Apl. Prof. Dr. Reinhold Kosfeld
Fachbereich Wirtschaftswissenschaften

Aufgaben zur Zeitreihenanalyse (Kap. 5)

Aufgabe 5.1

Welches Phänomen lässt sich mit ARCH-Prozessen modellieren und welche prognostische Relevanz besitzen sie?

Lösung:

Vor allem bei Finanzmarktdaten und Preisreihen erhält man häufig nach Bildung absoluter oder logarithmischer Differenzen aus den Niveaudaten unmittelbar oder nach Berücksichtigung einer ARMA-Struktur Autokorrelationen und partielle Autokorrelationen, die nur noch zufällig von null abweichen. Man betrachtet die differenzierten Reihen aus diesem Grund als Realisationen eines White-Noise-Prozesses.

Gleichzeitig zeigt sich aber oft für die quadrierten (und absoluten) Differenzen eine deutliche Autokorrelationsstruktur. Sie lässt sich auf eine Tendenz zeitlich konzentriert auftretender großer positiver und großer negativer Differenzen der Finanzmarktreihen zurückführen. Absolut große Ausschläge mit wechselndem Vorzeichen kommen durch Unsicherheiten über die Entwicklung auf den Finanzmärkten in bestimmten Phasen zustande. Sie werden durch Schocks (z.B. Immobilienmarktkrise, Fehlspekulationen, Bankenkonkurse) oder neu eintreffende Informationen (z.B. über Übernahmen, Fusionen, Entlassungen, Gewinn- und Verlustmeldungen) hervorgerufen. Man bezeichnet diese Eigenschaft als **Volatilitätsclusterung** (volatility clustering).

Das Phänomen der Volatilitäsclusterung lässt sich mit ARCH-Prozessen modellieren, die die Veränderlichkeit des Risikos in unterschiedlichen Finanzmarktphasen durch bedingte heteroskedastische Varianzen erfassen. Die aktuelle bedingte Varianz hängt von der Volatilität - gemessen durch die quadrierten Störgrößen - in der vorangegangenen und ggf. weiter zurückliegenden Perioden ab. Mit den zeitabhängigen Varianzen lässt sich das Prognoseintervall und damit das Ausmaß der Kursänderungen genauer abschätzen.

Aufgabe 5.2

Die Korrelogramme der Siemens-Aktie weisen im Zeitraum 2. Mai 1996 – 22.1.1999 (n = 712) folgende Autokorrelationen (ACF) und partiellen Autokorrelationen (PACF) für die Tageskurse (Y_t) und ihre Differenzen ΔY_t aus:

	Zeitreihe (Y_t)		Zeitreihe (ΔY_t)	
Lag	ACF von Y_t	PACF von Y_t	ACF von ΔY_t	PACF von ΔY_t
1	0.989	0.989	0.051	0.051
2	0.978	-0.055	-0.009	-0.011
3	0.966	-0.003	-0.007	-0.006
4	0.955	0.020	-0.008	-0.008
5	0.945	0.019	-0.025	-0.024
6	0.934	0.001	-0.023	-0.021
7	0.924	0.017	-0.009	-0.007
8	0.914	-0.009	0.037	0.037
9	0.904	-0.040	-0.012	-0.017
10	0.893	0.006	0.038	0.039

a) Geben Sie für die beiden Zeitreihen (Y_t) und (ΔY_t) auf der Grundlage des approximativen 95%-Konfidenzintervall unter Verwendung des einheitlichen Standardfehlers $\sqrt{1/n}$ die signifikanten Autokorrelationen und partiellen Autokorrelationen an?

