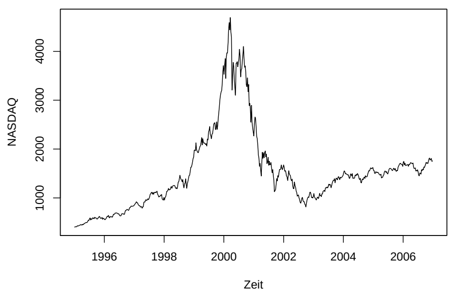
Kapitel 12

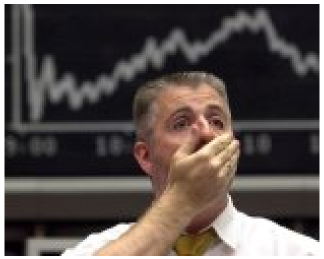
Stochastische Prozesse und Zeitreihenmodelle

Einleitung: .com-Blase an der NASDAQ

.com-Blase an der NASDAQ

Modellierung und Prognose von Aktienrenditen





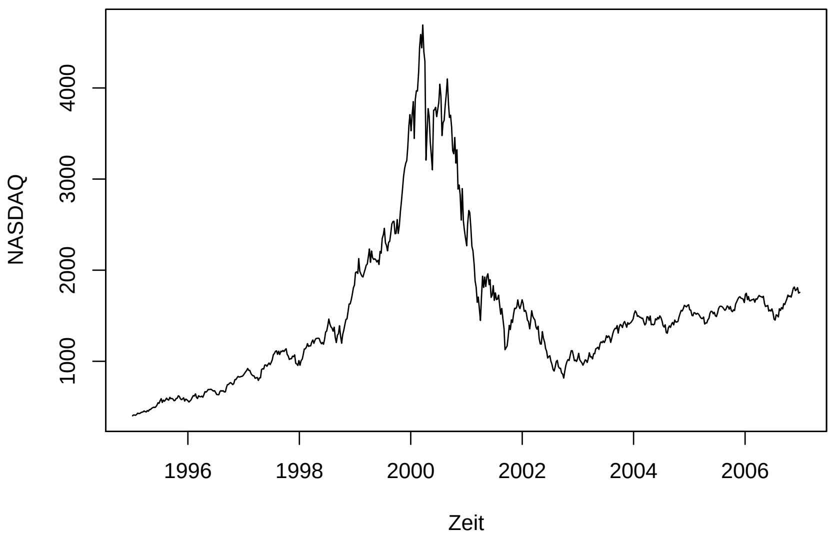
Daten

An der New Yorker Börse NASDAQ (National Association of Securities Dealers Automated Quotation System) sind die Hälfte der US-amerikanischen Aktiengesellschaften gelistet, u.a.: Adobe Systems, Amazon.com, Apple, Cisco, Dell, eBay, Google, Intel, Microsoft, Sun Microsystems, Yahoo!. Sie profitierte in der zweiten Hälfte der 90er Jahre besonders vom Boom der sogenannten .com-Blase, deren Platzen im März 2000 zu einem 75%-igen Wertverlust der zuletzt gekauften Aktien führte. Hier untersuchen wir die wöchentliche Zeitreihe des NASDAQ-100 Index (1995–2006).

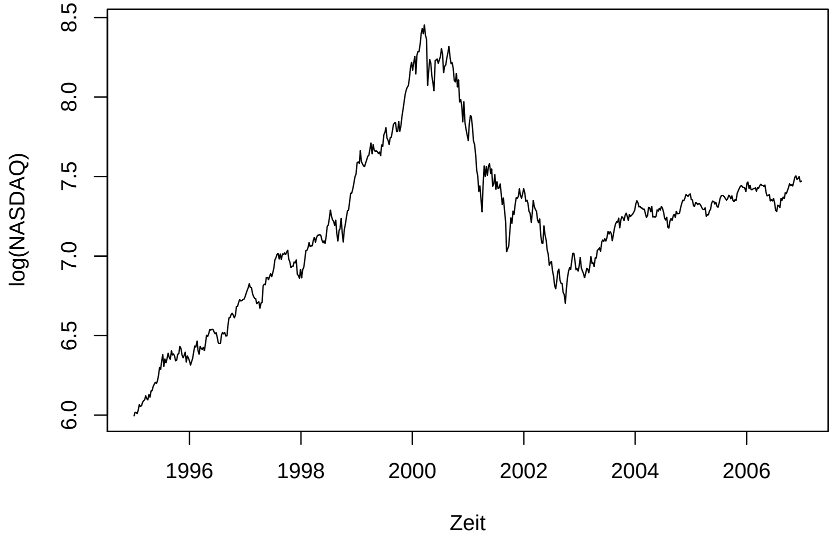
Frage: Wie kann man die Renditen des NASDAQ-Index bis Anfang 1998 (vor dem Platzen der .com-Blase) modellieren? Wie können damit zukünftige Renditen bsw. Kurse prognostiziert werden?

Quelle: Yahoo! Finance, `http://finance.yahoo.com/` (Instrument `^NDX`).

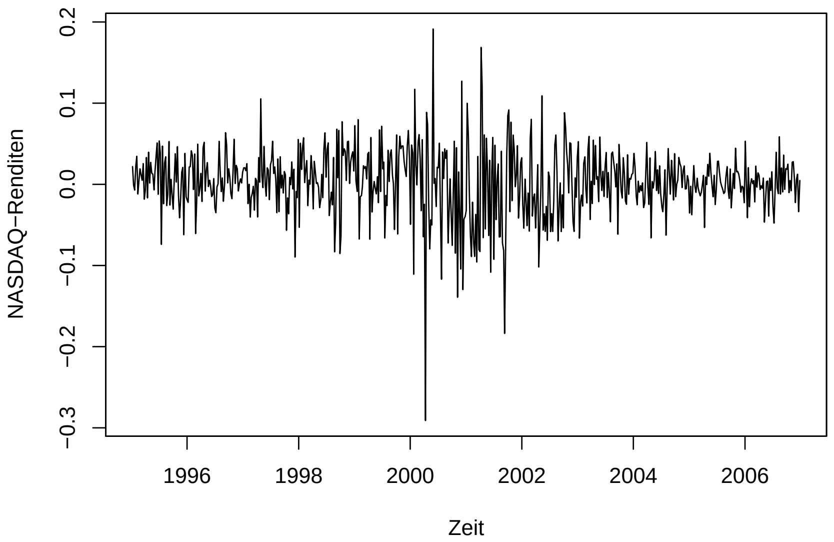
Daten: Preise



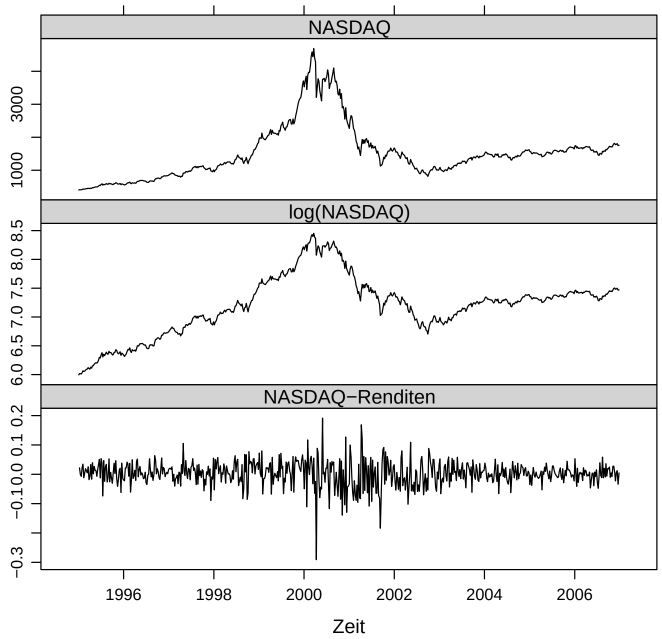
Daten: Log-Preise



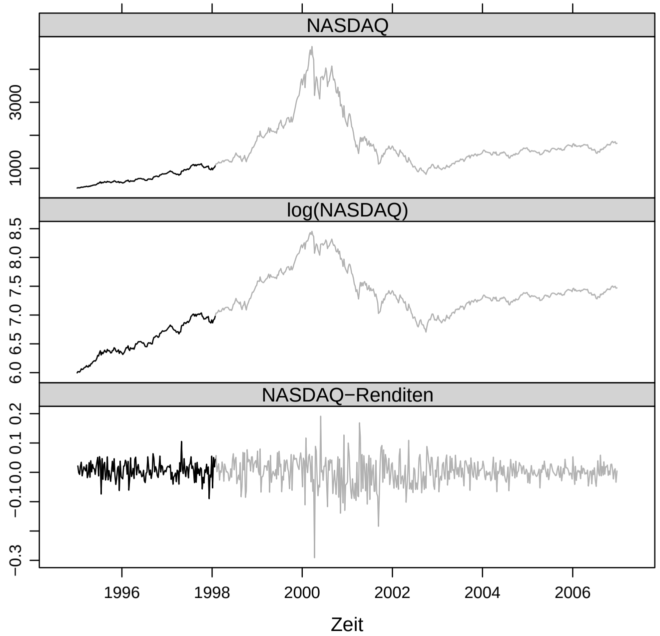
Daten: Renditen



Daten: Gesamtzeitraum



Daten: Modellierungs- und Prognosezeitraum



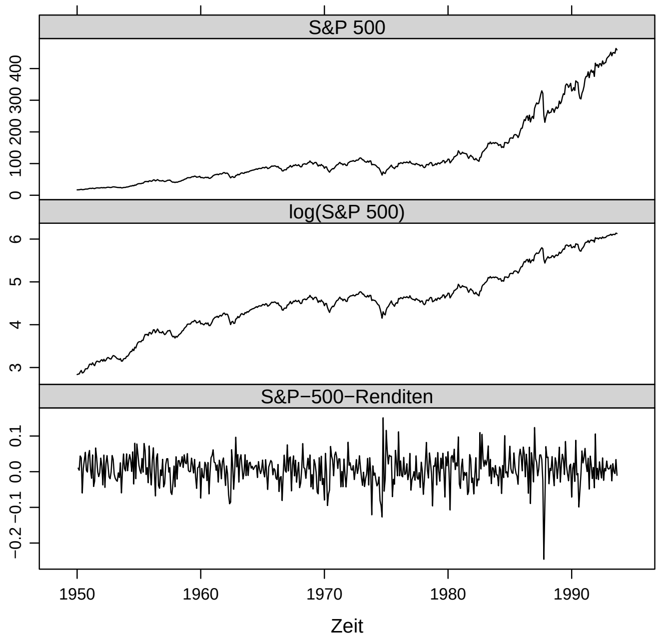
Standard & Poor's 500 Index

Weiterer Datensatz: Der Standard & Poor's 500 (kurz: S&P 500) ist ein Index der 500 größten Aktiengesellschaften, die an den beiden New Yorker Börsen NYSE (New York Stock Exchange) und NASDAQ gehandelt werden.

