{T} u_t \cdot u_{t-2} / \sum_{t=1}^{T} u_t^2 = \frac{-19'151.64}{455'103.84} = -0.042$$

3.2 Korrelogramm

Grafische Darstellung der Folge der Autokorrelationskoeffizienten (r_k) in Abhängigkeit von den Lags k in einem Koordinatensystem.

Mit Hilfe eines Korrelogramms lassen sich die Abhängigkeitsstrukturen einer Zeitreihe aufdecken.



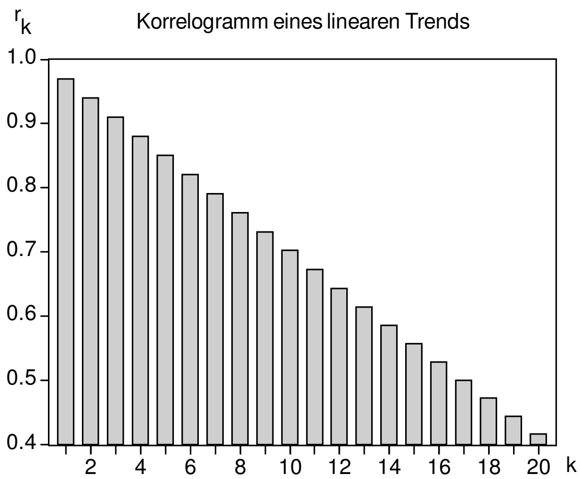
Gauss-Prozess:

Unabhängige normalverteilte Zufallsvariablen

Korrelogramme bei unterschiedlichen Zeitreihenmustern

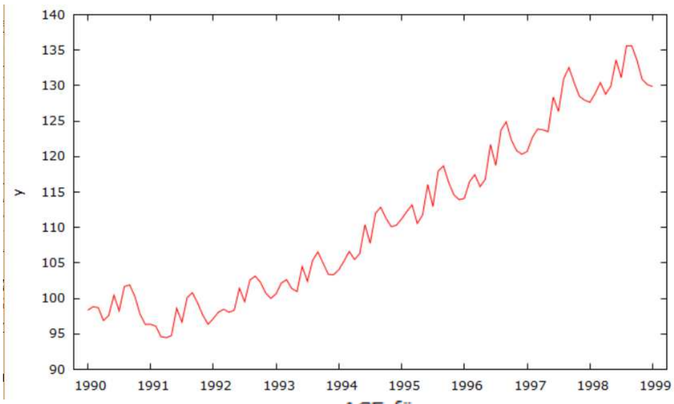
Trendbehaftete Zeitreihen

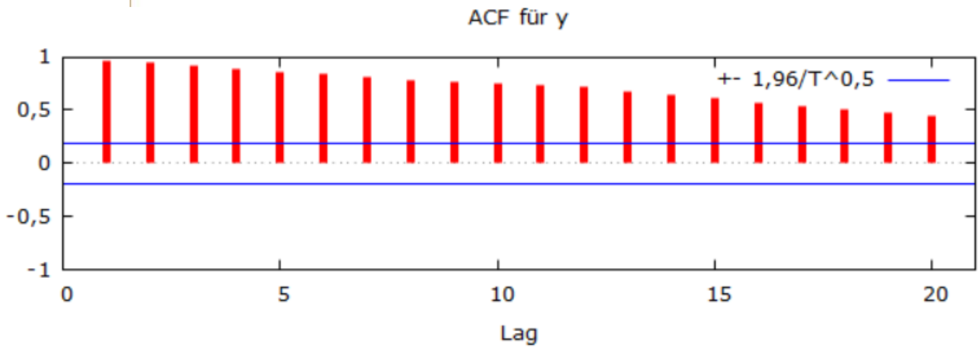
Wenn die Zeitreihe einen Trend enthält, verschwinden die empirischen Autokorrelationen r_k im Allgemeinen erst für sehr grosse Lags. Einer Beobachtung auf einer Seite des Gesamtmittels folgt dann nämlich aufgrund des Trends eine grosse Anzahl von Zeitreihenwerten auf der gleichen Seite. Da der Trend alle anderen Eigenschaften überdeckt, bietet ein Korrelogramm bei trendbehafteten Zeitreihen nur wenig Informationen → Trendbereinigung muss vorgeschaltet werden.