Lösung:

Approximatives 95%-Konfidenzintervall:

$$\begin{aligned}\left[-2\cdot\sqrt{1/n};\ 2\cdot\sqrt{1/n}\right] &= \left[-2\cdot\sqrt{1/712};\ 2\cdot\sqrt{1/712}\right]\\ &= \left[-2\cdot 0{,}03748;\ 2\cdot 0{,}03748\right]\\ &= \left[-0{,}075;\ 0{,}075\right]\end{aligned}$$

Signifikante Autokorrelationen und partielle Autokorrelationen ($\hat{=}$ $\hat{\rho}_\tau$ und $\hat{\phi}_{\tau\tau}$ außerhalb des Konfidenzintervalls)			
ACF von Y_t	PACF von Y_t	ACF von ΔY_t	PACF von ΔY_t
alle 10 Autokorrelationen	$\hat{\phi}_{11} = 0{,}989$	-	-

b) Warum lässt sich aus den in Teil a) wiedergegebenen Verläufen der beiden ACFs und PACFs auf einen Random Walk (stochastischer Trend) und nicht auf einen linearen Trend (deterministischer Trend) schließen?

Lösung:

Bei einem deterministischen und stochastischen Trend ist eine langsam fallende ACF der Kursreihe (Y_t) mit lange signifikant bleibenden Autokorrelationen zu erwarten. Da in beiden Fällen der Trend durch Differenzenbildung eliminiert wird, müssen die Autokorrelationen der differenzierten Reihe (ΔY_t) verschwinden, wenn die Ursprungsreihe außer einem Trend keine Struktur mehr enthält.

Während die ACF bei einem deterministischen und stochastischen Trend den gleichen Verlauf hat, gilt dies nicht für die PACF. Bei einem Random Walk muss die PACF nach einem Lag von 1 abbrechen, da es sich hierbei um den Extremfall eine AR(1)-Prozesses handelt, bei der der autoregressive Parameter ϕ_1 gegen 1 geht:

(1) $$Y_t = (\phi_1 = 1) \cdot Y_{t-1} + U_t \qquad \text{(Random Walk).}$$

Die differenzierte Reihe,

(2) $$\Delta Y_t = Y_t - Y_{t-1} = U_t \qquad \text{(White-Noise-Prozess)}$$

entspricht der White-Noise-Zufallsvariablen U_t, deren theoretische partielle Autokorrelationen alle gleich null sind.

Bei einem linearen Trend (deterministischer Trend) unterscheiden sich zwar nicht die ACFs von denen eines Random Walks (stochastischer Trend), wohl aber die PACFs. Hier soll allein die PACF der differenzierten Reihe in Falle eines linearen Trends betrachtet werden.

(3) $$Y_t = \alpha + \beta \cdot t + U_t \qquad \text{(linearer Trend).}$$

Bildet man hier die ersten Differenzen, erhält man

(4) $$\begin{aligned} \Delta Y_t &= Y_t - Y_{t-1} = \alpha + \beta \cdot t + U_t - (\alpha + \beta \cdot (t-1)) + U_{t-1} \\ &= \beta \cdot t + U_t - \beta \cdot (t-1) - U_{t-1} \\ &= \beta + U_t - U_{t-1}. \qquad \text{(MA(1) – Prozess)} \end{aligned}$$

Im Unterschied zu (2) ist (4) kein White-Noise-Prozess, sondern gibt den Extremfall eines MA(1)-Prozesses mit $\theta_1 = 1$ wieder (nicht-invertierbarer MA(1)-Prozess). Im Unterschied zu einem White-Noise-Prozess verschwinden die partiellen Autokorrelationen bei einem MA(1)-Prozess nicht.

Da die differenzierte Kursreihe der Siemens-Aktien keine signifikanten partiellen Autokorrelationen enthält, lässt sie sich als Realisation eines White-Noise-Prozesses und nicht eines MA(1)-Prozesses auffassen. Daher lässt sich aus den Verläufen der ACFs und PACFs der Siemens-Aktie nicht auf einen linearen Trend (deterministischen Trend), sondern auf einen Random Walk (stochastischen Trend) schließen.