Über den gewählten Zeitraum 1950(1) bis 1993(9) zeigt die S&P 500 Reihe ein exponentielles Wachstum. Dieses basiert wie die Entwicklung eines Guthabens eines Sparbuches auf dem Zinseszinseffekt.

Mittels Logarithmus transformiert man das multiplikative Bildungsgesetz in ein additives. Der transformierte Pfad wird dadurch ungefähr linear.

S&P 500 Index



[Stochastische Prozesse und Zeitreihenmodelle]

Preise und Renditen

Aktienkurs und Rendite

Wir bezeichnen einen beobachteten Aktienkurs mit P_t und den logarithmierten mit $p_t = \log(P_t)$. Aus der Mathematik Vorlesung ist bekannt, dass die Rendite (in stetiger Zeit), r_t, aus dem Logarithmus des Quotienten berechnet wird, $r_t = \log(P_t/P_{t-1})$. Damit ist r_t die Differenz der logarithmierten Kurse aufeinander folgender Zeitpunkte:

$$r_t = \log(P_t) - \log(P_{t-1}) = p_t - p_{t-1}$$

Umgekehrt ergibt sich für den aktuellen logarithmierten Kurs

$$p_t = p_{t-1} + r_t.$$

Der aktuelle logarithmierte Kurs ist der logarithmierte Kurs der Vorperiode plus der aktuellen Rendite. Das Modell für den Kurs P_t selbst ist

$$P_t = P_{t-1} \cdot \exp(r_t).$$

Hypothese effizienter Märkte

Die Hypothese effizienter Märkte besagt, dass Finanzmärkte informationseffizient sind. Marktpreise von Aktien enthalten bereits die gesamte verfügbare Information über zukünftige Gewinnaussichten. Formal bedeutet das:

"Die beste Prognose für den morgigen Aktienkurs ist der heutige Kurs plus einer durchschnittlichen Rendite."

Damit ist die 1-Schritt Prognose, $\hat{p}_{t+1}$, gleich dem zuletzt beobachteten Wert plus einer Konstanten, der durchschnittlichen Rendite β_0:

$$\hat{p}_{t+1} = p_t + \beta_0.$$

Die Renditen sind nicht prognostizierbar.

Hypothese effizienter Märkte

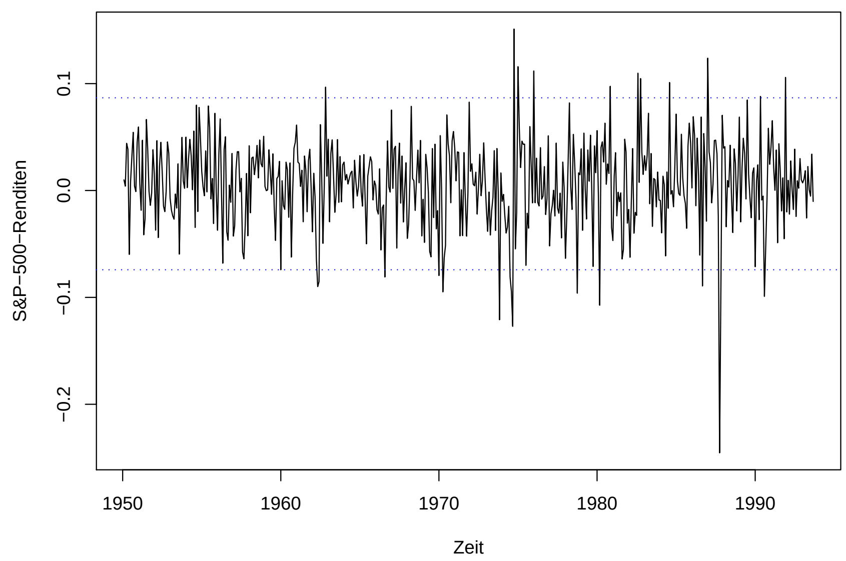
Renditen r_t sind nicht prognostizierbar, wenn die Vergangenheit der Reihe $r_{t-1}, r_{t-2}, \ldots, r_1$, (oder weitere Erklärungsvariablen, die wir aber hier nicht weiter diskutieren) nicht zur Verbesserung der Prognose verwendet werden können. Das einzig sinnvolle Regressionsmodell für die Renditen lautet dann:

$$r_t = \beta_0 + \epsilon_t.$$

Anmerkung: Dieses Modell ist bereits aus früheren Kapiteln bekannt.

β_0 wird als die durchschnittliche Rendite geschätzt (Kapitel 4) bzw. völlig äquivalent als Konstante in einem Regressionsmodell $r_t \sim 1$ (Kapitel 10).

Renditen des S&P 500, Monatsdaten



Renditen des S&P 500

Zusammen mit der Renditenreihe haben wir ein 95% Intervall um das Mittel eingezeichnet. Die Renditenreihe verläuft in etwa horizontal. Das Mittel ist leicht positiv, umgeben von einer großen Streuung.

Die Zeitreihe hat die Länge $n = 524$, das Mittel $\bar{r} = 0.00628$ und die Standardabweichung $s_{r,n-1} = 0.04104$. Wir können testen, ob die erwartete Rendite von 0 abweicht (Kapitel 4): $H_0 : \mu_r = 0$, $H_A : \mu_r \neq 0$.

$$T = \frac{\bar{r} - 0}{\sqrt{s^2_{r,n-1}/n}} = \frac{0.0063}{\sqrt{0.041^2/524}} = 3.5051 > 1.96$$

Die Nullhypothese wird klar zugunsten von $\mu_r > 0$ verworfen.

Die durchschnittliche Monatsrendite für die Periode 1950(1) bis 1993(9) beträgt 0.00628 (0.628%). Das ergibt eine durchschnittliche Jahresrendite von 7.54% ($= 0.00628 \cdot 100 \cdot 12$).

Beispiel: T-Test

Gegeben sind die Werte eines Aktienindex, P_t, für die Periode 1978-03-13 bis 1978-05-28, tägliche Schlusskurse.
Daraus ergibt sich für die Renditenreihe das Mittel $\bar{r} = 0.00057$, die Standardabweichung $s_{r,n-1} = 0.01669$ und die Länge $n = 77$.

Prüfen Sie die Hypothese, dass die erwarteten Renditen gleich 0 sind.

Beispiel: T-Test

Gegeben sind die Werte eines Aktienindex, P_t, für die Periode 1978-03-13 bis 1978-05-28, tägliche Schlusskurse.
Daraus ergibt sich für die Renditenreihe das Mittel $\bar{r} = 0.00057$, die Standardabweichung $s_{r,n-1} = 0.01669$ und die Länge $n = 77$.

Prüfen Sie die Hypothese, dass die erwarteten Renditen gleich 0 sind.

Für die Teststatistik ergibt sich

$$T = \frac{\bar{r} - 0}{\sqrt{s^2_{n-1}/n}} = \frac{0.00057}{\sqrt{0.01669^2/77}} = 0.300.$$

Beispiel: T-Test

Welche der Aussagen sind richtig? (Signifikanzniveau 5%)

(a) Eine geschätzte Renditengleichung der Form $r_t = \beta_0 + \epsilon_t$ würde eine signifikante Konstante liefern.

(b) Der Absolutbetrag der Teststatistik ist größer als 1.33.

(c) Der Test weist nach, dass die erwarteten Renditen ungleich 0 sind.

(d) Die Teststatistik ist größer als 0.13.

(e) Die Nullhypothese wird beibehalten.

Beispiel: T-Test

Welche der Aussagen sind richtig? (Signifikanzniveau 5%)

(a) Eine geschätzte Renditengleichung der Form $r_t = \beta_0 + \epsilon_t$ würde eine signifikante Konstante liefern. falsch

(b) Der Absolutbetrag der Teststatistik ist größer als 1.33. falsch

(c) Der Test weist nach, dass die erwarteten Renditen ungleich 0 sind. falsch

(d) Die Teststatistik ist größer als 0.13. richtig

(e) Die Nullhypothese wird beibehalten. richtig

Prognostizierbarkeit

Renditen sind r_t prognostizierbar, wenn die Vergangenheit der Reihe zur Verbesserung der Prognose verwendet werden kann, d.h. wenn man das Mittel von r_t aus verzögerten Werten (**lags**) $r_{t-1}, r_{t-2}, \ldots$ vorhersagen kann.