Beispiel

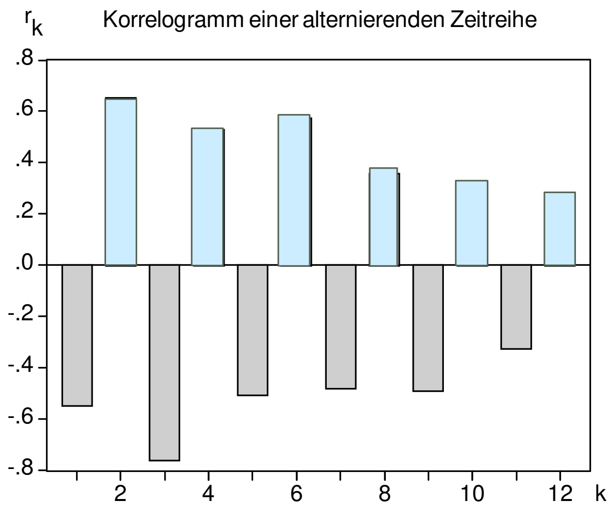
- Anzahl neuer registrierten Autos





Alternierende Zeitreihen

Bei **alternierenden Zeitreihen** liegen aufeinanderfolgende Werte auf verschiedenen Seiten des Gesamtmittelwerts. Bei einem ungeraden Lag ist der Autokorrelationskoeffizient negativ, da die Beobachtungen auf verschiedenen Seiten des Mittelwerts liegen. Entsprechend tendieren die Beobachtungen bei einem geraden Lag auf der gleichen Seite des Mittelwerts zu liegen, so dass der Autokorrelationskoeffizient positiv ist.

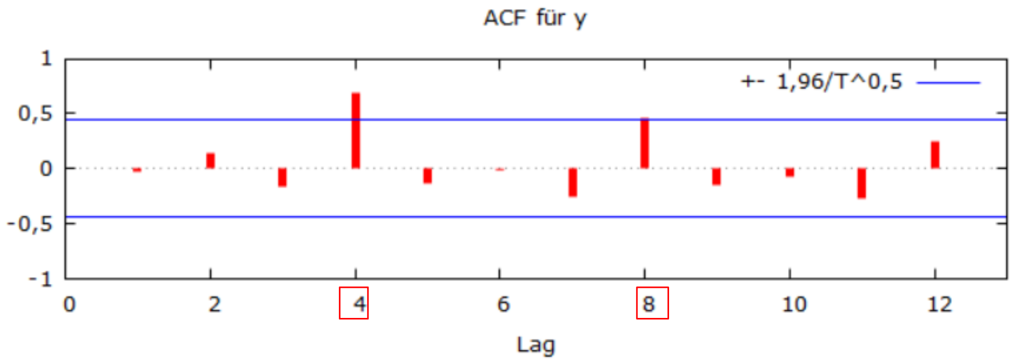


Saisonschwankungen

Saisonschwankungen findet man im Korrelogramm mit der gleichen Frequenz wieder.

Bei Monatsdaten ist z.B. r_6 absolut gross und negativ, während r_{12} gross und positiv ausfällt.

Korrelogramm der Zeitreihe Lohn und Gehälter je Beschäftigten



Das Korrelogramm vermittelt bei saisonalen Daten wenig zusätzliche Informationen, da Saisonschwankungen bereits in dem Zeitreihenplot gut erkennbar sind. Sehr nützliche Informationen liefert dagegen ein Korrelogramm der saisonbereinigten Zeitreihen

Kurzzeit-Korrelationen

Bei nicht-trendbehafteten (oder trendbereinigten) Zeitreihen findet man häufig einen grossen Wert für r_1, dem zwei oder drei signifikant von 0 verschiedene Autokorrelationskoeffizienten folgen. Diese sind aber sukzessive kleiner. Die übrigen Autokorrelationen, d.h. die r_k-Werte für grössere Lags k, liegen dagegen nahe bei 0.

In der zugrunde liegenden Zeitreihe besteht die Tendenz, dass einer überdurchschnittlich (unterdurchschnittlich) grossen Beobachtung einige wenige, d.h. etwa eine, zwei, drei überdurchschnittlich (unterdurchschnittlich) grosse Beobachtungen folgen. Ein solches Verhalten der Zeitreihenwerte wird als **Kurzzeit-Korrelation** bezeichnet.

Nullhypothese: Zufallsprozess

Testen der Nullhypothese H_0: reiner Zufallsprozess der Residualkomponente

$$H_0: \rho_1 = \rho_2 = \ldots = \rho_k = 0 \;\; \rightarrow \text{Populationsparameter}$$

Wenn einer von 20 Autokorrelationskoefizienten ausserhalb des 95%-Konfidenzintervalls liegt → H_0 noch nicht verwerfen, wie es in unserem Beispiel der Fall ist.