Aufgabe 5.3

Für die differenzierten Zeitreihen $(\Delta Y_t)^2$ und $(\Delta \log Y_t)^2$ der Siemens-Aktie ergeben sich im Zeitraum 2. Mai 1996 – 22.1.1999 (n = 712) folgende Korrelogramme:

	Zeitreihe $(\Delta Y_t)^2$		Zeitreihe $(\Delta Y_t)^2$	
Lag	ACF von $(\Delta Y_t)^2$	PACF von $(\Delta Y_t)^2$	ACF von $(\Delta \log Y_t)^2$	PACF von $(\Delta \log Y_t)^2$
1	0.084	0.084	0.096	0.096
2	0.034	0.027	0.044	0.035
3	0.077	0.072	0.057	0.050
4	0.049	0.036	0.053	0.042
5	0.036	0.026	0.028	0.016
6	0.027	0.015	0.031	0.022
7	0.082	0.072	0.072	0.062
8	0.067	0.049	0.076	0.060
9	0.005	-0.012	0.016	-0.004
10	0.054	0.040	0.085	0.073

a) Was geben die differenzierten Zeitreihen $(\Delta Y_t)^2$ und $(\Delta \log Y_t)^2$ inhaltlich wieder?

Lösung:

Zeitreihe $(\Delta Y_t)^2$: quadrierte Kursdifferenzen der Siemens-Aktie

Zeitreihe $(\Delta \log Y_t)^2$: quadrierte Renditen ($\hat{=}$ quadrierte relative Kursveränderungen) der Siemens-Aktie

b) Geben Sie für die beiden Zeitreihen $(\Delta Y_t)^2$ und $(\Delta \log Y_t)^2$ auf der Grundlage des in Aufg. 5.2a) bestimmten approximativen 95%-Konfidenzintervalls die signifikanten Autokorrelationen und partiellen Autokorrelationen an?

Lösung:

Approximatives 95%-Konfidenzintervall: $\left[-2 \cdot \sqrt{1/n};\ 2 \cdot \sqrt{1/n}\right] = \left[-0{,}075;\ 0{,}075\right]$

Signifikante Autokorrelationen und partielle Autokorrelationen ($\hat{=}$ $\hat{\rho}_\tau$ und $\hat{\phi}_{\tau\tau}$ außerhalb des Konfidenzintervalls)			
ACF von $(\Delta Y_t)^2$	PACF von $(\Delta Y_t)^2$	ACF $(\Delta \log Y_t)^2$	PACF von $(\Delta \log Y_t)^2$
$\hat{\rho}_1 = 0{,}084$	$\hat{\phi}_{11} = 0{,}084$	$\hat{\rho}_1 = 0{,}096$	$\hat{\phi}_{11} = 0{,}096$
$\hat{\rho}_3 = 0{,}077$		$\hat{\rho}_8 = 0{,}076$	
$\hat{\rho}_7 = 0{,}082$		$\hat{\rho}_{10} = 0{,}085$	

c) Welche Konsequenzen ergeben sich aus den Unterschieden der Korrelogramme der differenzierten Zeitreihe (ΔY_t) und der differenzierten Zeitreihe $(\Delta Y_t)^2$ bzw. $(\Delta \log Y_t)^2$ für die Modellierung und Prognose der Tageskurse der Siemens-Aktie?

Lösung:

Da in dem Korrelogramm der Zeitreihe der ersten Differenzen, (ΔY_t), keine Struktur mehr vorhanden ist, sind die Kursveränderungen der Siemens-Aktie unkorreliert. Das bedeutet, dass das beste lineare Zeitreihenmodell für die Kurse der Siemens-Aktie der Random Walk ist. Ein ARIMA-Modell kann keine zusätzliche Erklärung der Kursentwicklung leisten.