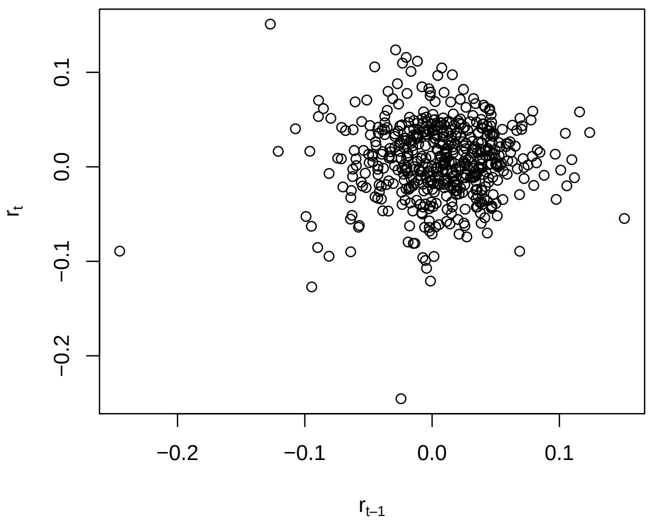
Im einfachsten Fall verwendet man zur Vorhersage des Werts zum Zeitpunkt t den vorigen Wert der Reihe (also Zeitpunkt $t-1$):

$$r_t = \beta_0 + \beta_1 \cdot r_{t-1} + \epsilon_t.$$

Dieses Modell nennt man **autoregressives Modell erster Ordnung**, kurz **AR(1)**.

Anmerkung: Abhängigkeiten können auch komplexer sein, bspw. verzögerte Werte, die weiter in der Vergangeheit liegen (autoregressive Modelle höherer Ordnung) oder andere Variablen führen zu einer verbesserten Prognose.

S&P 500: r_t vs. r_{t-1}



Die Korrelation zwischen r_t und r_{t-1} beträgt 0.027.

Autokorrelationskoeffizient, Korrelogramm

Die Korrelationskoeffizienten

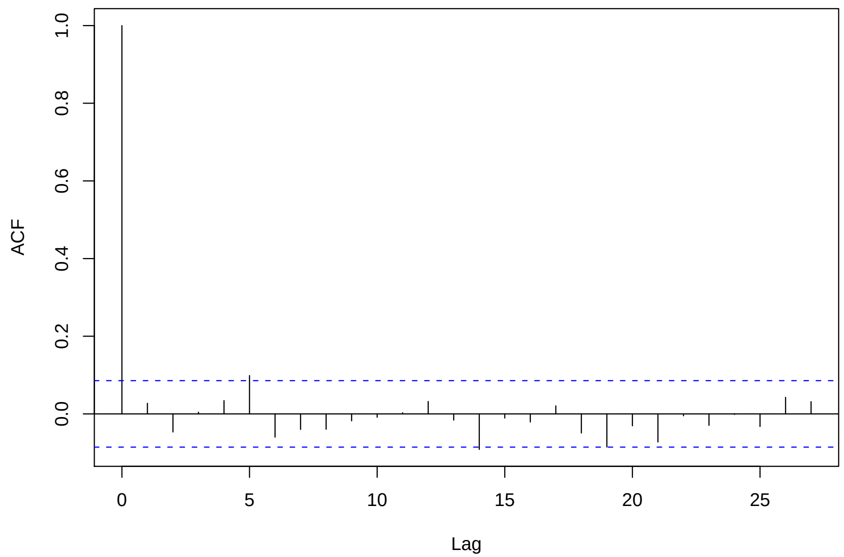
$$\rho_k = Cor(r_t, r_{t-k}), \quad k = 0, 1, 2, \ldots$$

heißen **Autokorrelationskoeffizienten k-ter Ordnung**.

Der erste Wert des Korrelogramms ρ_0 ist immer 1, da $Cor(r_t, r_t) = 1$ ist (siehe Kapitel 9).

Alle empirischen Autokorrelationskoeffizienten $\hat{\rho}_k$ zusammen nennt man **Korrelogramm**: $\hat{\rho}_0, \hat{\rho}_1, \hat{\rho}_2, \ldots$.

S&P-500-Renditen: Korrelogramm



S&P-500-Renditen: Korrelogramm

Das Korrelogramm zeigt hier die empirischen Autokorrelationskoeffizienten $\hat{\rho}_0, \hat{\rho}_1, \ldots, \hat{\rho}_{27}$. Parallel zur x-Achse befinden sich außerdem die beiden Grenzen der punktweisen 95%-Intervalle für die Nullhypothesen $\rho_k = 0$.

Auch wenn alle Nullhypothesen zutreffen, liegen im Durchschnitt 5% der empirischen Autokorrelationen $\hat{\rho}_k$ außerhalb der Intervalle. D.h. wenn alle $\hat{\rho}_k$ innerhalb der Intervalle liegen, sind alle Autokorrelationen nahe bei 0. Wenn einige wenige die Grenzen etwas überschreiten, spricht das noch nicht unbedingt dagegen, dass die Autokorrelationen 0 sind.

Das Korrelogramm der S&P-500-Renditen zeigt nur geringfügige Abweichungen von $\rho_k = 0$ für $k = 1, 2, \ldots, 27$. (Bei $k = 0$ ist der Wert per Definition eins.)

S&P-500-Renditen: Regression

Zur Modellprüfung mit Hilfe von statistischen Tests verwenden wir die Modellselektionsmethoden aus Kapitel 10:
Will man überprüfen, ob man zur Erklärung der Renditen des S&P 500 verzögerte (lagged) Renditen der Ordnung eins verwenden soll, kann man die T-Größe verwenden. Wir schätzen das autoregressive Modell erster Ordnung $r_t \sim r_{t-1}$:

$$r_t = \beta_0 + \beta_1 \cdot r_{t-1} + \epsilon_t$$

```
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      0.00610    0.00182    3.36  0.00085
lag(rsp500, -1)  0.02732    0.04380    0.62  0.53306
```

Für den Test $H_0 : \beta_1 = 0$, $H_A : \beta_1 \neq 0$ erhalten wir einen p-Wert von 0.533, so dass wir die Nullhypothese beibehalten.

S&P-500-Renditen: Regression

Damit vereinfachen wir das Modell zu $r_t \sim 1$:

$$r_t = \beta_0 + \epsilon_t.$$

Dies Modell ist mit der Hypothese effizienter Märkte verträglich.

Schätzung des Mittelwerts ergibt (wie bereits gezeigt):

$$r_t = 0.00628 + \epsilon_t,$$

Die Regressionszusammenfassung ist:

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.00628    0.00179    3.51   0.0005
```

Beispiel: Regression

Gegeben sind die Werte für einen Aktienindex, P_t, für die Periode 1998-11-08 bis 1999-01-17, tägliche Schlusskurse. Es werden die Renditen, r_t, berechnet.
Wir passen an die Renditen ein AR(1) Modell an:

$$r_t = \beta_0 + \beta_1 \cdot r_{t-1} + \epsilon_t.$$

Die Regressionszusammenfassung ist:

```
Model: r ~ lag(r, -1)

             Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.003      0.002   1.931    0.058
lag(r, -1)    -0.012      0.122  -0.099    0.921
```

Beurteilen Sie das geschätzte Modell.

Beispiel: Regression

Welche der Aussagen sind richtig? (Signifikanzniveau 5%)

(a) Es kann nachgewiesen werden, dass β_1 ungleich 0 ist.

(b) Die Konstante ist nicht signifikant von 0 verschieden.

(c) Vergangene Renditewerte tragen nicht signifikant zur Verbesserung der Prognose von r_t bei.

(d) Das Vorzeichen des ersten Autokorrelationskoeffizienten und das von β_1 ist immer gleich.

(e) Die Hypothese, dass der Prozess ein AR(0) ist, kann verworfen werden.

Beispiel: Regression

Welche der Aussagen sind richtig? (Signifikanzniveau 5%)

(a) Es kann nachgewiesen werden, dass β_1 ungleich 0 ist. falsch

(b) Die Konstante ist nicht signifikant von 0 verschieden. richtig

(c) Vergangene Renditewerte tragen nicht signifikant zur Verbesserung der Prognose von r_t bei. richtig

(d) Das Vorzeichen des ersten Autokorrelationskoeffizienten und das von β_1 ist immer gleich. richtig

(e) Die Hypothese, dass der Prozess ein AR(0) ist, kann verworfen werden. falsch
AR(0) meint $r_t = \beta_0 + \epsilon_t$ (siehe folgender Abschnitt).

[Stochastische Prozesse und Zeitreihenmodelle]

Stochastische Prozesse

Stochastische Prozesse: Autoregressive Modelle

Ein **stochastischer Prozess** Y_t ist eine Sequenz von Zufallsvariablen und damit gleichsam das theoretische Gegenstück zu einer empirischen Zeitreihe. Somit können stochastische Prozesse als Modelle für Zeitreihen verwendet werden.

Eine besonders wichtige Klasse solcher Modelle sind autoregressive Modelle. Sie erklären einen Prozess durch die eigene Vergangenheit:

$$Y_t = \beta_0 + \beta_1 \cdot Y_{t-1} + \ldots + \beta_p \cdot Y_{t-p} + \epsilon_t$$

Dies ist ein **autoregressives Modell der Ordnung** p, kurz **AR(**p**)**.

Beispiele: Für $Y_t = r_t$ haben wir solche Modelle bereits verwendet.

AR(0): $Y_t = \beta_0 + \epsilon_t$

AR(1): $Y_t = \beta_0 + \beta_1 \cdot Y_{t-1} + \epsilon_t$

AR(2): $Y_t = \beta_0 + \beta_1 \cdot Y_{t-1} + \beta_2 \cdot Y_{t-2} + \epsilon_t$

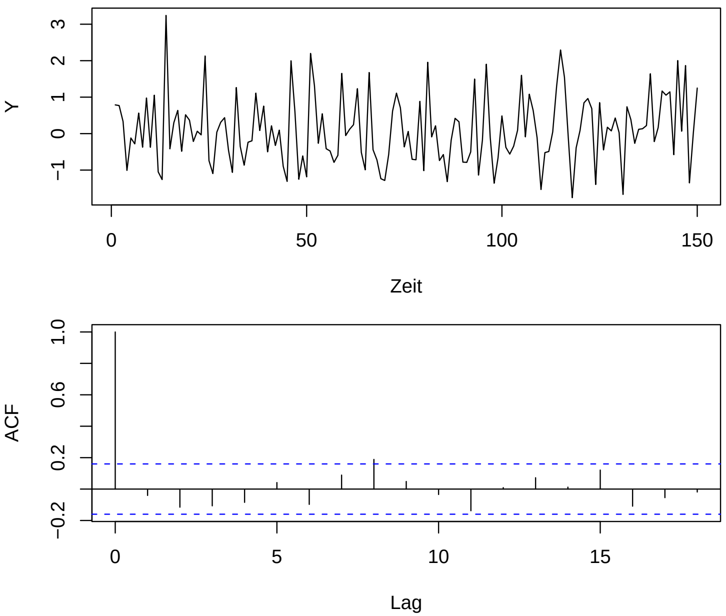
Weißes Rauschen

Die Störgröße ϵ_t in diesen Modellen sollte ein stochastischer Prozess sein mit

- Mittel 0: $E(\epsilon_t) = 0$,
- konstante Varianz: $V(\epsilon_t) = \sigma_\epsilon^2$,
- keine Autokorrelation: $\rho_k = 0$ für $k = 1, 2, \ldots$.

Einen solchen Prozess nennt man **weißes Rauschen** (**white noise**).

Weißes Rauschen



Stationäre Prozesse

Weißes Rauschen ist ein Spezialfall eines allgemeineren Typs stochastischer Prozesse, nämlich **stationärer Prozesse**. Diese sind charakterisiert durch

- konstantes Mittel,
- konstante Varianz,
- Autokorrelationen, die über die Zeit stabil bleiben, d.h. nur von der zeitlichen Distanz k der Beobachtungen abhängen.

Daher sind auch AR(0) Prozesse stationär:

$$Y_t = \beta_0 + \epsilon_t,$$