Denn dies ist bei einem Signifikanzniveau von 5% durchaus zu erwarten. Hierdurch können Schwierigkeiten bei der Identifikation des datenerzeugenden Prozesses auftreten.

Konfidenzintervall: Beispiel

Es wird überprüft, ob die Residuen des linearen Trendmodells der Zeitreihe des Kfz-Bestands als Realisationen eines reinen Zufallsprozesses aufgefasst werden können.

Approx. 95%-Konfidenzintervall für r_1 und r_2 mit T = 9 (Jahre):

$[-2\sqrt{1/T}, 2\sqrt{1/T}] = [-2\sqrt{1/9}, 2\sqrt{1/9}] = [-0.667; 0.667]$

Autokorrelationskoeffizient 1. Ordnung der Residuen:

$r_1 = 0.528 \in [-0.667; 0.667]$

Autokorrelationskoeffizient 2. Ordnung der Residuen:

$r_2 = -0.042 \in [-0.667; 0.667]$

Da r_1 und r_2 im approx. 95%-Konfidenzintervall liegen, kann die Nullhypothese eines reinen Zufallsprozesses der Residuen des linearen Trendmodells **nicht** abgelehnt werden.

Hinweis: Die Überprüfung hat illustrativen Charakter, da sich eine Ablehnung der Nullhypothese angesichts des geringen Stichprobenumfangs trotz des hohen r_1-Wertes als nicht möglich erweist.

Portmanteau-Tests

Portmanteau-Test: prüft, ob Informationen in den Autokorrelationskoeffizienten bis zu einem Lag k enthaltenen sind.

Box-Pierce-Teststatistik: $Q(k) = T\sum_{j=1}^{k} r_j^2 \quad \to \chi^2$

Bei grossem T approximativ chi-Quadrat mit k-m Freiheitsgraden verteilt.

m = Anzahl der geschätzten Modellparameter und k = Anzahl Lags

Wenn Residuen u_t = reiner Zufallsprozess → r_k niedrige Werte → Prüfgrösse Q niedrig.

Die Nullhypothese eines reinen Zufallsprozesses (White-Noise-Prozess) wird abgelehnt, wenn die Box-Pierce-Statistik Q bei einem Signifikanzniveau α das (1- α)-Quantil der χ^2-Verteilung mit k-m Freiheitsgraden überschreitet:

$$Q(k) > \chi^2_{k-m;1-\alpha} \Rightarrow H_0 \text{ ablehnen}$$

→ noch eine Systematik in den Residuen vorhanden. Das zugrunde gelegte Zeitreihenmodell erfasst nicht alle systematischen Komponenten der Zeitreihe.

Ljung-Box test

Verteilungsapproximation beim Box-Pierce-Test verbessert sich nur langsam mit wachsendem T. Der Test ist extrem konservativ, d.h. die Nullhypothese eines reinen Zufallsprozesses wird im Allgemeinen zu lange beibehalten.

Ljung und Box Test → verbesserte Approximation der Verteilung der Teststatistik an die χ^2-Verteilung. Anstelle der Residualkorrelationen r_k verwenden sie die asymptotisch äquivalenten Grössen $\sqrt{\frac{T+2}{T-j}} \cdot r_j$

Ljung-Box-Teststatistik $Q^*(k) = T(T+2)\sum_{j=1}^{k} \frac{r_j^2}{T-j}$

Q* ist ebenso wie Q approximativ χ^2-verteilt mit k-m Freiheitsgraden. Mit dem Ljung-Box-Test lassen sich unter Umständen inadäquate Zeitreihenmodelle aufdecken, die bei Anwendung des Box-Pierce-Tests möglicherweise übersehen werden.