Angesichts der vorhandenen Autokorrelationsstrukturen in den Zeitreihen der quadrierten Differenzen, $(\Delta Y_t)^2$, und quadrierten logarithmierten Differenzen, $(\Delta \log Y_t)^2$, sind die Kursveränderungen und Aktienrenditen jedoch nicht unabhängig voneinander. Aufgrund der zeitlichen Abhängigkeit der quadrierten Differenzen lässt sich durch ein nichtlineares Zeitreihenmodells der Erklärungsgehalt eines Random Walks verbessern. Da die logarithmierten quadrierten Differenzen als Residuenquadrate des Random Walks interpretiert werden können, ist insbesondere eine zeitvariablen Modellierung der bedingten Varianz der Zeitreihe nahe liegend. Hierzu ist die Klasse der ARCH- und GARCH-Modelle entwickelt worden.

Interessiert man sich stärker für die Renditen statt der Kursveränderungen einer Aktie, wird die bedingte Varianz in Abhängigkeit von den logarithmierten quadrierten Differenzen modelliert. In diesem Fall werden die erklärenden Größen der ARCH-Modelle als Residuen eines Random Walks in den logarithmierten Aktienkursen betrachtet.

Mit Hilfe der ARCH- und GARCH-Modelle lässt sich das in der Aktienkursentwicklung vorhandene Risiko zeitvariabel messen. Außerdem lässt sich mit dieser Modellklasse das Prognoseintervall genauer abschätzen als bei Annahme einer konstanten Varianz. Die zeitvariablen Varianzen geben Auskunft über das Ausmaß von Kursveränderungen und enthalten damit Informationen zur Verbesserung der Intervallprognose.

Aufgabe 5.4

Gegeben ist der EViews-Output des Lagrange-Multiplier-Tests auf ARCH-Effekte in den Renditen der Siemens-Aktie im Zeitraum 2. Mai 1996 – 22.1.1999:

ARCH Test:

F-statistic	6.682217	Probability	0.009937
Obs*R-squared	6.638439	Probability	0.009980

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Sample (adjusted): 5/06/1996 1/22/1999
Included observations: 710 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000422	5.52E-05	7.646919	0.0000
RESID^2(-1)	0.096701	0.037408	2.584999	0.0099

R-squared	0.009350	Mean dependent var	0.000467
Adjusted R-squared	0.007951	S.D. dependent var	0.001400
S.E. of regression	0.001394	Akaike info criterion	-10.31003
Sum squared resid	0.001376	Schwarz criterion	-10.29717
Log likelihood	3662.059	F-statistic	6.682217
Durbin-Watson stat	2.006670	Prob(F-statistic)	0.009937

a) Testen Sie die Renditen der Siemens-Aktie unter Verwendung des kritischen Werts und des p-Werts auf ARCH(1)-Effekte ($\alpha = 0{,}05$)!

Lösung:

Lagrange-Multiplier-Test (LM-Test) auf ARCH-Effekte

Signifikanzniveau: $\alpha = 0{,}05$

Nullhypothese H_0: $\alpha_1 = 0$ (keine zeitvariablen Varianzen)

● **Test mit kritischem Wert**

- Prüfgrüße: $LM(ARCH(1)) = n \cdot R^2 = 710 \cdot 0{,}009350 = 6{,}6385$

R^2: Determinationskoeffizient der Testgleichung (Hilfsregression)
n: Stichprobenumfang (nach Adjustierung)

Der berechnete Wert der Prüfgröße entspricht – bis auf die Rundung - dem von EViews unter Obs*R-squared ausgewiesenen Wert.

- Kritischer Wert ($\alpha = 0{,}05$): $\chi^2_{1;0,95} = 3{,}841$

- Testentscheidung:

LM(ARCH(1)) = 6.6385 > $\chi^2_{1;0,95}$ = 3,841 => H_0 ablehnen

- Interpretation:

Die Nullhypothese einer konstanten bedingten Varianz der Renditen der Siemens-Aktie wird durch den LM-Test für ARCH-Effekte bei einem Lag von 1 verworfen. Die Alternativhypothese einer bedingt heteroskedastischen Störvarianz ist bei der Wahl eines Lags von 1 auf dem 5%-Niveau statistisch gesichert.