denn sie sind ein Mittelwert β_0 plus ein weißes Rauschen ϵ_t. Sie haben also gleiche Varianz und gleiche Autokorrelationen wie weißes Rauschen, nur einen anderen (aber konstanten) Mittelwert.

Stationäre Prozesse

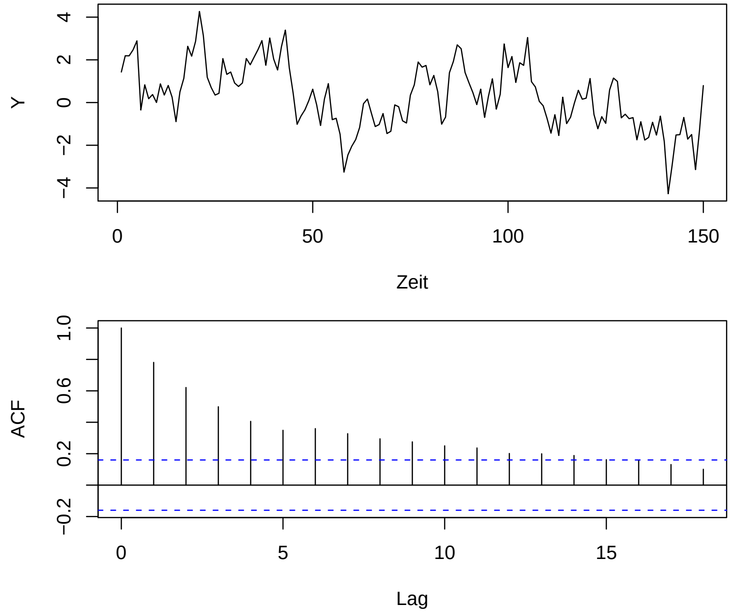
AR(1) Prozesse sind für $|\beta_1| < 1$ stationär.

$$Y_t = \beta_0 + \beta_1 \cdot Y_{t-1} + \epsilon_t$$

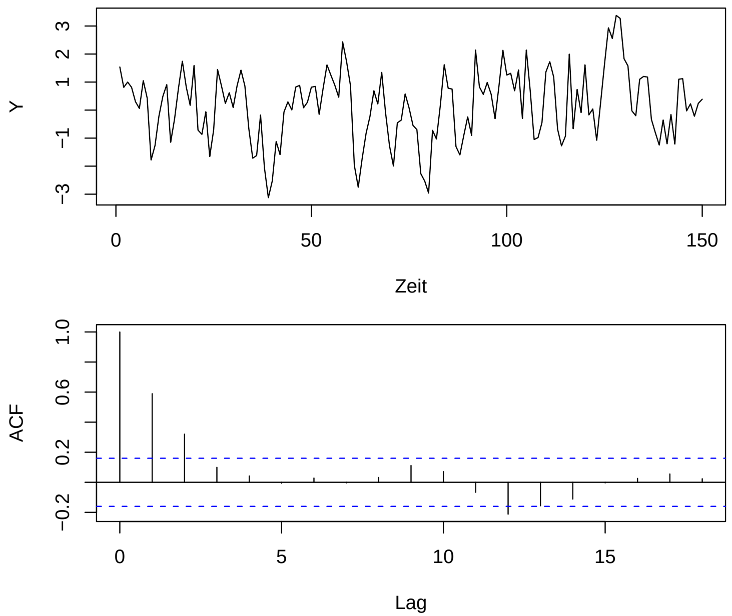
Man kann zeigen, dass der Erwartungswert und die Varianz konstant sind. Die zugehörigen Korrelogramme klingen geometrisch wie β_1^k ab.

- Ist β_1 positiv, sind alle Autokorrelationskoeffizienten positiv.
- Ist β_1 negativ, so haben die Autokorrelationskoeffizienten alternierende Vorzeichen.

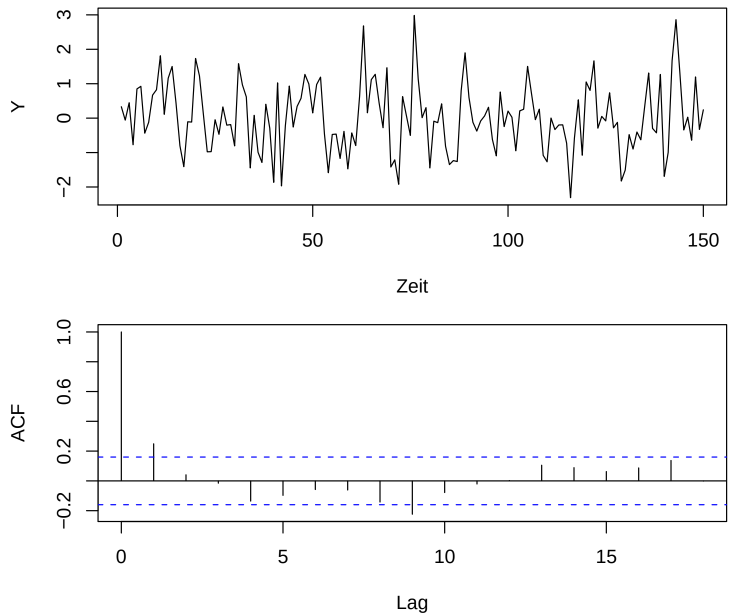
AR(1) Prozess mit $\beta_1 = 0.8$



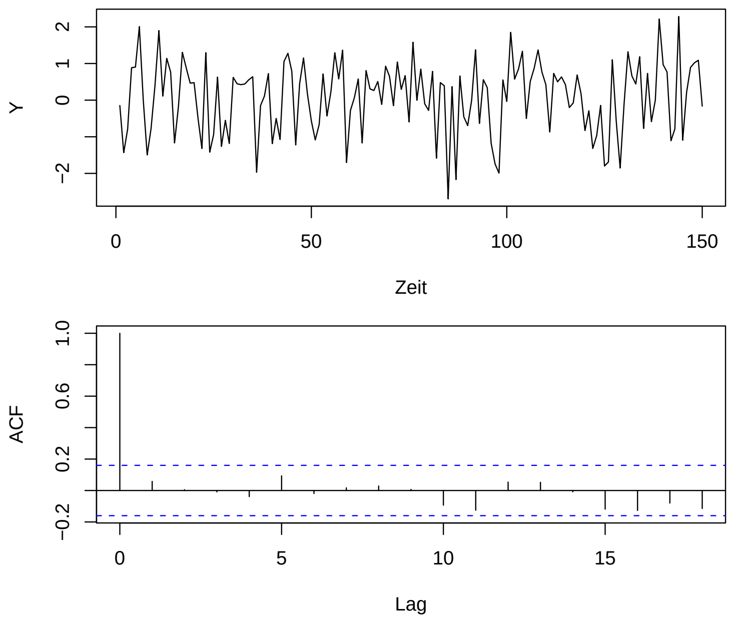
AR(1) Prozess mit $\beta_1 = 0.6$



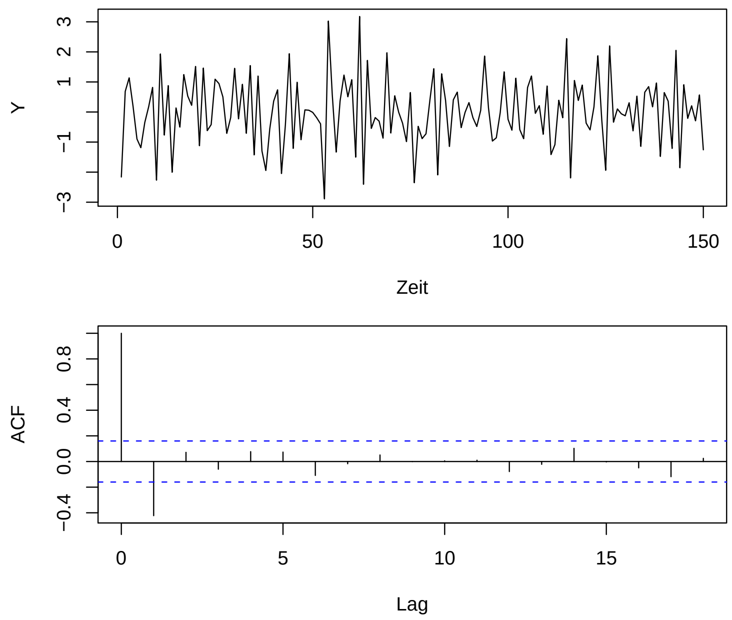
AR(1) Prozess mit $\beta_1 = 0.4$



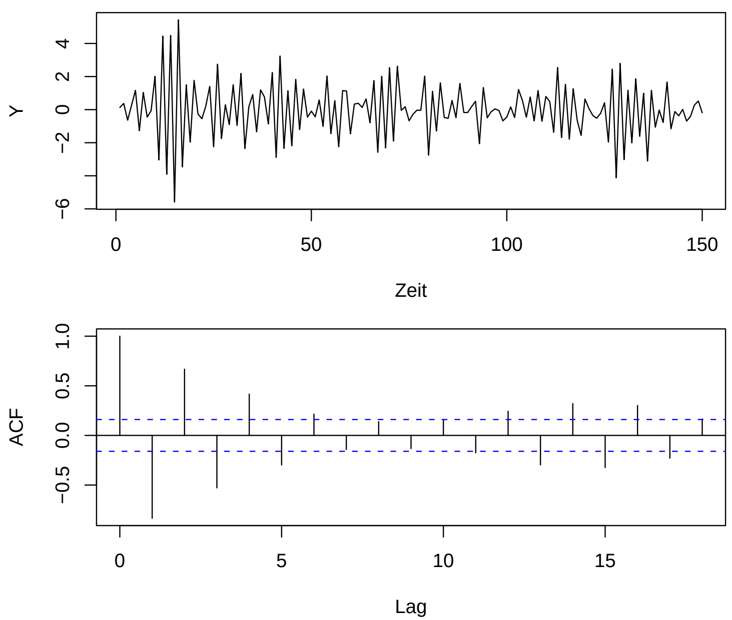
AR(1) Prozess mit $\beta_1 = 0$ (= AR(0))



AR(1) Prozess mit $\beta_1 = -0.4$



AR(1) Prozess mit $\beta_1 = -0.8$



Random Walk

AR(1) Prozesse sind nur für $|\beta_1| < 1$ stationär. Für $\beta_1 = 1$ hingegen erhält man einen nicht-stationären Prozess, eine sogenannte **Zufallsbewegung** oder **Random Walk**:

$$Y_t = \beta_0 + Y_{t-1} + \epsilon_t.$$

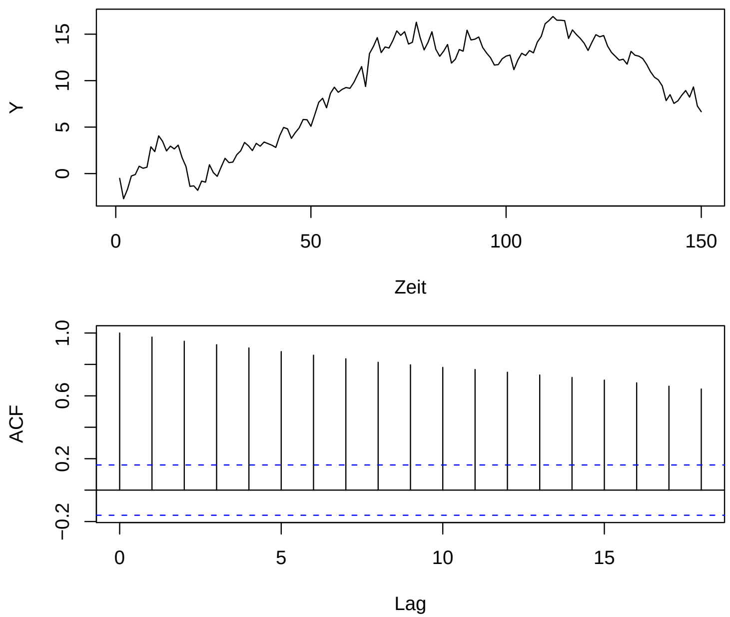
Ist $\beta_0 = 0$ ist es ein Random Walk **ohne Drift**, für $\beta_0 \neq 0$ **mit Drift**.

Dieses Modell haben wir bereits für die logarithmierten Aktienpreise $Y_t = p_t = \log(P_t)$ unter der Effizienzhypothese kennengelernt.

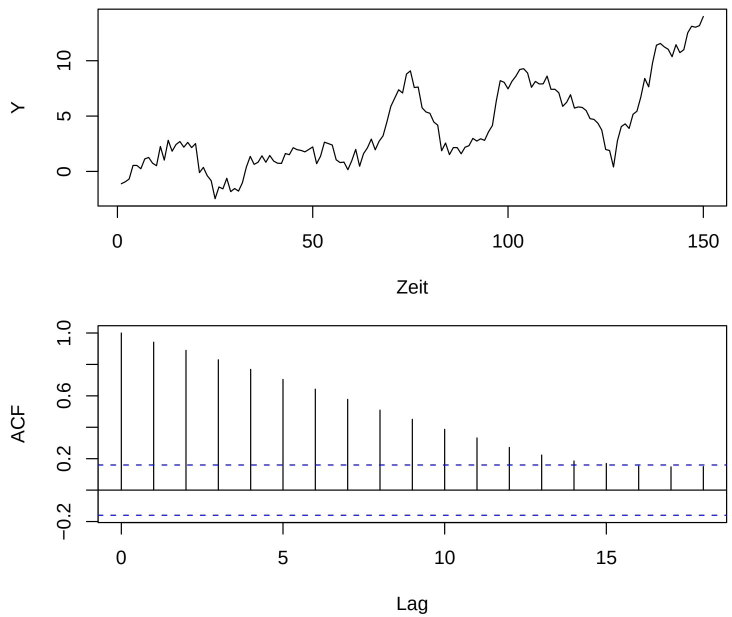
Ein Random Walk ist nicht stationär, weil die Varianz mit der Zeit steigt. Das 95%-Intervall, in dem sich der Pfad bewegt, liegt ausgehend vom Startwert symmetrisch um den Drift und verbreitert sich mit $\sqrt{t}$. Bei einem Random Walk mit Drift steigt zusätzlich das Mittel mit der Zeit, ist also auch nicht konstant.

Das Korrelogramm eines Random Walk geht mit steigendem k nur langsam (linear) gegen 0.

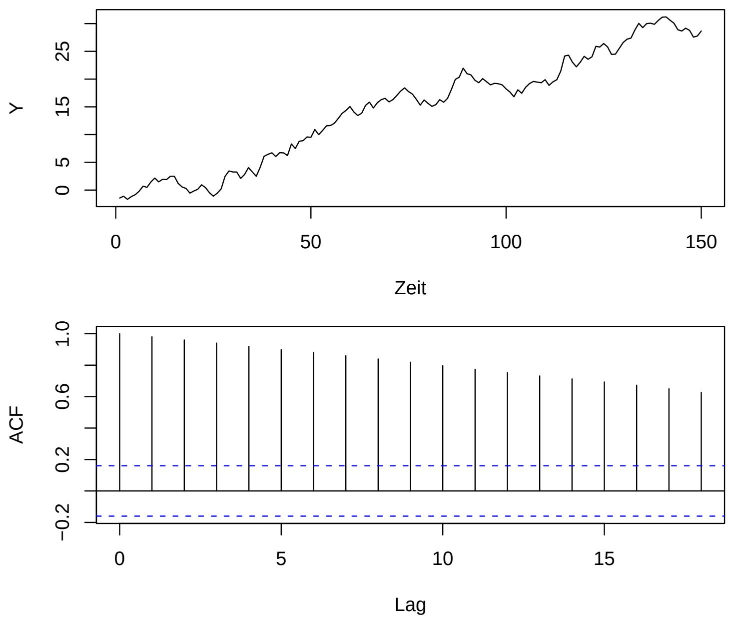
Random Walk ohne Drift $\beta_0 = 0$



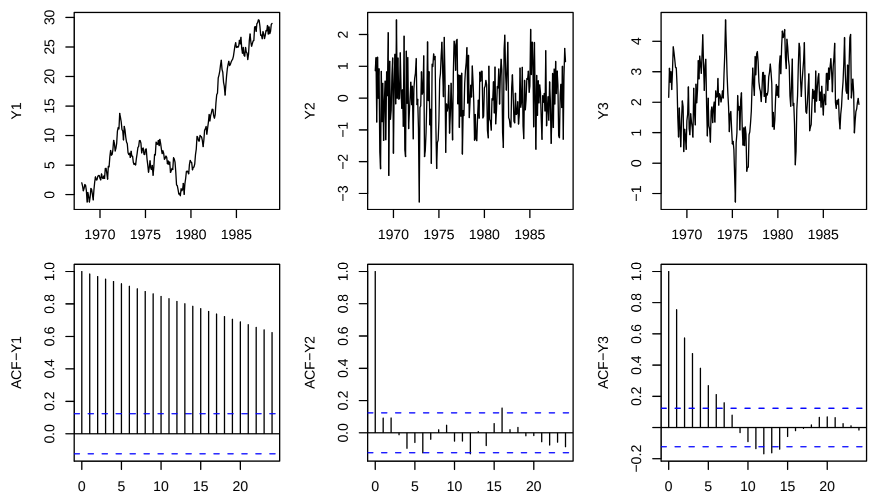
Random Walk ohne Drift $\beta_0 = 0$



Random Walk mit Drift $\beta_0 = 0.2$



Beispiel: Stochastische Prozesse



Beispiel: Stochastische Prozesse

Es wurden je ein Pfad eines weißen Rauschens, eines AR(1) und eines Random Walks generiert. Ordnen Sie jedem Graphen das richtige Modell zu.
Welche der folgenden Aussagen sind richtig?

(a) Y1 ist weißes Rauschen.
(b) Y2 gehorcht einem AR(1).
(c) Y3 zeigt das Verhalten eines Random Walks.
(d) Y1 ist weißes Rauschen, weil die meisten Autokorrelationskoeffizienten nahe bei 0 liegen.
(e) Der Koeffizient β_1 des AR(1) Prozess ist positiv.

Beispiel: Stochastische Prozesse

Es wurden je ein Pfad eines weißen Rauschens, eines AR(1) und eines Random Walks generiert. Ordnen Sie jedem Graphen das richtige Modell zu.
Welche der folgenden Aussagen sind richtig?

(a) Y1 ist weißes Rauschen. falsch
(b) Y2 gehorcht einem AR(1). falsch