Beispiel Portmanteau-Test

Portmanteau-Test anhand der Zeitreihe des Kfz-Bestands:

H_0: $r_1 = r_2 = 0$

Box-Pierce-Teststatistik $Q(2) = T\sum_{j=1}^{2} r_j^2 = 9 \cdot (0.659^2 + 0.342^2) = 9 \cdot 0.551245 = 4.961$

Kritischer Wert (α=0,05): $\chi^2_{2;0.95} = 5.99$

Freiheitsgrade: FG = K-m = 2 (m = 0, keine Schätzung)

Testentscheidung: $Q(2) = 4.961 < \chi^2_{2;0.95} = 5.99 \Rightarrow H_0$ annehmen

Ljung-Box-Test

Teststatistik: $Q^*(2) = T(T+2)\sum_{j=1}^{2} \frac{r_j^2}{n-j} = 9 \cdot (9+2) \cdot \left(\frac{0.659^2}{9-1} + \frac{0.342^2}{9-2}\right)$

$= 99 \cdot (0.054285 + 0.016709) = 99 \cdot 0.070994 = 7.028$

Kritischer Wert (α=0.05): $\chi^2_{2;0.95} = 5.99$

Testentscheidung: $Q^*(2) = 7.028 > \chi^2_{2;0.95} = 5.99 \Rightarrow H_0$ ablehnen

Beispiel: Eviews-Output

Korrelogramm und Ljung-Box-Test eines reinen Zufallsprozesses

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
.*| . |	.*| . |	1	-0.130	-0.130	1.7284	0.189
. | . |	. | . |	2	0.017	0.001	1.7601	0.415
. | . |	. | . |	3	-0.051	-0.050	2.0345	0.565
. | . |	. | . |	4	0.044	0.032	2.2420	0.691
. | . |	. | . |	5	0.026	0.037	2.3131	0.804
.*| . |	.*| . |	6	-0.068	-0.064	2.8146	0.832
.*| . |	.*| . |	7	-0.101	-0.117	3.9225	0.789
. | . |	. | . |	8	0.036	0.010	4.0629	0.851
. |** |	. |** |	9	0.241	0.250	10.594	0.305
. | . |	. | . |	10	-0.051	0.008	10.894	0.366
. | . |	. | . |	11	-0.007	-0.013	10.900	0.452
.*| . |	.*| . |	12	-0.072	-0.063	11.502	0.486
. | . |	. | . |	13	0.016	-0.039	11.531	0.566
.*| . |	.*| . |	14	-0.106	-0.135	12.863	0.537
.*| . |	.*| . |	15	-0.072	-0.071	13.490	0.565
. | . |	. |*. |	16	0.043	0.097	13.719	0.620
. | . |	. | . |	17	0.001	0.000	13.719	0.687
.*| . |	.*| . |	18	-0.086	-0.180	14.637	0.687
.*| . |	.*| . |	19	-0.058	-0.102	15.057	0.719
.*| . |	**| . |	20	-0.177	-0.231	19.039	0.519

$Q^*(6) = 2.81$

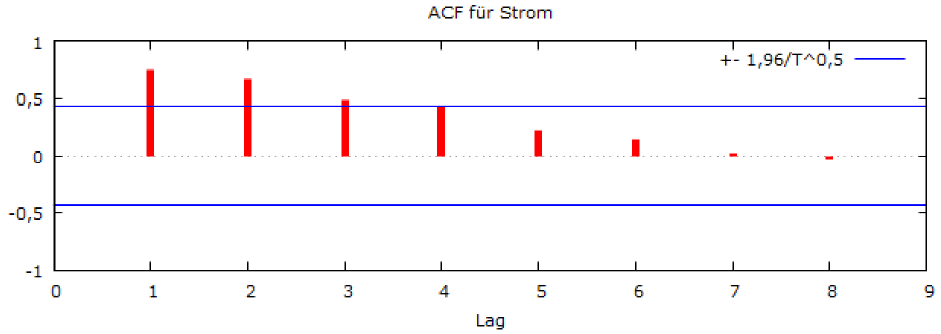
$P(Q6 \geq 2.81) = 83\%$

H_0: $\rho_1 = \ldots = \rho_6 = 0$

→ H_0 nicht ablehen

Wenn p < 5% → H_0 verwerfen

Beispiel: gretl-Output




```
Autokorrelationsfunktion für Strom
***, **, * bezeichnen Signifikanz zum 1%, 5%, 10%-Niveau
mit Standardfehler 1/T^0,5

  LAG      ACF          PACF         Q-Stat. [p-Wert]

    1   0,7509  ***   0,7509 ***    13,0559  [0,000]
    2   0,6710  ***   0,2458        24,0623  [0,000]
    3   0,4823  **   -0,2007        30,0830  [0,000]
    4   0,4318  *     0,1128        35,2106  [0,000]
    5   0,2264       -0,2740        36,7146  [0,000]
    6   0,1412       -0,0423        37,3415  [0,000]
    7   0,0137       -0,0063        37,3479  [0,000]
    8  -0,0346       -0,0478        37,3918  [0,000]
```