● **Test mit p-Wert**

- p-Wert: p = 0,009980

Der p-Wert des LM-Tests auf ARCH(1)-Effekte wird von EViews unter Probability ausgewiesen. Er gibt die Wahrscheinlichkeit dafür an, dass der Wert der Prüfgröße LM(ARCH) von 6,6385 (=Obs*R-squared) bei Gültigkeit der Nullhypothese überschritten wird (=tatsächliches Signifikanzniveau).

- Signifikanzniveau: $\alpha = 0,05$

α ist das nominale Signifikanzniveau, das vorgegeben werden muss. Es gibt an, wie hoch die Wahrscheinlichkeit bei dem Test maximal sein kann, die Nullhypothese zu verwerfen, wenn sie tatsächlich richtig ist (= Irrtumswahrscheinlichkeit).

- Testentscheidung:

$p = 0,009980 < \alpha = 0,05$ => H_0 ablehnen

- Interpretation:

Die Nullhypothese wird beibehalten, solange Werte der χ^2-Verteilung, die mindestens so groß sind wie der Wert der Prüfgröße, eine Wahrscheinlichkeit von mindestens 5% haben bei Gültigkeit der Nullhypothese einzutreten. Haben sie eine Wahrscheinlichkeit, die kleiner als 5% ist, wird ihre Abweichung von Wert 0 nicht mehr als Zufallsabweichung, sondern als substanzielle (= signifikante) Abweichung interpretiert, was in unserem Fall gegeben ist. Die Nullhypothese einer zeitlich konstanten Störvarianz der Renditen der Siemens-Aktie wird daher zugunsten eines zeitvariablen Varianzprozesses mit einem Lag von 1 (ARCH(1)-Prozess) verworfen.

b) Geben Sie die geschätzte Testgleichung an und interpretieren Sie sie!

Lösung:

Die Testgleichung (= Hilfsregression) des LM-Tests auf ARCH(1)-Effekte lautet:

$$\hat{u}_t^2 = 0{,}0004 + 0{,}0967 \cdot \hat{u}_{t-1}^2$$

t-Werte: (7,647) (2,585)

Sie wird mit der gewöhnlichen Methode der kleinsten Quadrate (OLS-Methode) geschätzt und gibt die Abhängigkeit der aktuellen quadrierten Residuen von den um eine Periode verzögerten Residuenquadrate an. Die Residuen sind hier die Renditen (= Residuen eines Random Walks der logarithmierten Tageskurse) in Abweichung von ihrem Mittelwert.

Aufgrund des hohen Anzahl von Freiheitsgraden (=710–2=708) sind für einzelne Signifikanztests der geschätzten Koeffizienten der Testgleichung die t-Werte mit dem 97,5%-Quantil der Standardnormalverteilung vergleichen. Bei einem Signifikanzniveau von 0,05 ist der kritische Wert $z_{0,975}$ gleich 1,96, so dass sich die beiden geschätzten Koeffizienten $\hat{\alpha}_1$ und $\hat{\alpha}_2$ der Testgleichung als signifikant erweisen.

Aufgabe 5.5

Aufgrund des in Aufg. 5.4a ermittelten Testergebnisses ist für die Renditen der Siemens-Aktie folgende Modellschätzung vorgenommen worden:

Dependent Variable: D(LOG(SIE))
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 5/03/1996 1/22/1999
Included observations: 711 after adjustments
Convergence achieved after 13 iterations
Variance backcast: ON
GARCH = C(2) + C(3)*RESID(-1)^2

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000589	0.000803	0.733698	0.4631
Variance Equation				
C	0.000405	1.09E-05	37.12332	0.0000
RESID(-1)^2	0.139585	0.033968	4.109264	0.0000

R-squared	-0.000045	Mean dependent var	0.000444
Adjusted R-squared	-0.002870	S.D. dependent var	0.021621
S.E. of regression	0.021652	Akaike info criterion	-4.850389
Sum squared resid	0.331904	Schwarz criterion	-4.831120
Log likelihood	1727.313	Durbin-Watson stat	1.912639

a) Geben Sie die beiden geschätzten Regressionsgleichungen für die Renditen unter Einbeziehung des Störterms wieder!