(c) Y3 zeigt das Verhalten eines Random Walks. falsch
(d) Y1 ist weißes Rauschen, weil die meisten Autokorrelationskoeffizienten nahe bei 0 liegen. falsch
(e) Der Koeffizient β_1 des AR(1) Prozess ist positiv. richtig

Prognose eines AR(1)

Gegeben sind $Y_1, Y_2, \ldots, Y_{n-1}, Y_n$. Gesucht sind die

- 1-Schritt Prognose: $\hat{Y}_{n+1}$

$$Y_{n+1} = \beta_0 + \beta_1 \cdot Y_n + \epsilon_{n+1}$$

Ersetzt man den unbekannten Fehler ϵ_{n+1} durch seinen Erwartungswert $E(\epsilon_{n+1}) = 0$, erhält man

$$\hat{Y}_{n+1} = \beta_0 + \beta_1 \cdot Y_n$$

- 2-Schritt Prognose: $\hat{Y}_{n+2}$

$$Y_{n+2} = \beta_0 + \beta_1 \cdot Y_{n+1} + \epsilon_{n+2}$$

Die unbekannten Werte werden wieder durch Erwartungen ersetzt:

$$\begin{aligned} \hat{Y}_{n+2} &= \beta_0 + \beta_1 \cdot \hat{Y}_{n+1} + 0 \\ &= \beta_0 + \beta_1 \cdot (\beta_0 + \beta_1 \cdot Y_n) \\ &= \beta_0 \cdot (1 + \beta_1) + \beta_1^2 \cdot Y_n \end{aligned}$$

Prognose eines AR(1)

- k-Schritt Progose: $\hat{Y}_{n+k}$
Analog erhalten wir durch wiederholtes Einsetzen von $\hat{Y}_{n+j}$, $j = 1, \ldots, k-1$,

$$\hat{Y}_{n+k} = \beta_0 \cdot (1 + \beta_1 + \ldots + \beta_1^{k-1}) + \beta_1^k \cdot Y_n$$

Da hier $-1 < \beta_1 < 1$ gilt, (der Fall $\beta_1 = 1$ wird gesondert behandelt), verschwindet der Einfluss der letzten Beobachtung Y_n mit steigendem k.

Beispiel: Prognose eines AR(1)

Gegeben sei ein AR(1):

$$Y_t = -0.0007 - 0.10 \cdot Y_{t-1} + \epsilon_t$$

Der Wert des letzten Beobachtungszeitpunkts beträgt $Y_n = 0.125$.

Berechnen Sie die 2-Schritt Prognose.

Beispiel: Prognose eines AR(1)

Gegeben sei ein AR(1):

$$Y_t = -0.0007 - 0.10 \cdot Y_{t-1} + \epsilon_t$$

Der Wert des letzten Beobachtungszeitpunkts beträgt $Y_n = 0.125$.

Berechnen Sie die 2-Schritt Prognose.

Die 1-Schritt Prognose:

$$\hat{Y}_{n+1} = -0.0007 - 0.10 \cdot 0.125 = -0.0132$$

Die 2-Schritt Prognose:

$$\hat{Y}_{n+2} = -0.0007 - 0.10 \cdot \hat{Y}_{n+1} = 0.00062$$

Prognose der Preise

Das Bildungsgesetz für die logarithmierte Reihe der Preise $p_t = \log(P_t)$ lautet

$$p_t = p_{t-1} + r_t, \qquad t = 1, \ldots, n$$

- Die 1-Schritt Prognose ausgehend vom Zeitpunkt n ist daher

$$\hat{p}_{n+1} = p_n + \hat{r}_{n+1}$$

  $\hat{r}_{n+1}$ ist die 1-Schritt Prognose für die Renditen r_{n+1}, bspw. aus einem AR(1).

- Für die Prognose von $p_{n+2} = p_{n+1} + r_{n+2}$ ersetzen wir alle Variablen durch ihre Prognosen.

$$\hat{p}_{n+2} = \hat{p}_{n+1} + \hat{r}_{n+2}$$

  Die Prognose $\hat{p}_{n+1}$ ist bereits bekannt. Eingesetzt ergibt das

$$\hat{p}_{n+2} = (p_n + \hat{r}_{n+1}) + \hat{r}_{n+2} = p_n + (\hat{r}_{n+1} + \hat{r}_{n+2})$$

Prognose der Preise

- Die k-Schritt Prognose ausgehend von n ergibt sich daher als

$$\hat{p}_{n+k} = p_n + (\hat{r}_{n+1} + \hat{r}_{n+2} + \ldots + \hat{r}_{n+k})$$

Die Prognose für p_{n+k} ist der letzte Wert p_n plus der Summe aller prognostizierten Renditen, $\hat{r}_{n+1}, \ldots, \hat{r}_{n+k}$.

Die Prognose für P_{n+k}, dem Niveau der Reihe in $(n+k)$, ergibt sich einfach durch $\exp(\hat{p}_{n+k})$.

$$\hat{P}_{n+k} = \exp(\hat{p}_{n+k})$$

In prognostizierten Renditen ausgedrückt:

$$\begin{aligned}\hat{P}_{n+k} &= \exp(p_n) \cdot \exp(\hat{r}_{n+1}) \cdot \ldots \cdot \exp(\hat{r}_{n+k}) \\ &= P_n \cdot \prod_{j=1}^{k} \exp(\hat{r}_{n+j}).\end{aligned}$$

Beispiel: Preisprognose

Gegeben sind wöchentliche Daten der Reihe des Dow-Jones-Industrial Index, P_t, und dessen Renditen, r_t, für die Periode 11 Jan 1999 bis 28 Jun 1999. Es wurden bereits die 1-, 2- und 3-Schritt Prognosen der Renditen berechnet.

	Beobachtungsperiode			Prognoseperiode		
Datum	...	21 Jun	28 Jun	06 Jul	12 Jul	19 Jul
Preise	...	10552.6	11139.2			
Renditen	...	-0.028	0.054	-0.010	0.015	0.005

Berechnen Sie die Prognose des Dow-Jones-Industrial Index für 19 Jul, $\hat{P}_{n+3}$.

Beispiel: Preisprognose

Die Summe der Prognosen der Renditen ist

$$\begin{aligned}\sum_{j=1}^{3} \hat{r}_{n+j} &= \hat{r}_{n+1} + \hat{r}_{n+2} + \hat{r}_{n+3} = \\ &= -0.010 + 0.015 + 0.005 = 0.010.\end{aligned}$$

Damit ist

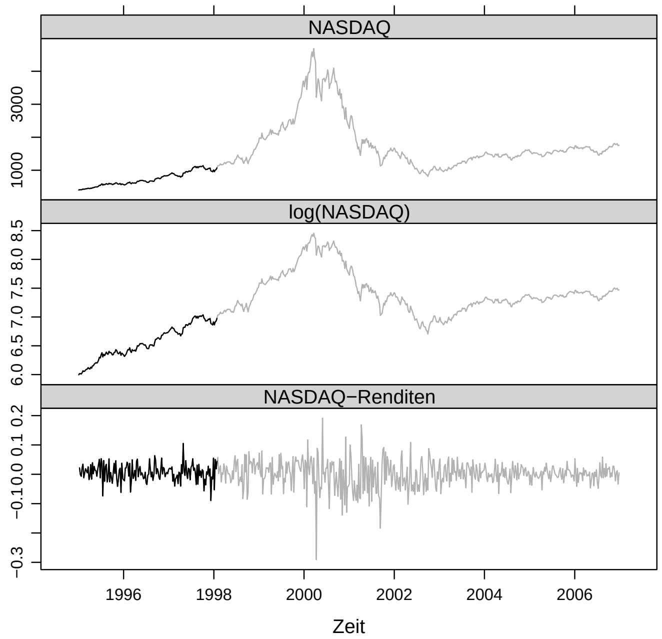
$$\begin{aligned}\hat{P}_{n+3} &= P_n \cdot \exp\left(\sum_{j=1}^{3} \hat{r}_{n+j}\right) \\ &= 11139.2 \cdot \exp(0.010) = 11251.200.\end{aligned}$$

NASDAQ-100

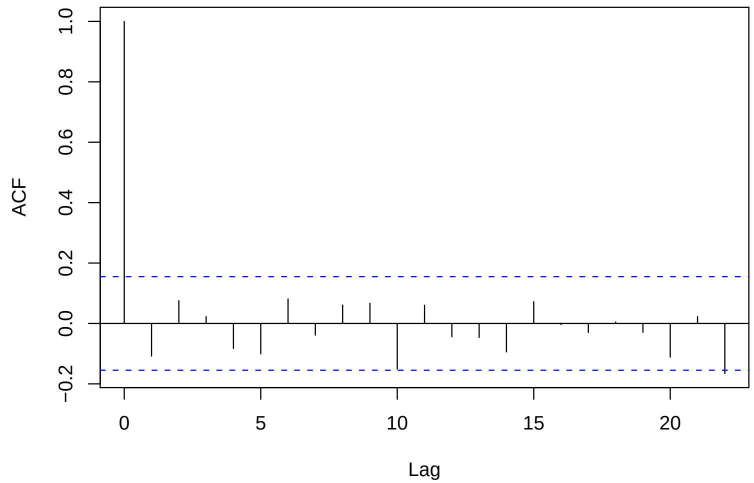
Angenommen wir befinden uns im Januar 1998 und fragen uns, welche Entwicklung des NASDAQ-100 in den nächsten Monaten zu erwarten ist. Wir verwenden dafür die beobachtete Zeitreihe (Januar 1995 bis Januar 1998) und gehen folgendermaßen vor:

- Visualisierung der Zeitreihe (Preise, und zugehörige log-Preise und Renditen),
- Berechnung und Visualisierung des Korrelogramms (der Renditen),
- Schätzung und Vergleich eines AR(1)- und AR(0)-Modells (für die Renditen),
- Prognose mit Hilfe des gewählten Modells für Februar 1998 bis Dezember 2001.

NASDAQ-100



NASDAQ-100



NASDAQ-100

AR(1)-Modell: $r_t \sim r_{t-1}$

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.00667	0.00235	2.85	0.005
lag(rndx, -1)	-0.10895	0.07971	-1.37	0.174

Der Autokorrelationskoeffizient (der Ordnung 1) beträgt -0.109 und ist nicht signifikant von 0 verschieden.

Die verzögerten Beobachtungen r_{t-1} tragen also nicht signifikant zur Verbesserung der Prognose bei. Wir passen daher das einfachere AR(0)-Modell an.

NASDAQ-100

AR(0)-Modell: $r_t \sim 1$

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.00603	0.00230	2.62	0.0097

Die mittlere Wochenrendite ist also $0.006 = 0.6\%$ und signifikant größer als 0.

Die zugehörige mittlere Jahresrendite ist entsprechend $52 \cdot 0.00603 = 31.4\%$.

NASDAQ-100

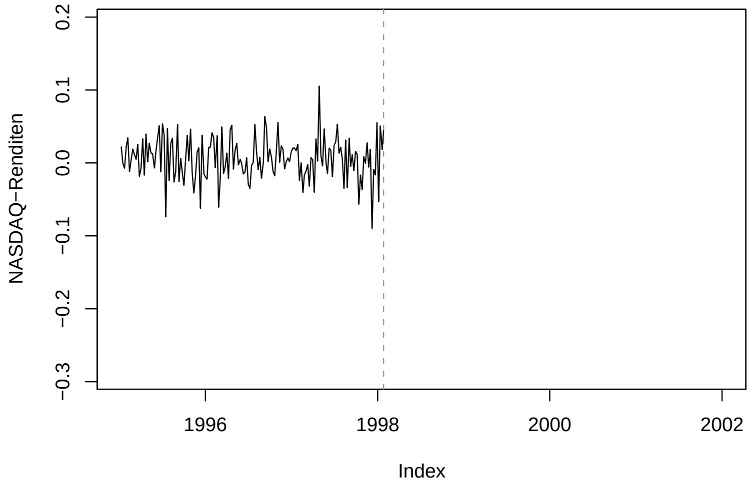
Erinnerung: Da sich die beiden Modelle nur in einem Koeffizienten (nämlich der Autokorrelation der Ordnung 1) unterscheiden, genügte hier der T-Test dieses Koeffizienten zum Vergleich der beiden Modelle.

Völlig äquivalent hätte man den Modellvergleich auch über die zugehörige ANOVA durchführen können.

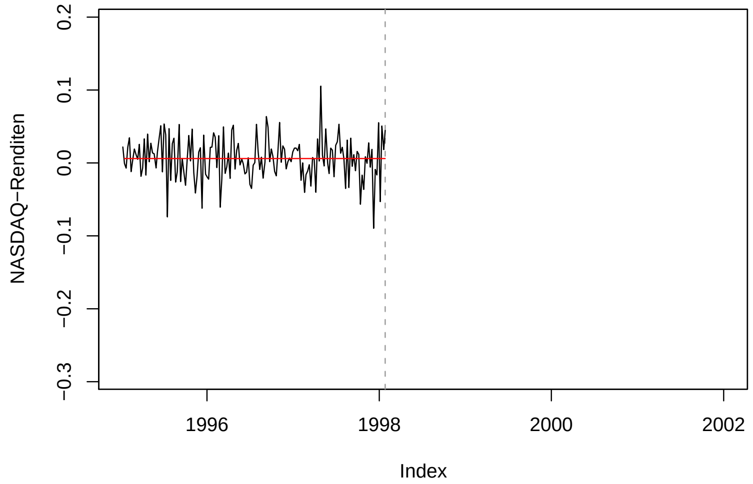
```
Analysis of Variance Table