Lösung:

- Mean Equation (1. Regressionsgleichung mit der abhängigen Variabler $\Delta\log(Y_t)$)

$$\Delta\log(Y_t) = 0{,}0006 + \varepsilon_t \cdot \hat{\sigma}_t^{1/2}$$

t-Wert: (0,734)

- Variance Equation (2. Regressionsgleichung mit der abhängigen Variablen σ_t^2)

$$\hat{\sigma}_t^2 = 0{,}0004 + 0{,}1396 \cdot U_{t-1}^2$$

t-Werte: (37,123) (4,109)

b) Welches numerisch spezifizierte Zeitreihenmodell wird durch die in Teil a) gegebenen Regressionsgleichungen für die logarithmierte Zeitreihe der Tageskurse der Siemens-Aktie impliziert?

Lösung:

- Mean Equation (1. Regressionsgleichung mit der abhängigen Variabler $\log(Y_t)$)

$$\log(Y_t) = \log(Y_t) + 0{,}0006 + \varepsilon_t \cdot \hat{\sigma}_t^{1/2} \qquad \text{(Random Walk mit ARCH-Effekt))}$$

t-Wert: (0,734)

- Variance Equation (2. Regressionsgleichung mit der abhängigen Variablen σ_t^2)

$$\hat{\sigma}_t^2 = 0{,}0004 + 0{,}1396 \cdot U_{t-1}^2$$

t-Werte: (37,123) (4,109)

c) Aus welchen Gründen kann sich die „Variance Equation“ von der in Aufg. 5.4 ausgewiesenen „Test Equation“ unterscheiden?

Lösung:

Die beiden Gleichungen werden mit unterschiedlichen Methoden geschätzt. Zur Schätzung der „Test Equation“ wird die gewöhnliche Methode der kleinsten Quadrate (OLS-Methode) eingesetzt, wobei als abhängige Variable die quadrierten Residuen, $\hat{u}_t^2$, verwendet werden. Die „Variance Equation“ wird dagegen mit der Maximum-Likelihood-Methode (ML-Methode) geschätzt. Hierbei wird eine iterative Schätzung der zeitvariablen Varianz, σ_t^2, als abhängige Variable vorgenommen. In jeder Iteration werden verbesserte Schätzer der Regressionskoeffizienten ermittelt. Das Verfahren wird beendet, wenn das Konvergenzkriterium erfüllt ist, d.h. wenn die aufeinander folgender Parameterschätzer um weniger als eine vorgegebene Konstante δ (z.B. $\delta = 0{,}0001$) voneinander abweichen.

Aufgabe 5.6

Im Falle von ARCH-Effekten erfolgt die Schätzung des ökonometrischen Eingleichungsmodells unter Anwendung der Maximum-Likelihood (ML)-Methode! Erläutern Sie das Schätzprinzip der ML-Methode!

Lösung:

Die gemeinsame Wahrscheinlichkeits- bzw. Dichtefunktion ist im Falle unabhängiger Zufallsvariablen $X_1, X_2, \ldots, X_n$ gleich dem Produkt der einzelnen Wahrscheinlichkeits- bzw. Dichtefunktionen:

$$f(x_1, x_2, \ldots, x_n \mid \boldsymbol{\theta}) = f_1(x_1 \mid \boldsymbol{\theta}) \cdot f_2(x_1 \mid \boldsymbol{\theta}) \cdot \ldots \cdot f_n(x_1 \mid \boldsymbol{\theta})$$

Wenn die Parameter der Wahrscheinlichkeitsverteilung $\theta_1, \theta_2, \ldots, \theta_k$ – angeordnet in dem Vektor θ – gegeben sind, lässt sich die Wahrscheinlichkeit für eine Realisation beliebiger x-Werte bestimmen. Dies ist die Vorgehensweise in der Wahrscheinlichkeitsrechnung.

In der induktiven Statistik ist dagegen eine andere Situation gegeben. Hier geht es um eine Schätzung der unbekannten Parameter $\theta_1, \theta_2, \ldots, \theta_k$ aus einer Zufallsstichprobe $X_1, X_2, \ldots, X_n$. Nach Ziehen der Stichprobe liegen die x-Werte $x_1, x_2, \ldots, x_n$ vor, d.h. sie sind fest und können nicht mehr variiert werden. Unbekannt bleiben dagegen die Parameter $\theta_1, \theta_2, \ldots, \theta_k$ der Grundgesamtheit, die durch die Stichprobe geschätzt werden sollen. Gedanklich könnte man die Parameter $\theta_1, \theta_2, \ldots, \theta_k$ variieren und fragen, welche Werte sie annehmen müssten, damit die gezogene Stichprobe $x_1, x_2, \ldots, x_n$ die größte Wahrscheinlichkeit bzw. Dichte hat. Um diese gegenüber der Wahrscheinlichkeitsrechnung umgekehrte Situation variabler Parameter herauszustellen, spricht man von einer **Likelihood-Funktion**:

$$L(\boldsymbol{\theta} \mid x_1, x_2, \ldots, x_n) = f_1(x_1 \mid \boldsymbol{\theta}) \cdot f_2(x_1 \mid \boldsymbol{\theta}) \cdot \ldots \cdot f_n(x_1 \mid \boldsymbol{\theta}).$$

Für gleiche Parameterwerte $\boldsymbol{\theta}$ und gleiche x-Werte nehmen die Likelihood-Funktion L und die gemeinsame Wahrscheinlichkeits- und Dichtefunktion f identische Werte an. Während bei der Wahrscheinlichkeits- und Dichtefunktion jedoch die Parameterwerte $\boldsymbol{\theta}$ konstant und die x-Werte variabel sind, liegen bei der Likelihood-Funktion die x-Werte aufgrund der gegebenen Stichprobe fest, wohingegen die unbekannten Parameterwerte $\boldsymbol{\theta}$ gesucht und damit als variabel betrachtet werden.

Mit der **Maximum-Likelihood-Methode** (ML-Methode) werden die Parameterwerte $\boldsymbol{\theta}$ so geschätzt, dass unter ihnen die gezogene Stichprobe die größte Wahrscheinlichkeit besitzt. Hat man z.B. für den Parameter θ_1 den Wert 5 gewählt und der Wert 7 würde den Wert der Likelihood-Funktion erhöhen, dann hätte die gegebene Stichprobe unter dem Parameter $\theta_1 = 7$ eine größere Wahrscheinlichkeit realisiert zu werden. Man sucht genau nach denjenigen Parametern, die zu einer maximalen Wahrscheinlichkeit bzw. Dichte der gegebenen Stichprobe führen. Hierzu ist die Likelihood-Funktion unter Anwendung der Differentialrechnung zu maximieren. Aufgrund der besseren analytischen Auswertungsmöglichkeit arbeitet man in der Regel mit der logarithmierten Likelihood-Funktion (log L). Diejenigen Parameterschätzer $\hat{\boldsymbol{\theta}}_{ML}$, die die log Likelihood-Funktion (log L) und damit auch die Likelihood-Funktion (L) maximieren,

$$L(\hat{\boldsymbol{\theta}}_{ML} \mid x_1, x_2, \ldots, x_n) = \underset{\boldsymbol{\theta}}{\text{Max}}\ L(\boldsymbol{\theta} \mid x_1, x_2, \ldots, x_n)$$

heißen **Maximum-Likelihood-Schätzer** (ML-Schätzer).