Model 1: rndx ~ 1
Model 2: rndx ~ lag(rndx, -1)
  Res.Df      RSS  Df Sum of Sq      F Pr(>F)
1    158 0.133394
2    157 0.131825   1  0.001569 1.8683 0.1736
```

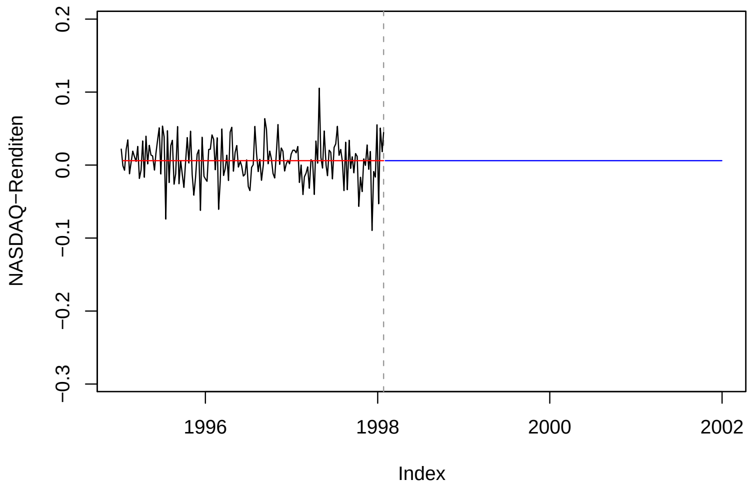
NASDAQ-100



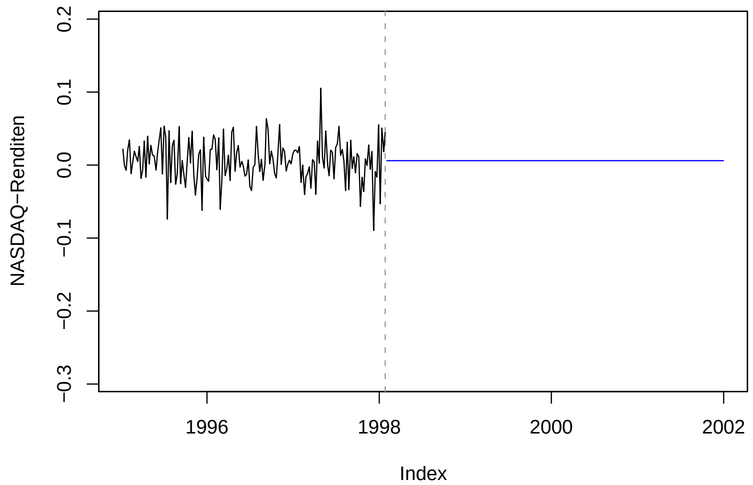
NASDAQ-100



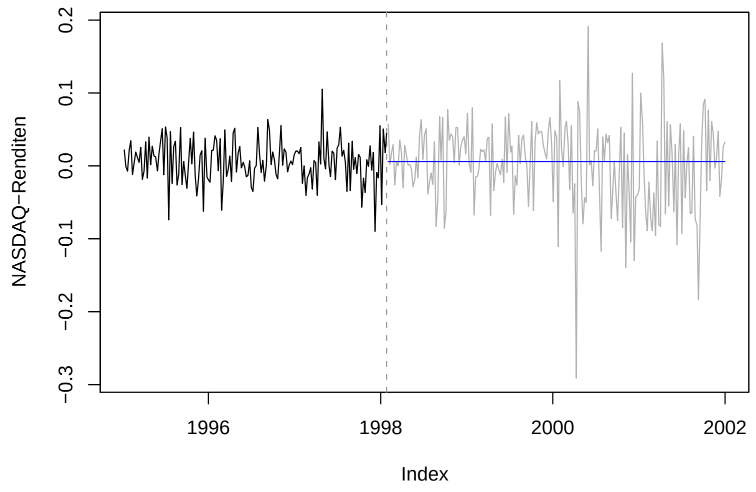
NASDAQ-100



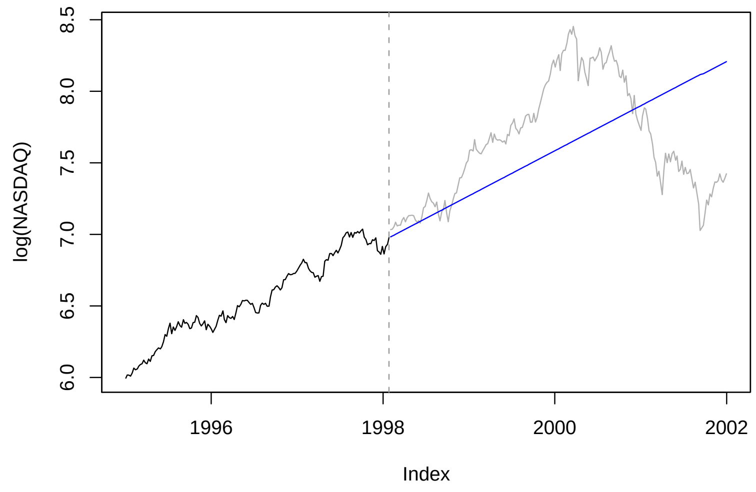
NASDAQ-100



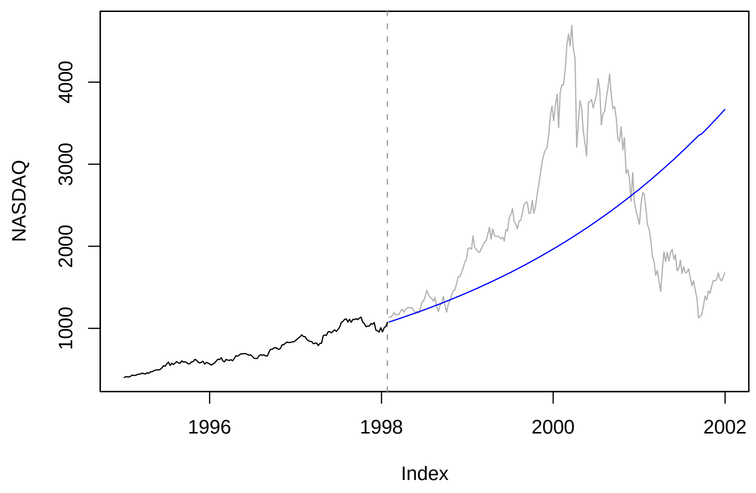
NASDAQ-100



NASDAQ-100



NASDAQ-100



NASDAQ-100

Die Prognose der mittleren Rendite aus dem AR(0)-Modell bildet die Entwicklung der Rendite im Jahr 1998 zunächst noch gut ab. Ab Ende 1999 wird die Rendite dann systematisch unterschätzt und ab März 2000 systematisch überschätzt. Dies ist noch deutlicher auf der Skala der Preise bzw. log-Preise zu sehen.

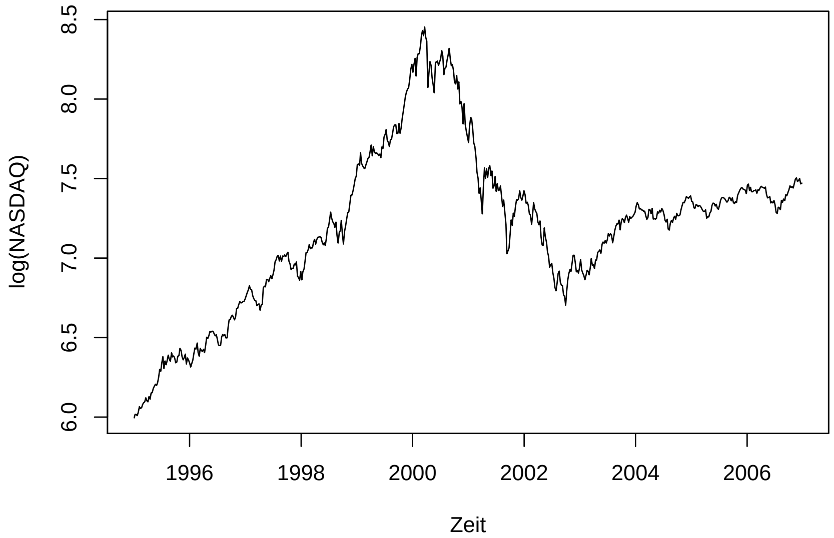
Der offensichtliche Grund ist, dass die Entwicklung des NASDAQ-100 nicht durch ein stabiles Modell erklärt werden kann. Sowohl die Boom-Phase der .com-Blase als auch ihr Zerplatzen konnte nicht aus den vergangenen Renditen erwartet werden.

NASDAQ-100

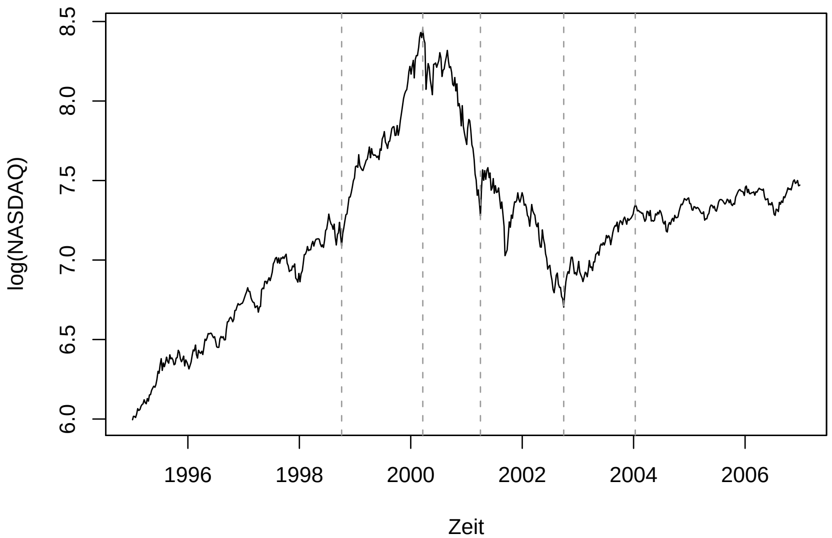
Rückblickend, d.h. hier unter Verwendung der Daten bis Ende 2006, können solche Strukturveränderungen modelliert werden, indem man unterschiedliche AR(0)-Modelle für unterschiedliche Zeitperioden anpasst.

Die Bruchpunkte können mit Hilfe von KQ-Verfahren (die weit über den Inhalt der LV hinausgehen würden) datengetrieben bestimmt werden.

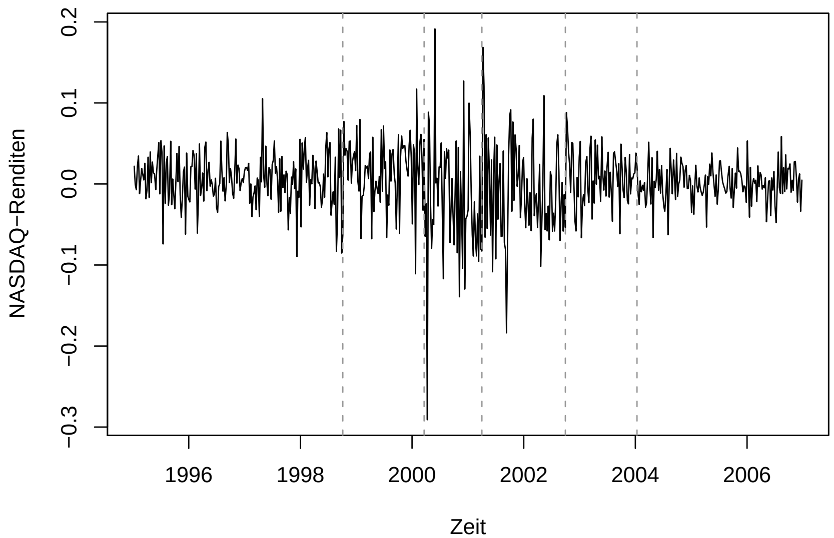
NASDAQ-100



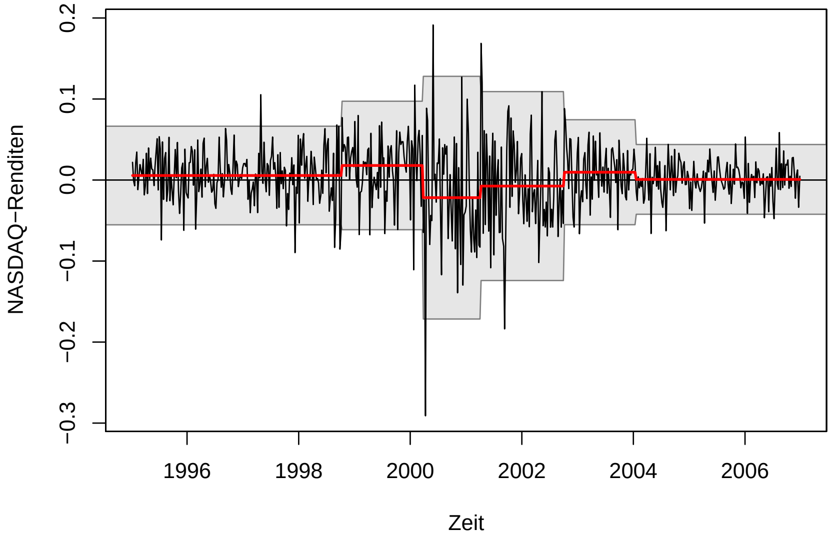
NASDAQ-100



NASDAQ-100



NASDAQ-100



NASDAQ-100

Die mittlere Jahresrendite und ihre zugehörgige Standardabweichung sind (in Prozent):

Anfang	Ende	$\bar{r}$	$s_{r,n-1}$
1995-01-09	1998-10-05	29.0	161.4
1998-10-12	2000-03-20	93.4	210.4
2000-03-27	2001-04-02	−113.2	397.3
2001-04-09	2002-09-30	−38.8	309.3
2002-10-07	2004-01-12	50.0	171.9
2004-01-20	2006-12-26	4.2	114.2

Inhalt

- Notation
- Zusammenhang zwischen Aktienkursen und Renditen
- Effiziente Markthypothese
- Autoregressive Modelle der Ordnung p, AR(p)
- Autokorrelationskoeffizient, Korrelogramm
- Stationäre Prozesse, white noise, AR(p)
- Nicht-stationäre Prozesse, random walk
- Modellierung und Prognose nicht-stationärer Prozesse